Approaches to address missing KS2 data in LSYPE2: user guide

July 2016

Catherine Saunders, Bryn Garrod & Alex Sutherland

RAND EUROPE

Preface

The Second Longitudinal Study of Young People in England (LSYPE2), which started in 2013, was designed in order to understand the compulsory education, school-to-work transitions, careers and lives of young people. Although it is of rich academic interest, the key purpose for this dataset is to provide a resource for evidence-based policy development. A significant barrier to achieving this purpose could be the ‘missingness’ present in LSYPE2, owing to a boycott of Key Stage 2 (KS2) testing in 2010. In 2010, 15,518 maintained schools were expected to administer KS2 tests, but 4,005 (26 per cent) of these did not administer it. Boycotts of national tests leave gaps in pupils’ attainment records and, in the case of LSYPE2, threaten to undermine a large-scale longitudinal study with substantial policy relevance. This project sought to find a way to calculate values for pupils who attended schools that boycotted KS2 tests in 2010 and/or partly mitigate the effect of the boycott on this study.

The Department for Education (DfE) commissioned RAND Europe, in collaboration with Professor Vignoles at the University of Cambridge and Professor Brunton-Smith at the University of Warwick, to explore and develop a strategy to address this missing data relating to the boycott.

We thank responders to the consultation sent out during this work (Tom Benton, Jake Anders, Steve Strand, Andy Ross) and we thank, in particular, Professor Harvey Goldstein for his engagement in and comments on the work. We are also very grateful to Clare Baker and Sarah Lasher at the Department for Education for their support. Any errors are the responsibility of the authors alone.

During this work, we created an inverse probability weight (for the 8,684 out of 11,823 pupils with responses at Wave 1 who had linked KS2 results) and a dataset of imputed values for KS2 attainment among pupils who attended boycott schools (for the 8,882 pupils who remain in LSYPE2 at Wave 3). In this report, we present advice to analysts faced with the issue of missing data and we provide guidance for users of these datasets. Full technical details of how these data were created are presented in the accompanying technical report.

Dr Alex Sutherland
Research Leader
RAND Europe
Tel: +44(0)1223 353 329
Email: alex_sutherland@rand.org

Contents

Abbreviations and acronyms

CI	confidence interval
DfE	Department for Education
IPW	inverse probability weight(ing)
KS1	Key Stage 1
KS2	Key Stage 2
KS4	Key Stage 4
LSYPE2	The Second Longitudinal Study of Young People in England
MI	multiple imputation
NPD	National Pupil Database
SE	standard error
TA	teacher assessment

Introduction

This user guide presents recommendations and guidance for analysts and researchers who are using data from the Second Longitudinal Study of Young People in England cohort (LSYPE2) together with linked Key Stage 2 (KS2) attainment data from the National Pupil Database. It accompanies the report *Missing Key Stage 2 data in LSYPE2: technical report* (herein 'the technical report'). The technical report describes the analyses and approaches taken to the development of a KS2 dataset using the statistical methods of multiple imputation (MI) and inverse probability weighting (IPW) – both of which can be used by researchers to address issues of possible bias. In addition, MI also addresses the issue of efficiency in research involving the KS2 test results in LSYPE2. Further resources that may support understanding of this document are available online (Bartlett and Carpenter, 2013a, Bartlett and Carpenter, 2013b).

Background

Missing data is a challenge for any researcher using observational data; therefore, different approaches should be used to examine the consequences of missing data in any analysis. The challenges associated with Key Stage 2 (KS2) data that are missing as a result of the 2010 KS2 boycott are a particularly important case. In 2010, 15,518 maintained schools were expected to administer KS2 tests, but 4,005 (26 per cent) of these did not administer it. The Second Longitudinal Study of Young People in England (LSYPE2), now with three completed survey waves, recruited pupils in year 9 who should have taken their KS2 tests in 2010. Of the 8,882 pupils in LSYPE2 who responded to the survey at wave 3 and gave consent for linkage,[1] 6,606 had a KS2 average points score, but 2,276 pupils did not have this information.

