PROJECT TITLE	Covid-19 disruptions: attainment gaps and primary school responses
EVALUATOR (INSTITUTION)	FFT and TeacherTapp
PRINCIPAL INVESTIGATOR(S)	Becky Allen, Rob Coe, Laura James, Dave Thomson, Ben Weidmann
STUDY PLAN AUTHOR(S)	Rob Coe, Dave Thomson, Ben Weidmann
STUDY DESIGN	Observational study
PUPIL AGE RANGE AND KEY STAGE	Key Stage 2 (7-10)
NUMBER OF SCHOOLS	140
NUMBER OF PUPILS	c20,000
PRIMARY OUTCOME MEASURE AND SOURCE	PIRA, PUMA, NTS (Rising Stars; Hodder)
SECONDARY OUTCOME MEASURE AND SOURCE	NA

Study Plan version history

VERSION	DATE	REASON FOR REVISION
1.2 [*latest*]	30 June 2021	Updated publication plan, which will have a single addendum (published after summer assessments in 2021) rather than two addenda. This version also adds an additional teacher survey to capture school responses to the second period of partial closures, as well as research questions 2c and 3c.
1.0 [*original*]	2nd Dec 2020	*[leave blank for the original version]*

Table of contents

Background and study rationale

In March 2020 schools in England were closed to the majority of pupils because of the Covid-19 pandemic. Pupils' learning was supported remotely, with different schools adopting a range of different approaches. From September 2020, most pupils were back in school, though disruption and periods of absence or remote learning continued for many.

This study primarily aims to estimate how the attainment gap between disadvantaged (Pupil Premium) pupils and their peers changed between November 2019 and December 2020, during different periods of remote and face-to-face schooling and with different approaches to each in different schools. We plan to use a sample of primary schools in England that were already using Rising Stars assessments to monitor their pupils' progress, and to complement these with additional assessments and teacher surveys about their school's approach to remote learning. An extension to the study may use later assessments in the spring and summer of 2021 to track subsequent changes in the gap.

The study will provide direct empirical evidence about the impact of the pandemic on the disadvantage attainment gap and on the approaches to remote learning that were associated with mitigating that impact. We hope it may be possible to translate the findings into advice for schools about how best to support their disadvantaged pupils in any future periods of remote or disrupted learning.

The study is primarily motivated by four factors. First, existing evidence is a poor guide to understanding the impact of covid-related disruptions on attainment gaps. Much of the quantitative evidence we have relies on studies of 'summer learning loss', as illustrated by the by the EEF rapid evidence assessment (EEF, 2020a). The 'summer learning loss' literature examines the impact on attainment gaps of students taking planned holidays between school years. As noted in the review, missing school due to summer holidays may not be a good analogy for the experience of children in England during covid-related disruptions. Moreover, existing empirical scholarship has other limitations: there are relatively few studies; evidence mostly comes from the US; and estimates are consistent with a wide range of impacts on the FSM learning gap.

Second, there is a scarcity of information about the types of supports schools and teachers provided (and are providing) during the period affected by covid. This study hopes to address this gap through a teacher survey conducted by TeacherTapp, which will provide a snapshot of how schools supported students during the disrupted period from March-June. An additional teacher survey will provide data on how schools are responding to challenges of in-person schooling in Autumn 2020.

Third, existing evidence does not provide a strong basis for recommendations to schools about how to mitigate learning loss and gap-widening during closures and partial closure. The EEF's rapid evidence assessment on remote learning provides a helpful starting point (EEF, 2020b). But existing evidence provides at best an analogy to the covid-disruption period, in which distance learning was unplanned, teachers lacked training or experience and access to technology was variable.

Last, partial or intermittent closure of schools may be ongoing, with potentially large variations over time or across regions. Feedback about the costs of covid-related closures, and the effectiveness of different teacher practices, may lead to concrete advice for schools and policymakers.

Quantitative analysis

The study has three aims:

1. Estimate the attainment gap between FSM students and their peers before and after the covid disruption
2. Describe how schools responded to covid-related disruptions and their consequences
3. Explore associations between school responses and FSM attainment gaps

These aims correspond to the following set of research questions.

Research questions

1. Estimate FSM attainment gaps

RQ1a) How did the attainment gap between disadvantaged pupils and their peers change during the first period of (partial) closures due to Covid (Nov 2019-Sep 2020)?

RQ1b) How did the attainment gap between disadvantaged pupils and their peers change after a term of (largely) in-person schooling (Sep-Dec 2020)?

RQ1c) How did the attainment gap between disadvantaged pupils and their peers change from Dec 2020 to March 2021? [to be addressed in an addendum, published after the main report]?

RQ1d) How did the attainment gaps between disadvantaged pupils and their peers change from March 2021 to May 2021? [to be addressed in an addendum, published after the main report]

Note: the different elements of RQ1 will be discussed simultaneously. For example, the study will describe how estimated changes in attainment gaps from Nov 2019-Sep 2020 are associated with subsequent changes in gaps (e.g. from Sep 2020 to Dec 2020).

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

RQ2a) How did schools respond during the first closure period (Mar-July 2020) to support pupils working from home?

RQ2b) How did schools respond during the re-opening period (Sep-Dec 2020) to compensate for lost instruction time during March-July 2020?

RQ2c) How did schools respond during the second closure period (Jan-March 2021) to support pupils working from home?

3. Explore associations between FSM attainment gaps and school responses

RQ3a) What is the association between school responses during the first closure period (Mar-July 2020) and changes in the FSM-attainment gap when children returned to in-person schooling (from Nov-2019 to Sep 2020)?