There are two key concerns associated with missing data. The first is *bias*. If young people surveyed in LSYPE2 who have missing KS2 test results are systematically different from those with KS2 test results, then it is possible that results based only on those young people with available test results may be different (*biased*) compared with results if KS2 tests had been available for all cohort members. The second concern is *efficiency*. If researchers using these data considered only those pupils with a KS2 test score available, a large amount of information would have to be excluded from analyses, leading to more uncertainty about results (which would be reflected in larger standard errors (SE)). Furthermore, listwise deletion of data (AKA complete-case analysis) is also

[1] The 8,882 includes those cases where both the main parent and the young person gave consent to linkage of the survey to information held on the pupil, including attainment at KS2, in the National Pupil Database (NPD) and where linkage to information about whether the pupil attended a boycott school at KS2 was possible.

not cost effective, as it means that information from up to one quarter of LSYPE2 participants is being discarded.

In the first section of this report, we describe general strategies and considerations for all researchers who need to consider missing KS2 attainment in their research. In the second section, we present Stata code and example analyses for analysts who will be using the MI and IPW variables we have produced.

Section 1: Addressing missing data in LSYPE2

There are several steps for researchers to consider when exploring missing data in their analysis of LSYPE2 data. In this section we include some introductory remarks for analysts, with practical comments and a brief discussion of the strengths and limitations, covering the following areas:

- possible reasons why data may be missing, accompanied by descriptive statistics and analyses where appropriate;
- conducting complete-case analyses;
- using the KS2 teacher assessment (TA) results;
- excluding KS2 attainment from the analysis; and
- using multiple imputation or other approaches to address missing data.

In section two, we describe how analysts could use the imputed KS2 dataset and IPW that we developed to address missing data – in particular, missing data related to the 2010 KS2 boycott.

General pointers for analysts on working with LSYPE2 and boycott data

This section provides advice for analysts working with LSYPE2 cohort data who will be thinking about how to address the issue of KS2 data that are missing due to the boycott.

Identify boycott schools

The variable KS2FAAT10_CONTFLAG from the National Pupil Database will tell you if the pupil attended a primary school with results that were 'significantly affected' by the 2010 boycott (CONTFLAG means ‘contingency flag’). This indicates whether or not a pupil attended a boycott school. Use this linked variable from the National Pupil Database (NPD) to explore whether or not a cohort member attended a boycott school in 2010. Examples of this type of analysis are presented in our technical report, where we explore the characteristics of pupils who attended boycott schools.

As a technical point for analysts it is worth noting that 167 pupils in the LSYPE2 cohort are identified using the variable KS2FAAT10_CONTFLAG as having attended a boycott school but have a test score available. There are several possible mechanisms for this. There would be some pupils where the boycott flag was set who had ‘B’ (below the level of the test) or ‘T’ (working at the level of the test but unable to access it) – the school supplies these codes when registering pupils for the test to indicate that they will not take it so these were known before the boycott took place. In addition, there were also a small number of boycott schools which took part in the single level test pilot in maths. Pupils in these schools would have results in maths but not English (since the maths single level tests took place later and were not part of the boycott). There may also be a small number of pupils who were dual registered, perhaps at a mainstream school but spending some time at a special school or a PRU. Their results would be shown against their main school but they may have actually taken the tests at a different school. If their main school boycott the tests but the other didn't then it would appear that they have results from a boycott school. Finally, there will be those pupils who moved schools close to test week. Pupil results would normally be included at the school where they took the tests but schools sometimes apply during the checking exercise to move results back to the old school.

Think through missingness in your sample

A first step in thinking through missingness, for any analysis, is to describe what data may be missing in your intended analysis sample. A next step is to consider reasons why data may be missing. Missing data concerns not just KS2 data associated with the boycott in 2010. Data may also be missing due to attrition at later survey waves or due to lack of consent to linkage. In LSYPE2, the variables NPDlinkYP and NPlinkMP identify whether the young person or parent, respectively, consented to linkage. For data to be linked, both must have consented.

Describe differences between your subsamples

Consider reporting descriptive statistics in your analysis sample for young people with and without missing data (and by the different types of missing data, through either the boycott, attrition, or lack of consent to linkage), to describe whether there are any differences in the characteristics of young people with and without missing data.