RQ3b) What is the association between school responses during the re-opening period (Sep-Dec 2020) and changes in the FSM-attainment gap when children returned to in-person schooling (Sep-Dec 2020)?

RQ3c) What is the association between school responses during the second closure period (Jan-March 2021) and changes in the FSM-attainment gap from Dec 2020 to March 2021?

Participants

A convenience sample has been selected for this project.[1] The stages of selection were:

1. Identifying schools which had administered Rising Stars tests in Year Groups 2, 3, 4 or 5 in Autumn 2019 using lists provided by Rising Stars
2. Writing to schools to seek volunteers who agreed to
    a. Administer tests in September 2020 and November/ December 2020
    b. Allow us to collect test data from the test provider
    c. Allow us to collect pupil context data via Aspire Data Exchange

This resulted in a sample of 145 schools. Assuming an average of 44 pupils per year group, this will yield a maximum pupil-level sample size of 25,000 prior to any attrition.

The table below compares participating schools to other state-funded mainstream schools with 2019 Key Stage 2 data. Overall, the recruited sample has similar mean values on observable characteristics. That said, the recruited sample has proportionately fewer schools among in the North of England (North East, North West, Yorkshire and the Humber) and the level of attainment in participating schools tends to be marginally lower.

	Participants	Other Schools
% junior schools	8%	7%
% London	29%	26%
% North	24%	32%
Cohort size	44.6	41.6
KS1 APS	16.2	16.3
% FSM6	31%	30%
% EAL	20%	16%
% expected standard RWM	62%	66%
% higher standard RWM	9%	10%

Outcome measures and other data

The data for the project will be collected directly from schools and the test provider, Rising Stars.

Test data

The main outcome measures will be standardised scores in termly reading and maths tests.

All schools in the sample conducted termly tests in Autumn 2019 using tests provided by RS Assessment, namely PIRA/ PUMA or NTS Assessments (national test style) reading and maths papers.

These are widely used termly tests taken by over 6,000 primary schools in England, typically for pupils in Years 1 to 6. Ordinarily, we would have expected at least a majority of these schools to have used these tests under “business as usual”.

Both types of test provide coverage of the revised national curriculum for each year group in reading and maths. The tests are taken in class towards the end of each term and typically take around 45 to 50 minutes, although schools have the flexibility to decide when to use them. The tests can be taken online although we would expect the majority to be taken using pen and paper. Similar to the arrangements for Key Stage 2 tests, we would expect all pupils in a

[1] Note that schools were offered a financial incentive to participate in the study: 250 pounds to be paid in January, and a further 100 pounds, paid at the conclusion of the study in May/June 2021.

year group to take the test, with the exception of any pupils who were absent or were unable to access the test for another reason. Scripts are marked by teachers using published mark schemes and guidance and the scores for each pupil can be entered into MARK, an online reporting and analysis tool provided by RS Assessment.

PIRA and PUMA tests were standardised using a representative sample of schools in 2014/15 and 2013/14 respectively, including results from at least 1,000 pupils per test. Raw test scores are converted into standardised scores and other age-related outcomes based on these standardisation samples. The technical manuals report reliability coefficients (Cronbach's Alpha) around 0.9 for all the termly tests we plan to use. 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.

Ordinarily, pupils would have taken further termly tests in Spring and Summer 2020. Schools were asked to test pupils using the Summer 2020 papers in September 2020. For example, pupils in Year 4 were asked to take Summer 2020 Year 3 papers. They will then take the suite of termly tests appropriate for each year group as normal.

For the purposes of this project, we will use raw scores and standardised scores for each test. These will be collected through the Aspire Pupil Tracking (APT) system, either directly from MARK (RS Assessment's online analysis tool) or by schools entering the data into APT directly.

We expect the majority of schools to use PIRA and PUMA tests although a small number may use the NTS Assessment tests. A further complication will be the introduction of a new suite of PIRA and PUMA tests in the 2020/21 year, updated and standardised in 2019/20, which some schools may choose to use. This means that we will have to work with different tests, all standardised at different times. Therefore, we will standardise scores across the different tests.

Linear equating is the simplest, and preferred, approach. This method attempts to equate tests by applying a linear transformation to scores from one or both tests, so that the means and standard deviations of the population who have taken both are equal. These 'standardised' scores from different tests may then be treated as interchangeable. However, in cases where relationship between test scores is non-linear, or where score may not be on interval scales, an equipercentile linking approach, based on ranks and smoothing, may be more appropriate (Kolen & Brennan, 2013). We will use this second approach as a robustness check.[2]

Context and Attendance Data

Data on pupil characteristics, such as disadvantage, gender and ethnicity, will be collected from schools via Aspire Data Exchange. The group of pupils classified as disadvantaged for the purposes of the analysis will be identified based on data extracts from October 2020. However, we will separately identify pupils joining the disadvantage group at a later date. This may lead to further exploratory analysis, depending on pupil numbers.

We will also attempt to collect sessional attendance data from September 2020 onwards for participating schools.

Teacher Survey Data

We used Google Forms to collect responses from teachers to an initial set of questions about how they supported their pupils during the initial closure period. This survey took place in Summer 2020 and covered types of work set, feedback given and interactions with pupils and families.

[2] A large dataset of results from scores entered into MARK from over 3,000 schools in the last three years also provides a recent comparison of national attainment in these tests.

The survey was based on questions previously asked in Teacher Tapp. Teacher Tapp is a teacher survey app, used by around 10,000 teachers in England each week. Teachers who use the app are asked 3 daily questions about their practice. On any given day, around 7,000 teachers answer these questions.