Advice for analysts on why data may be missing

The first steps for analysts using information from LSYPE2 are to consider and describe the sample that will be used for each analysis and to understand the different ways in which data can be missing.

For example, 13,100 pupils were recruited at wave 1 of LSYPE2. For 11,823 of them, linked NPD data on whether or not they attended a boycott school are available. By wave 3 of LSYPE2, 10,010 young people remained in the study and were interviewed. Consent for linkage to NPD data was given for 8,882 of these pupils.

Missing data occur across LSYPE2 for a variety of reasons. Although missing KS2 attainment as a result of the boycott in 2010 is the focus of this work, there are other important reasons why data may be missing that analysts should bear in mind. One particularly important issue for any cohort study is attrition, where young people recruited at wave 1 withdraw from the study at later survey waves. In LSYPE2, to ensure adequate sample sizes are maintained, groups of young people more likely to withdraw were over-sampled in wave 1 (TNS BMRB, 2013).

However, there are other reasons why KS2 test scores may be missing, which are particularly relevant to this study. For example:

- consent to linkage was not obtained;
- linkage was not possible; or
- the pupil did not take the KS2 tests because the pupil was living abroad or attending an independent school in 2010.

It is also important to consider missing data in other variables. Some survey questions, particularly those asking about more sensitive topics, may not be answered by all young people who respond at a particular survey wave.

Advice for analysts on conducting a complete-case analysis

Our statistical analyses reported in the technical report found that very few LSYPE2 pupil or household characteristics predict a pupil having attended a boycott school at KS2. Although attending a boycott school may be related to unobserved factors that then influence pupil achievement, this lack of prediction somewhat reduces concerns about unrepresentative KS2 missingness in this cohort. In short, the lack of a systematic relationship to missingness of pupil-level predictors reinforces that the LSYPE2 cohort is representative of the wider population from which it was drawn. This is a key finding because it suggests that things that predict KS2 attainment do not predict missingness and that bias from analyses based on complete cases – that is, analyses excluding pupils who attended boycott schools – is therefore likely to be small.

Further, this lack of strong, consistent associations between pupil characteristics and school-level decisions to boycott is to be expected. The ‘missingness mechanism’ for KS2 test results in the LSYPE2 cohort occurred at the school level rather than at the individual pupil level, because it was the schools – or rather their head teachers – who chose to boycott the tests or not.

Advice for analysts on complete-case analyses

For analyses using KS2 attainment, a complete-case analysis may be appropriate. In order to arrive at unbiased estimates, analyses should adjust for variables that predict missingness (primarily KS2 school characteristics) where appropriate (see technical report). However, estimates based on complete cases will still be inefficient. In these analyses we would effectively be ‘throwing away’ one quarter of the sample. This reduces the power to detect statistically significant associations, and increases the chance of making Type 2 errors of inference (i.e. falsely accepting the null hypothesis). Therefore, although complete-case analysis results may be unbiased, there is a case for using multiple imputation estimates to improve efficiency. Efficiency is particularly important for researchers who wish to make statistical inferences to the LSYPE2 population (e.g. through the use of significance testing).

Advice for analysts on using the KS2 teacher assessment (TA) results

During the boycott, only the KS2 external tests were boycotted. This means that for pupils who attended a boycott school in 2010, KS2 TA results will still be available. These might be considered a viable substitute for test results. Using TA results instead of KS2 test results is an approach that has been taken in some analyses of KS2 performance – for example, looking at progress from KS2 to KS4 (Ofsted and Department for Education, 2015).

However, there are two main limitations to replacing the missing KS2 results with TA results. The first is that external test results are considered a more reliable and objective measure of pupil attainment than teacher assessments. The second limitation, related to this, is that using TA results alone may actually introduce bias. The reason is that in 2010 only 37% of schools provided TA scores before the external test results were released (Bew, 2011), which suggests that some schools may use test results as some kind of internal moderation for the TA. Internal analysis by the Department for Education (DfE) used to produce the KS2 Statistical Releases (Department for Education, 2010a, Department for Education, 2010b) identified that teacher assessments in boycott schools in the year of the boycott were likely to be appreciably higher than those in non-boycott schools compared with the years before and after the boycott, returning to similar levels in the following years.[2] What the DfE analysis tells us is that KS2 TA scores were slightly higher in boycott compared with non-boycott schools in 2010. This means that using TA scores for all pupils is also likely to lead to analysis results that are biased.