A second teacher took place in Autumn 2020 (in December 2020). The primary goal of the second survey was to describe how teachers and schools are supporting students after their return to 'in-person' schooling.

In response to a second, unforeseen period of partial school closures (Jan-March 2021) a third teacher survey was added to this study. This survey mirrored the first survey and asked teachers about the ways in which they supported their pupils while schools were closed.

Published School Level Data

Further relevant school characteristics to control for in models, such as inspection ratings, region, historic attainment and percentage disadvantage, will be sourced from published government datasets.

Analysis overview

Our analysis will be conducted in R, with multilevel models fit using the lme4 package.

Research aim 1: estimating gaps

Let Y_{ijgt}^{k} be a standardized attainment score for pupil i, in school j, for cohort g, at time t on domain k (reading or maths). Scores will be standardized across different tests using two methods i) linear equating; and ii) ranks. Linear equating is our preferred approach. However, as noted below, we will examine whether our results are sensitive to using ranks in robustness checks.

Let F_i be a binary indicator equal to 1 if student i was classified as FSM6 when children returned to school (in Autumn 2020).[3]

Define the following time points:

- T_0 = average date of assessment in autumn 2019
- T_1 = average date of first assessment after schools open (Sep/Oct in 2020)
- T_2 = average date of second assessment in autumn 2020 (Dec 2020)
- T_3 = average date of spring assessment in 2021 (Mar 2021)
- T_4 = average date of summer assessment in 2021 (May 2021)

Define "the attainment gap" to be the difference in Y between the mean attainment of FSM6 pupils and their peers. Let this difference be given by:

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

Our first research question is to estimate G_{tg}^{k} at different points in time.

Let the change in G_{tg}^{k} from Nov 2019 to Sep 2020 be given by:

[3] We expect the proportion of children who are designated FSM6 in October 2020 will be larger than in autumn 2019, due to the coronavirus. This change in the composition of the two groups (FSM6 and non-FSM6) would influence our estimate of how the FSM gap has changed over time. To control for this, our headline results will define FSM6 at a single point in time – October 2020. This will keep the percentage of FSM children constant at different time points and allow 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 – at the start of our study period but unfortunately we do not have access to this data. Note also that we plan on keeping track of FSM status over the course of the 2020-21 year, and will report on how this changes in our sample in the addenda (see Timeline for publishing details).

$$\Delta^{\mathrm{k}}_{T_1 g} = G^{k}_{T_1 g} - G^{k}_{T_0 g} \quad (2)$$

To estimate $\Delta^{\mathrm{k}}_{T_1 g}$, and address RQ1a, we will fit model 1a using R:

$$Y^{k}_{ijgT_1} - Y^{k}_{ijgT_0} = \alpha_j + \delta_g \cdot g + \Delta^{\mathrm{k}}_{T_1 g} F_i \cdot g + s_0(t^0_j) + s(t^1_j) + e_{ijg} \quad (model\ 1a)$$

Where:

- t^0_j is the number of days between T_0 and the administration of the Rising Star assessment in school j in autumn 2019. t^1_j is the number of days between T_1 and the first administration of Rising Star assessments in school j in autumn 2020.
- s_0 and s_1 represent flexible functions (cubic splines) and are included to take account of variations in the timing of tests across schools.
- g is the year level of pupils in 2020/21 (i.e. their cohort)
- $\delta^k_g + \Delta^{\mathrm{k}}_g$ is the average change in Y (from to T_0 to T_1) for FSM kids in grade g in subject $k, \in (maths, English)$
- δ^k_g is the average change in Y (from to T_0 to T_1) for non-FSM kids in grade g in subject k ∈ $(maths, English)$
- α_j is a random effect for school j, to account for clustering of pupils within schools
- e_{ijg} are normally-distributed disturbances at the pupil level

Our analysis for RQ1b will be analogous, but compare different time points:

$$Y^{k}_{ijgT_2} - Y^{k}_{ijgT_1} = \delta_g \cdot g + \Delta^{\mathrm{k}}_g F_i \cdot g + s_2(t^2_j) + s_1(t^1_j) + e_{ijg} \quad (model\ 1b)$$

In model 1b, t^2_j is the number days, for school j, between September 1st and the date of the Rising Stars assessment at the end of the autumn term (Dec 2020). For research questions 1c and 1d, our primary analysis will mirror models 1a and 1b:

$$Y^{k}_{ijgT_3} - Y^{k}_{ijgT_2} = \delta_g \cdot g + \Delta^{\mathrm{k}}_g F_i \cdot g + s_3(t^3_j) + s_2(t^2_j) + e_{ijg} \quad (model\ 1c)$$

$$Y^{k}_{ijgT_4} - Y^{k}_{ijgT_3} = \delta_g \cdot g + \Delta^{\mathrm{k}}_g F_i \cdot g + s_4(t^4_j) + s_3(t^3_j) + e_{ijg} \quad (model\ 1d)$$

Research aim 2: Describe how schools responded to covid-related closures (and their consequences)

RQ2a) How did schools respond during the closure period (Mar-May 2020) to support pupils working from home?

As of the time of writing, analysis of the teacher survey (delivered in July 2020, and covering the period from March to May 2020) was almost complete.

To answer RQ2a, we:

- Surveyed a sample of 511 teachers from 138 schools (7 schools from our sample of145 provided no responses to the teacher survey).
    - The process for recruiting teachers worked as follows. When schools signed up they were asked to indicate which year groups would be participating in the research. Schools were then asked for teacher names and email addresses

that taught the relevant year groups. Schools nominated up to 3 teachers per year group.
    - The survey primarily focused on: (i) background questions about a teacher's job role and class; (ii) details of their remote learning provision during March to May.
- Undertook an exploratory factor analysis (EFA), based on 15 questions (included in Appendix 1 of this document). Note that the EFA will be described in detail in the final report.
- Specified 5 constructs, using simple calculations based on raw survey responses. Each of these constructs
    - (i) reflected one of the primary factors identified in the EFA
    - (ii) provided relatively consistent responses at the level of the school (based on estimates of the ICC of responses in a multilevel analysis).[4]
    - (iii) is calculated based on the initial survey responses, to help make the construct interpretable
- The 5 constructs cover:
    1. Using a technology platform
    2. Timetabling
    3. Work submission regularity
    4. Speaking to students on the phone
    5. Live lessons and video
- The questions related to each of these constructs is illustrated in Appendix 1.
- These 5 school-level constructs describe school responses during March-May 2020. For the remainder of the analysis, we label these $Z_1, \dots Z_5$. Collectively these variables are labelled $\boldsymbol{Z}$.

In our report, we will describe the analyses leading to the creation of the 5 constructs in $\boldsymbol{Z}$. This will include descriptions of the survey instrument, the EFA and ICC analyses. We will also report on the covariance of $\boldsymbol{Z}$ along with its marginal distributions.

To answer RQ2b we:

- Released a survey to the recruited sample. For RQ2b, this was the second teacher survey, released in December 2020. For RQ2c this was the third teacher survey, released in March 2021.
- Followed a similar procedure to that described for RQ2a (in terms of both analysis and reporting)
- Specified 5 school-level constructs, $\tilde{Z}_1, \dots, \tilde{Z}_5$, collected into matrix $\widetilde{\boldsymbol{Z}}$
- Reported on all analyses leading to the creating of 5 constructs in $\widetilde{\boldsymbol{Z}}$, along with the distributions of $\widetilde{\boldsymbol{Z}}$.

To answer RQ2c we:

- Released a survey to the recruited sample. For RQ2c this was the third teacher survey, released in March 2021.
- Followed a similar procedure to that described for RQ2a (in terms of both analysis and reporting)
- Specified 5 school-level constructs, $Z_1^*, \dots, Z_5^*$, collected into matrix $\boldsymbol{Z}^*$.
- Will report on the distributions of $\boldsymbol{Z}^*$

[4] Details will be provided in the report but are beyond the scope of the study plan.

Note that for the 5 questions in Teacher Survey 3 where the prompts mirror those in Teacher Survey 1, we will compare mean responses for the sample of teachers who responded to both surveys.[5]

Research aim 3: Explore associations between school responses and FSM attainment gaps

We will address RQ3a and RQ3b using multilevel models. Consider model 3a, in which cohort and subject indices have been suppressed for simplicity:

$$Y_{ijT_1} - Y_{ijT_0} = \alpha_j + \Delta_j F_i' + \boldsymbol{\beta X_{ij}} + s_0(t_j^0) + s_1(t_j^1) + e_{ij} \quad (model\ 3a)$$

$$\alpha_j = \delta_a + \boldsymbol{\phi_0 Z_j} + \gamma_j$$

$$\Delta_j = \Delta_\alpha + \boldsymbol{\phi_1 Z_j} + \eta_j$$

Where:

- F_i' is a zero-centered individual indicator of FSM status for pupil i
- δ_a is the mean change in Y for non-FSM children in a typical school
- $\Delta_a + \delta_a$ is the mean gain score for FSM children in a typical school
- $\boldsymbol{X_{ij}}$ is a matrix of student and school controls (at the pupil level: gender, ethnic background, first language; at the school level: inspection ratings, historic attainment and percentage disadvantage)
- s_0 and s_1 represent splines
- $\boldsymbol{Z_j}$ is a matrix of 5 zero-centered constructs at the school level, describing school activity from Mar-July 2020, defined from $\boldsymbol{Z}$ (in research aim 2).
- $\boldsymbol{\phi_0}$ is a vector of 5 parameters, measuring the association between school-level constructs and the mean change in Y for non-FSM students
- $\boldsymbol{\phi_1}$ is a vector of 5 parameters, measuring the association between school-level constructs and the mean change in Y for FSM students (i.e. the association between school-level constructs 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 $\boldsymbol{X}$ **and Z**

The purpose of model 3a is to answer the following question:

[5] These are the following questions from "Teacher Survey 3" in the Appendix: Q2, Q3, Q4, Q6 and Q7.

- What are the associations between the different elements of $\boldsymbol{Z}$ and school-level changes in the FSM-achievement gap from Autumn 2019 to Sep 2020?

We will present point estimates of $\boldsymbol{\phi_1}$ along with 95% confidence intervals, using profile likelihood.