[2] This was explored by looking at the improvement in share of pupils at level 4+ in both English and maths by school between 2009 and 2010 and then between 2010 and 2011. The analysis shows that there was substantially higher (1.1 percentage points) increase in the share between 2009 and 2010 for boycott schools than for non-boycott schools in English, followed by 0.7 percentage point *smaller* increase between 2010 and 2011.

Advice for analysts on excluding KS2 attainment or substituting with other KS data

In many situations, excluding measures with high levels of missingness from analysis may be possible. If KS2 attainment is not central to a given analysis, then excluding this measure might be one option. But because KS2 attainment is a key measure in an educational cohort, this is unlikely to be feasible in many situations. However, it is an approach that should be considered. For example, for later survey waves (after wave 3), using KS4 attainment alone may be a 'good enough' approach to accounting for prior attainment in some analyses. For earlier survey waves, using KS1 attainment alone is also a possibility.

Advice for analysts on multiple imputation or other approaches to addressing missing data

Many researchers may want to carry out their own analyses to address missing data. The technical report includes a description of the issues we identified during the course of this project, and of how we developed MI variables and an IPW to address missing KS2 data related to the KS2 boycott. Analysts who are working on their own imputation may find this a helpful resource. For those intending to use the MI data or IPW weights created by the main project, the following sections set out how to do this.

Analysts should also consider whether it is appropriate to use the MI data sets provided here for their analyses. In particular, if analyses involve variables that are not those included in the imputation model then use of the MI data sets requires caution. The imputation model included the following variables:

Pupil measures

- Free school meals at any point in the last 6 years (NPD)
- Female (Survey, Wave 1)
- Ethnicity (Survey, Wave 1)
- Special Educational Needs (NPD)

Household measures

- Highest household qualification-level (Survey, Wave 1)
- Household non-English as a first language (Survey, Wave 1)
- Household size (Survey, Wave 2)

Attainment

- KS4 average point score (NPD)
- KS4 average point score - squared (derived)

School

- The mean point score at KS2 of pupils at the school where they took KS4 exams (NPD)
- KS2 school mean attainment in 2009 (NPD)
- % EAL pupils in the KS2 school in 2010 (NPD)
- Mean IDACI of pupils at KS2 school in 2010 (NPD)
- % pupils with SEN in the KS2 school in 2010 (NPD)
- % pupils at the school the pupils attended at time of sampling with missing KS2 data because of the boycott (NPD)

Outcomes

- Attainment; KS4
- Ambitions and future plans; Higher education aspirations (Survey, Wave 1)
- Bullying; Having been bullied in the last 12 months (Survey, Wave 2)
- Mental health, well-being and non-cognitive skills; GHQ (Survey, Wave 2)
- Participation in risky behaviours (Survey derived variable, Wave 1)

Section 2: How to use the IPW and MI variables

Recommendations for all analyses using the IPW and MI variables

We recommend that any analyses using the IPW or MI datasets be clearly labelled as such in any report. We also recommend that a complete-case analysis be carried out as a further sensitivity analysis. It may be helpful to compare the results from each of the three sets of analyses (weighted, MI and complete-case) to evaluate the results.

If the results are similar, we would probably recommend reporting only the complete-case results, but if there are differences (either in the magnitude of the estimate or in the size of the standard errors), we would recommend reporting both sets of analyses and highlighting where differences occur.

Data sets

The imputations and the inverse probability weights are stored in two separate Stata datasets, **MI_for_user.dta** and **ipw_for_user.dta**. The structures of the files are outlined below. Both files contain the new data generated as part of our work, and none of the original raw data; we assume that the user already has any existing data that she or he wishes to use. The user must have the variable surveyID_W1_ADM to link existing data with the imputations or inverse probability weights.

Sample sizes depend on approach used

A final point that it is important to highlight for users of the IPW and MI data sets is that LSYPE2 had 13,100 responses at Wave 1. The IPW and MI are based on different samples, and so analyses based on the two approaches may not be directly comparable.