The analysis for RQ3b used the same setup. The only changes were:

- Instead of $\boldsymbol{Z}$, the predictors in model 3b will be $\widetilde{\boldsymbol{Z}}$
- The LHS variable will be the change in achievement from Sep/Oct 2020 (T_1) to Dec 2020 (T_2)

$$Y_{ijT_1} - Y_{ijT_0} = \alpha_j + \Delta_j F_i' + \boldsymbol{\beta X_{ij}} + s_1(t_j^1) + s_2(t_j^2) + e_{ij} \quad (model\ 3b)$$

$$\alpha_j = \delta_a + \boldsymbol{\phi_0}\widetilde{\boldsymbol{Z}}_j + \gamma_j$$

$$\Delta_j = \Delta_\alpha + \boldsymbol{\phi_1}\widetilde{\boldsymbol{Z}}_j + \eta_j$$

The analysis for RQ3c followed the same plan. The only changes are:

- Instead of $\widetilde{\boldsymbol{Z}}$ the predictors in model 3c will be $\boldsymbol{Z}^*$
- The LHS variable will be the change in achievement from Dec 2020 (T_2) to March 2021 (T_3)

$$Y_{ijT_1} - Y_{ijT_0} = \alpha_j + \Delta_j F_i' + \boldsymbol{\beta X_{ij}} + s_1(t_j^1) + s_2(t_j^2) + e_{ij} \quad (model\ 3c)$$

$$\alpha_j = \delta_a + \boldsymbol{\phi_0}\boldsymbol{Z}_j^* + \gamma_j$$

$$\Delta_j = \Delta_\alpha + \boldsymbol{\phi_1}\boldsymbol{Z}_j^* + \eta_j$$

Robustness and further analysis

This study is taking place in a period where schools are facing rapid changes and large uncertainties. In this section we have done our best to identify several threats to our analysis, and briefly outline the mitigations we intend to pursue. However, there may well be other issues that require further analysis that are not specified here. As our study is exploratory and descriptive – rather than focused on testing a pre-specified causal hypothesis – we believe this is appropriate.

Threats to be addressed in robustness checks and further analysis:

1. Non-linear scaling of tests: e.g. ceiling effects could attenuate score differences and distort interpretation of changes.
    - Mitigation1a) investigate ceiling/floor effects and evidence of scale compression
    - Mitigation 1b) examine core results when scores on different tests are equated using equipercentile equating (matching ranks) vs linear equating (matching mean and SD). Note that, before having seen attainment data, linear equating is our preferred approach.

2. School respondent sample may not be representative of national population
    - Mitigation 2a) compare school-level observed characteristics for sample against national data (e.g. extend the analysis presented on p8 of this document)

3. Teacher respondent sample not representative of all teachers at the school. This could result from bias in the selection of teacher respondents, or because teacher sample is too small and/or there is high within-school variation by teacher

- Mitigation 3a) compare our sample with the wider TeacherTapp sample (who have also been asked items from the Teacher Survey).
- Mitigation 3b) calculate intra-cluster correlation for teachers in each school, and within each year group teaching team (for those schools with multiple teachers per year group).

4. Scaling of $\boldsymbol{Z}$ and $\widetilde{\boldsymbol{Z}}$. In our primary analysis we propose to treat these constructs as being interval scales, however they may not have interval properties.
    - Mitigation 4a) investigate the sensitivity of results to different ways of scaling these factors (eg using IRT or dichotomising)
    - Mitigation 4b) compare with results from ordinal regression models

5. Model specification. In our primary analysis for RQ1 we propose a relatively simple model to estimate the changes in FSM-attainment gaps across time. We will also fit a repeated-measures model to assess whether the results are robust to a different specification. In particular we will fit models 2a and 2b:

Model 2a)

$$Y_{it} = a_i + \eta_A MONTHS_{it}^A + \eta_B MONTHS_{it}^B + \eta_C MONTHS_{it}^C +$$

$$\tau_A MONTHS_{it}^A FSM_i + \tau_C MONTHS_{it}^B FSM_i + \tau_C MONTHS_{it}^C FSM_i + e_{it}$$

$$e_{it} \sim N(0, \sigma_Y^2) \perp a_i \sim N(\mu_0 + \tau_0 FSM_i, \sigma_a^2)$$

Where:

- α_i represents a student-specific random effect accounting for dependence within student
- $MONTHS_{it}^A$ represents the number of months between Sep 2nd 2019 and Mar 20th 2020 during which pupils were in school
- $MONTHS_{it}^B$ represents the number of months since Mar 20th 2020 during which children were not in school[6]
- $MONTHS_{it}^C$ represents the number of months since Sep 2nd 2020 during which pupils were back in school.

Using this model, we can estimate the FSM attainment gap at 5 different time points: T_0, T_1, T_2, T_3 and T_4. This is a simple model drawn from the literature on estimating achievement gaps (e.g. Quinn et al., 2016). To assess whether this model is appropriate, we will examine whether there is an association between the timing of tests in November 2019 and FSM status. If there is no association, then we will fit model 2a. If there is a clear association, then allowing for differing time trends between FSM students and their peers may be somewhat unstable. In this case, we will fit both model $2a$, and model $2a^*$. In model $2a^*$ we will remove the τ_A interaction term. Adding this constraint will potentially induce some small bias, but will reduce variance (i.e. reduce standard errors).

Finally, we will see if our results are affected by individual-level absences. We can do this by fitting model 2b

$$Y_{it} = a_i + \eta_A MONTHS_{it}^A + \eta_B MONTHS_{it}^B + \eta_C MONTHS_{it}^{C^*} +$$

$$\tau_A MONTHS_{it}^A FSM_i + \tau_C MONTHS_{it}^C FSM_i + \tau_C MONTHS_{it}^{C^*} FSM_i +$$

$$\eta_D ABSENCE_{it} + \tau_D ABSENCE_{it} FSM_i + e_{it}$$

[6] Unfortunately, we do not have attendance data for the period March 2020 to Sep 2020

$$e_{it} \sim N(0, \sigma_Y^2) \perp a_i \sim N(\mu_0 + \tau_0 FSM_i, \sigma_a^2)$$

Where:

- $MONTHS_{it}^{C^*}$ are the number of months that pupil i attended school since Sep 2nd 2020
- $ABSENCE_{it}$ are the number of months that pupil i missed due to absences since Sep 2020

Missing data

Missing data may arise in the following circumstances:

1. A school drops out of the project
2. A school does not administer tests due to COVID-related reasons (e.g. closure)
3. A school only administers the tests for a subset of year groups
4. A pupil leaves a school
5. A pupil does not take a test (e.g. due to absence, including COVID-related absences) at a school where tests are administered

Following each test administration, we will report on the number of missing records in each of the categories above using cross-tabulations. This will include breakdowns by pupil-level and school-level disadvantage, absence and attainment, and an assessment of the extent to which missing data is concentrated within particular schools.