The IPW should be used with the 8,684 out of 11,823 pupils with survey responses at Wave 1 who had linked KS2 results *and* who did *not* attend a boycott school.

The MI dataset should be used with the 8,882 pupils who remain in LSYPE2 at Wave 3, have consent for NPD linkage from both parents and pupil, *and* for whom linkage was possible (see Table 1 for a comparison of samples for IPW and MI approaches).

Table 1: comparison of sample sizes for missing data approaches

Analysis technique	Based on wave:	Total number of pupils in wave	Consent to linkage[3]	Total number of pupils with NPD linked data at that wave	Number of pupils with NPD linkage to the KS2 school-level boycott flag	Number of pupils who attended / did not attend a boycott school
IPW	1	13,100	12,504	11,887	11,823	3,139 / 8,684
MI	3	10,010	9,531	9,077	8,882	2,276 / 6,606

Description of Stata files

MI_for_user

This is the multiple imputation file which has 32 variables, as follows:

surveyID_W1_ADM identifies the observation and is used for merging the imputations with other data.

miss_KS2 is a binary 0–1 variable that indicates whether KS2_CVAAPS was missing in the dataset we used for the multiple imputation, with 1 indicating missing and 0 indicating present. When the imputation dataset is merged with other data, this variable will be missing for any observations not in the imputation dataset.

The variables **_n_KS2_CVAAPS** for n = 1 to 30 are the 30 imputations of KS2_CVAAPS. If KS2_CVAAPS was not missing – indicated by **miss_KS2** = 0 – then all 30 imputations are equal to the original value.

Of the 8,882 pupils in LSYPE2 who responded to the survey at wave 3 and gave consent for linkage, 6,606 had a KS2 average points score, but 2,276 pupils did not have this information. The MI data set imputes values for these 2,276 pupils. When the imputation data is merged with other data (for example Wave 1 data) and an ‘mi estimate’ command is subsequently run, these variables will be filled in with the value of KS2_CVAAPS for any observations not in the imputation dataset (or missing if KS2_CVAAPS is missing).

Note that, although multiple imputation by chained equations imputes 30 values for all variables with any missing data in the imputation model, we provide only the imputed values for **KS2_CVAAPS**.

ipw_for_user

This is the inverse probability weight file, which has three variables:

[3] NPlinkP_W1_MP_NPD “Yes” at wave 1; NPDConMP_W3_MP “Yes” or NPlinkP_W3_MP “Yes” and NPDConYP_W3_YP “Yes” or NPDlinkYP_W3_YP “Yes” at wave 3

surveyID_W1_ADM identifies the observation and is used for merging the inverse probability weights with other data.

IPW is the inverse probability weight (variable name: ipw_mi), calculated as the multiplicative inverse of the probability of the observation being present – that is, non-missing – in the dataset that we used for the missingness analysis (n = 11,823 observations). The IPW is created for the 8,684 out of 11,823 pupils with responses at Wave 1 who had linked KS2 results, and who did not attend a boycott school.

How to use the datasets

Technical details

In this section, we give example Stata[4] commands that may be useful (users will need to edit them as appropriate for their own analysis).

MI_for_user – combining datasets

Prepare the data that you wish to use in your analysis. This must include the variable surveyID_W1_ADM to allow matching.

Merge the multiple imputations using the following command:

```
. merge 1:1 surveyID_W1_ADM using "[your file's
location]/MI_for_user.dta"
```

In your data preparation, you may have dropped some observations that are present in the imputed data, or you may not have had these observations in the first place. Drop the imputed data, which will be of no use in your model, with the following command:

```
. drop if _merge == 2
```

The _merge variable does not tell you anything you don't know already. If you have an observation not in the imputed dataset, then **miss_KS2** will be missing – so you can drop it with the following command:

```
. drop _merge
```

Using the MI variables

Multiple imputation data analysis in Stata is similar to standard data analysis. The standard syntax applies.

Type

```
. help mi estimation
```

---

[4] It would be possible to analyse the MI data provided using other statistical software, for example, SPSS or R. An SPSS version of the MI data has also been provided. To convert from Stata to SPSS, we used the user-written command savespss [savespss "MI_default_SPSS.sav", codepage(1252) strlmax(32767)] (see http://www.radyakin.org/transfer/savespss/savespss.htm).

for a list of all the commands that you can use with the multiply imputed dataset. Any command using the MI KS2 variable works by typing

```
. mi estimate: [your model]
```

Using the IPW variable

In Stata, ‘pweights’, or sampling weights, are weights that denote the inverse of the probability that the observation is included because of the sampling design. The IPW is the inverse probability weight, calculated as the multiplicative inverse (reciprocal) of the probability of the observation being present – that is, non-missing – in the dataset that we used for the missingness analysis (11,823 observations). For this dataset, using the pweight option is appropriate. For other analyses, for example if using the weight with later waves of survey data, the iweight (with a less formal statistical definition) is the better approach.

Merge the IPW variable from the dataset ipw_for_user.dta in the same way as you did for the MI data.

Weights in Stata have the following general form, and they can be used with most commands:

```
[your model] [pweights=IPW_after_MI]
```

Stata help files and online documentation about weights can also be helpful for specific syntax.

Note that it is good practice to run your model on complete cases without weights as a sensitivity analysis.

Using the IPW in conjunction with other survey weights

The IPW was not created to reflect the full survey weighting (for sampling and attrition at each wave). In general, it could be appropriate to combine weights multiplicatively, but we have not explored whether it would be appropriate with this and other LSYPE2 weights. We recommend that the primary use of the IPW be to compare the findings with an unweighted analysis to assess the magnitude of any bias associated with the missing boycott data. So the steps would be to run models with and without IPW, assessing whether there are any substantive differences between the results. If there are none, that is, if the complete-case analysis appears unbiased, then it might be appropriate to use the survey weights for your analysis. (Note that our work on predictors of missingness in the technical report for this work suggests that pupil-level analyses are unbiased, because very few pupil characteristics were predictive of missingness).

Example analyses

Here we present output from example analyses using the imputed data sets, the inverse probability weight and a complete case analysis. It is worth highlighting again that the analysis data sets are different between the MI (Wave 3, linked data) and IPW (Wave 1, linked data). In Example analysis 1 (descriptive statistics) we present complete case analyses for both samples (IPW and MI) to allow direct comparison. In addition, these simple analyses allow users to explore whether they are able to directly replicate the results presented in this user guide in their own work.

In Example analyses 2 and 3 (stratified and regression analyses) the MI KS2 test scores, and IPW, were merged into the Wave 1 survey data, and the largest available data set was used for each analysis (with the inclusion of all possible pupils across waves and case wise deletion for missing data for variables other than KS2 test scores with missing values).[5] Sample sizes for each analysis vary (and are shown in the example output), and results are not directly comparable across approaches.

Example 1 – Descriptive statistics

By comparing the mean KS2 average point score with the MI and IPW estimates, we see that, while all three are very close, their standard errors differ. For the IPW, the SE is widest, and for MI, the SE is narrowest, as would be expected given the additional information used for MI. We explored whether the increase in SE associated with the IPW was plausible, using a limited simulation study, and we found it to be so in the tested scenarios. IPW analyses are inefficient, although a large, rather technical literature now exists that develops more efficient estimators based on IPW (Scharfstein et al., 1999). We have not explored their use here.

[5] This is a bespoke dataset provided to RAND Europe which includes NPD data, as a result, users should not expect to perfectly replicate sample sizes or results for each analysis.

Multiple imputation – Wave 3 linked data

```
. mi estimate: mean KS2_CVAAPS

Multiple-imputation estimates     Imputations     =         30
Mean estimation                   Number of obs   =      8,882
                                  Average RVI     =     0.1830
                                  Largest FMI     =     0.1561
                                  Complete DF     =       8881
DF adjustment:   Small sample     DF:     min     =   1,043.05
                                          avg     =   1,043.05
Within VCE type:     Analytic             max     =   1,043.05

--------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
  KS2_CVAAPS |   27.12297   .0529624      27.01905    27.22689
--------------------------------------------------------------
```

Complete case analysis – Wave 3 linked data

```
. mean KS2_CVAAPS

Mean estimation                   Number of obs   =      6,606

--------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
  KS2_CVAAPS |   27.19555   .0568015       27.0842     27.3069
--------------------------------------------------------------
```