If more than 10% of test scores are missing for a particular test administration, we will use logistic regression to test how well missingness can be predicted by the dimensions of the cross-tabulations. We will then test the sensitivity of our results by imputing missing data under the assumption of missing at random (MAR). We will use Multiple Imputation using Chained Equations (MICE) to impute datasets, using all the variables in $\boldsymbol{X_{ij}}$ listed on page 10. Following the rule-of-thumb recommended by White et al. (2011), the number of imputed datasets will be the same as the average percentage of missingness (at student level). Using the imputed datasets we will re-estimate model parameters combining the means and standard errors using Rubin's rules.

Ethics

- The project was reviewed and approved by the directors at FFT, including the Managing Director, Paul Charman and the Data Protection Officer and Senior Information Risk owner, Helen Robinson.
- The data collected for this project is part of FFT's on-going service to schools so there is no specific consent requested in addition to their FFT Aspire data sharing agreements.
- The data is not being linked to NPD and the data is not being archived.

Data protection

- Schools that were approached to be part of this research are all FFT Aspire schools. FFT Aspire schools already have a GDPR-compliant data sharing agreement in place between FFT and the school which covers the use of personal pupil and teacher data for research. Link to FFT Aspire privacy policy: https://fft.org.uk/gdpr/
- The legal basis for processing personal data is legitimate interests (to be able to undertake research for public benefit) and necessary for FFT customers legitimate

interest (analysis of pupil performance and requirements, school performance and ensuring equality of opportunity and treatment of pupils)

- https://fft.org.uk/privacy/ highlights that as part of the schools data sharing agreements they agree for their personal data to be used to undertake research into how education systems function
- The data collected will be anonymised and analysed by a team of researchers from FFT and Teacher Tapp. We will ensure that schools, teachers and pupils are not named or identified in any publications or write ups regarding the research
- FFT is the data controller and responsible for schools personal data. The datasets created for this research will be deleted following publication. The data will not be archived or linked to the National Pupil Database.

Personnel

Project Delivery Team	**EEF**
Laura James – Project Lead/Manager	Matthew Van Poortvliet – Project Lead until main phase sign off
Ben Weidmann – Project Lead/Reporting	Jennifer Stevenson – Project Lead until main phase sign off
Rob Coe - Reporting	Celeste Cheung – Project lead
Dave Thomson – Analysis and Reporting	
Natasha Plaister – Analysis and Reporting	
Becky Allen – Teacher survey lead	

Risks

Description	Likelihood	Magnitude	Counter measure
COVID causing national/ local lockdowns	H	H	Report on extent to which it occurs; test sensitivity of results to missing data
School retention	M	H	Close contact with participating schools ensuring assessments and surveys have been completed and data shared with FFT. FFT offering support with project tasks. Financial incentives, paid to schools after submitting data.

Timeline

Date	Activity	Staff responsible/ leading
Sep 30th 2020	Submit draft study plan	FFT, Teacher Tapp
September/October 2020	Communicate to schools about September assessments and survey of teachers in Autumn term	FFT
September 2020	Schools take September assessments (missed Summer paper)	Schools
September 2020	FFT to work with schools to set up test data and MIS data sharing with participating schools	FFT
October 2020	Communicate with schools about Autumn Assessments	FFT
October 2020	Produce and finalise second survey questions	Teacher Tapp and FFT
December 2020	Teachers take second teacher survey	Schools
December 2020	Schools take Autumn assessments	Schools
Dec2020/January 2021	Process data and run analysis	FFT
January/February 2021	Summary report for EEF and participating schools	FFT
March 2021	Teachers take third teacher survey	FFT

March/April 2021	Schools take Spring Assessments	Schools
June 2021	Schools take Summer Assessments	Schools
July 2021	Process data and run analysis	FFT
July 2021	Pay financial incentives to schools	FFT
August 2021	Addendum report for EEF and participating schools including Spring, Summer Assessments and third teacher survey analysis	FFT

Appendix: Teacher Surveys

Teacher Survey 1

(Preceding these questions, information on the teacher's school, class and job role were collected. Following these questions, we asked about the return to school and how many days each year group was able to attend in June and July.)

Note that quite a few of these questions will not be used in the analysis, as they showed within-school variation. This will be described in detail in the report.

Table 1: Survey items relating to learning during the first half of summer term 2020

(Note: the right column lists specific variable names that were used in the full analysis. In the final report, these variables will be described in more detail with a full analysis of the teacher survey).

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 (e.g. Google classroom, Microsoft Teams, Class Dojo, Seesaw, Firefly etc…) • Via a social media site (e.g. 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` This question was coded in conjunction with question 2. 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 (e.g. Google classroom, , Microsoft Teams, Class Dojo, Seesaw, Firefly etc…) • Via a social media site (e.g. 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
		`curric_div_slowed`
		Coded 1 if teacher some reduced content or slowed pace (0 otherwise).
		
		`curric_div_paused`

	• 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 (e.g. Oak National or BBC Bitesize topics) • None of the above statements align with our practice during lockdown • Not relevant / cannot answer	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 (e.g. 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.
	• 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_fully`
		Coded 1 if majority of resources and activities were created with parents in mind (0 otherwise).
		
		`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 (e.g. 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?
	• 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
		`feedback_praise`
		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
		`feedback_given`
		Binary indicator for whether or not the teacher encouraged work to regularly be submitted.
		
		`feedback_given_daily`

	• I gave no suggestion about how frequently work should be sent in • Not relevant / cannot answer	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.

Note: Cannot answer / not relevant responses generally coded as missing

Mapping of questions to constructs

Construct label	Description	Questions
Technology	Did teachers use a technology platform? Did the platform have chat functions? Was technology used for daily registration?	Q1, Q2, Q4, Q7, Q11
Timetabling	Did teachers use a timetable?	Q8
Work submission	Did teachers have any requirements for regular work submission? (ranges from no recommendation about work submission to daily requirements)	Q13
Phoning students	Did teachers speak to students (not parents) directly on the phone either one or twice during the half term	Q10

	(scored 1) or about once a week or more (scored 2).	
Live lessons and video	Did teachers hold lives lessons, or pre-recorded video of teacher?	Q9

Teacher Survey 2

Primary COVID-19 and learning loss (Dec 2020)

Teaching isolating pupils

1. Which of the following statements are true about your class and school so far since September? Tick any that apply
   - Individual pupils in my class have needed to isolate at home due to suspected COVID-19
   - My entire class has had a period of whole class isolation at home
   - Another class in the school has had a period of whole class isolation at home
   - At least one student in my CLASS has been tested as COVID-19 positive
   - At least one student in my SCHOOL has been tested as COVID-19 positive
   - None of the above have happened so far
   - Not relevant / cannot answer

2. Which of the following types of provision were available for an individual child who needed to isolate whilst their peers were in class? Tick any that apply
   - Streaming classroom lessons to the student's home
   - Recording videos of the class lesson for student to watch
   - Pre-recorded videos made by someone in the school
   - Providing link to instructional videos made by others (e.g. White Rose, Oak National, BBC)
   - A single portal, such as google classroom, where students can access AND submit work
   - Ability to collect printed worksheets from the school office, on request
   - Giving feedback on submitted work
   - None of the above
   - Not relevant – no students needed to isolate

3. Last term, how much time each day have you typically spent supporting remote learning for isolating students, whilst also teaching a class?
   - No time – no isolating students this term
   - 15 minutes a day, on average
   - 30 minutes a day, on average
   - 45 minutes a day, on average
   - 60 minutes a day, on average
   - 90 minutes a day, on average
   - More than 90 minutes each day, on average
   - Not relevant (e.g. I have not taught a class in person for the past few weeks)

4. At your school last term, which of the following types of provision were available whilst a whole class were isolating together? Tick any that apply
   - Daily online registration sessions
   - 'Live' online lessons, with or without student interaction
   - Pre-recorded videos made by someone at the school
   - Providing link to instructional videos made by others (e.g. White Rose, Oak National, BBC)
   - A single portal, such as google classroom, where students can access AND submit work
   - Ability to collect printed worksheets from the school office, on request
   - Giving feedback on submitted work
   - None of the above
   - Not relevant – no classes needed to isolate this term

Strategies to respond to missed learning

5. Has your school done any of the following to compensate for lost learning time last year? Tick any that apply

   - Removing items from the curriculum to help students get back on track
   - Running extra lunchtime or after-school classes to help students catch-up
   - Increasing small group/1-to-1 intervention time DURING lessons
   - Increasing small group/1-to-1 intervention time OUTSIDE lessons
   - Extending the school day to allow for greater lesson time
   - Opening the school to students during the holidays
   - Something else not listed
   - None of the above
   - Not relevant / cannot answer

6. 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.

   - Zero minutes - we have not reallocated lesson time at all
   - Extra 15 minutes per day on English and maths
   - Extra 30 minutes per day on English and maths
   - Extra 45 minutes per day on English and maths
   - Extra 60 minutes per day on English and maths
   - Extra 90 minutes per day on English and maths
   - More than 90 extra minutes per day on English and maths
   - Not relevant / cannot answer

7. 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
   - English
   - Maths
   - Science
   - History, Geography or RE
   - Art, Music or DT
   - Computing
   - Modern languages
   - I haven't taught any topics that were missed in summer term 2020 so far this year
   - Not relevant / cannot answer

8. Was your school making greater or less use of small group and individual face-to-face interventions (e.g. run by TAs) last term?
   - There are more intervention groups running than in a normal year so far
   - Similar number of intervention groups running compared to a normal year so far
   - There are fewer intervention groups running compared to a normal year so far
   - Not relevant / cannot answer

9. To what extent was missed learning over the past 9 months making it difficult for you to plan lessons last term?
   - A great extent - missed learning is a big problem for us
   - Some extent
   - A little extent
   - Not at all - missed learning has not been a big impediment to my lesson planning this term
   - Not relevant / cannot answer

10. In your current class, what proportion of pupils do you feel have fallen behind due to missed learning during summer term last year?
    - All of them
    - The majority of the class
    - About half of the class
    - Less than half of the class
    - Very few pupils
    - No pupils

Teacher Survey 3

Learning during the first half of summer term

This set of questions asks you about your provision for those in your class who were learning at home between January and March 2021. If you supported your class during this time then please think about provision for that class. If you did not have a class of your own (e.g. headteachers), please think about your school's provision for lower Key Stage Two children.

1. Did you use a technology platform (e.g. Google classroom, Microsoft Teams, Class Dojo, Seesaw, Firefly etc…) to communicate with families in any of the following ways? You may tick more than one response.
   - No, I did not use a technology platform for communication with families
   - To set pupils' work
   - To receive pupils' work submissions
   - To share examples of student work with everyone
   - We held a daily registration or check-in system
   - Parents or pupils typed messages for me reply to via the platform
   - I spoke to parents or pupils using video/audio chat via the platform
2. 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
3. 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 (or less)
   - I gave no suggestion about how frequently work should be sent in
4. 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, I spoke to them each once or twice during the half term
   - Yes, I spoke to them each several times during the half term
   - Yes, I spoke them to each about once a week (or more frequently)
5. Did you try teaching a 'live' lesson to your class in January, February or March (with students either speaking or on mute)?
   - No
   - No, but I did pre-record a video of myself for them to watch
   - Yes, I held one or two live lessons during this time
   - Yes, I held several live lessons during this time
   - Yes, I held a live lesson about once a week
   - Yes, I held one live lesson most days
   - Yes, I held more than one live lesson most days
6. Which best describes the type of feedback you felt able to give on pieces of work submitted by pupils?
   - 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, with general praise and encouragement on others

- I gave specific feedback intended to support learning on all/ALMOST ALL pieces of work

7. 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 ALTERED the school's curriculum during lockdown
8. Which of the following approaches did you use to contact families with children not participating in remote learning (tick any that apply)?
   - Phoning or emailing the parents
   - Visiting the family's home
   - Asking these children to come into school each day
   - None of the above - we have not been able to do anything to contact them
   - We have not had any children who did not participate
9. Who was responsible for supervising the pupils in your class who are attending school as keyworker or vulnerable children?
   - I am, and generally have a teaching assistant to support me in the class
   - I am, usually with no teaching assistant support
   - They are mostly supervised by another teacher
   - They are mostly supervised by teaching assistants or other (non-teaching) staff
10. Overall, how good do you feel the home learning experience was for your class between January and March? Tick the response that most closely aligns with your feelings.
    - Very successful - all or almost all pupils were consistently completing school work
    - 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

References

EEF. (2020a). *Impact of school closures on the attainment gap: rapid evidence assessment*. Education Endowment Foundation, available at https://educationendowmentfoundation.org.uk/evidence-summaries/evidence-reviews/school-closures-rapid-evidence-assessment/.

EEF. (2020b). *Remote Learning: Rapid Evidence Assessment*. Education Endowment Foundation, available at https://educationendowmentfoundation.org.uk/public/files/Publications/Covid-19_Resources/Remote_learning_evidence_review/Remote_Learning_Rapid_Evidence_Assessment.pdf.

Kolen, M. J., & Brennan, R. L. (2013). *Test equating: Methods and practices*: Springer Science & Business Media.

Quinn, D. M., Cooc, N., McIntyre, J., & Gomez, C. J. (2016). Seasonal dynamics of academic achievement inequality by socioeconomic status and race/ethnicity: Updating and extending past research with new national data. *Educational Researcher, 45*(8), 443-453.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: issues and guidance for practice. *Statistics in Medicine, 30*(4), 377-399.

Extension to original study plan (June 2021)

As described above, this project was initially designed to focus primarily on two periods:

1. The initial lockdown, beginning in March 2020, when schools were closed to a vast majority of pupils.
2. The autumn term of 2020, when schools were re-opened to all students.

However, as of January 2021, schools in England were once again closed to most pupils. In response, **we proposed to extend the project by conducting an additional teacher survey.** This allowed us to extend research goals 2 and 3 to cover this second period of school closures.

This brief addendum makes explicit the research questions that were added to the original study plan published on the EEF website.

Additional research questions and analysis

The project had always planned to examine the impact of covid disruptions on estimates of the attainment gap in March 2021 and May 2021 (see RQ1c and RQ1d).

We have added an additional teacher survey, thereby adding two research questions to the project:

- RQ2c: How did schools respond during the second period of school closures (Jan – Mar? 2021) to support pupils working from home?
- RQ3c What is the association between school responses during the second closure period (Jan-Mar 2021) and changes in the FSM-attainment gap when children returned to in-person schooling (from Dec-2020 to March 2021)?

The additional teacher survey (i.e. teacher survey 3) was administered to the same sample of teachers described above, and followed the same process as the previous two teacher surveys:

- Questions were developed and tested by Teacher Tapp, and focus on the same 5 constructs that were analysed in RQ2a (i.e. teacher survey 1).
- The analysis of the survey responses will follow the same procedure to that described for RQ2a and RQ2b.

We will use the same analytical approach for RQ3c as we used for RQ3a and RQ3b. Compared to model 3a, the changes will be as follows:

- Instead of $\boldsymbol{Z}$**,** the predictors in model 3c will be derived from the third teacher survey, denoted by $\boldsymbol{Z}^*$
- The LHS variable will the change in achievement from Dec 2020 (T_2) to March 2021 (T_3):

$$Y_{ijT_4} - Y_{ijT_2} = \alpha_j + \Delta_j F_i' + \boldsymbol{\beta X_{ij}} + s_1(t_j^4) + s_2(t_j^2) + e_{ij} \quad (model\ 3c)$$

$$\alpha_j = \delta_a + \boldsymbol{\phi_0 Z^*}_j + \gamma_j$$

$$\Delta_j = \Delta_\alpha + \boldsymbol{\phi_1 Z^*}_j + \eta_j$$

We will report the same outputs for RQ3a, RQ3b and RQ3c.