Inverse probability weighting – Wave 1 linked data

```
. mean KS2_CVAAPS [pweight=ipw_mi]

Mean estimation                   Number of obs   =      8,628

--------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
  KS2_CVAAPS |   26.96419   .0564713      26.85349    27.07489
--------------------------------------------------------------
```

Complete-case analysis – Wave 1 linked data[6]

```
. mean KS2_CVAAPS

Mean estimation                   Number of obs   =      8,795

--------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
  KS2_CVAAPS |   26.87711   .0499222      26.77926    26.97497
--------------------------------------------------------------
```

[6] Note this analysis sample includes the 167 observations identified as having attended a boycott school but for whom a KS2 test score is available, explaining the difference between the two sample sizes here.

Example 2 – Stratified analysis

In stratified analyses, all three sets of estimates are very close as well, but the CI are narrowest from those analyses using MI. In this example analysis, the mean KS2 average point score is stratified by a variable from the NPD, EVERFSM_6_SPR10, which identified whether or not the pupil had received free school meals in any of the six years before the boycott in 2010.

Multiple imputation[7]

```
. mi estimate: mean KS2_CVAAPS, over(EVERFSM_6_SPR10)

Multiple-imputation estimates     Imputations     =         30
Mean estimation                   Number of obs   =     11,069
                                  Average RVI     =     0.1069
                                  Largest FMI     =     0.1054
                                  Complete DF     =      11068
DF adjustment:   Small sample     DF:     min     =   2,086.28
                                          avg     =   2,401.40
Within VCE type:     Analytic             max     =   2,716.53

            0: EVERFSM_6_SPR10 = 0
            1: EVERFSM_6_SPR10 = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
           0 |   27.97773   .0532191      27.87336     28.0821
           1 |   25.10504   .0778648      24.95236    25.25772
--------------------------------------------------------------
```

Inverse probability weighting

```
. mean KS2_CVAAPS [pweight=ipw_mi], over(EVERFSM_6_SPR10)

Mean estimation                   Number of obs   =      8,627

            0: EVERFSM_6_SPR10 = 0
            1: EVERFSM_6_SPR10 = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
KS2_CVAAPS   |
           0 |    28.0402   .0643667      27.91403    28.16638
           1 |   25.28996   .0948068      25.10412    25.47581
--------------------------------------------------------------
```

7 The sample size in this analysis includes all pupils in the MI data set, plus pupils from Wave 1, not included in the MI data set.

Complete-case analysis

```
. mean KS2_CVAAPS, over(EVERFSM_6_SPR10)

Mean estimation                   Number of obs   =      8,793

            0: EVERFSM_6_SPR10 = 0
            1: EVERFSM_6_SPR10 = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
KS2_CVAAPS   |
           0 |     27.957   .0568092      27.84564    28.06836
           1 |   25.07258   .0847919      24.90637    25.23879
--------------------------------------------------------------
```

Example 3 – Regression model predicting KS4 performance

Noting that samples change across models, we see that coefficients for KS2 attainment are very similar from MI, IPW and complete-case models. The coefficients for some smaller ethnic groups change slightly between models, although the confidence intervals (CI) are wide.

Multiple imputation

```
. mi estimate: regress KS4_VAPTSC_PTQ_EE KS2_CVAAPS female EVERFSM_6_SPR10 i.ethnicity

Multiple-imputation estimates                   Imputations       =         30
Linear regression                               Number of obs     =      8,652
                                                Average RVI       =     0.2199
                                                Largest FMI       =     0.3623
                                                Complete DF       =       8642
DF adjustment:   Small sample                   DF:     min       =     219.02
                                                        avg       =   1,297.05
                                                        max       =   2,156.14
Model F test:       Equal FMI                   F(   9, 4156.4)   =    1196.38
Within VCE type:          OLS                   Prob > F          =     0.0000

--------------------------------------------------------------------------------
KS4_VAPTSC_PT~E |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------+---------------------------------------------------------------
     KS2_CVAAPS |    14.3284   .1538595    93.13   0.000     14.02663    14.63017
         female |   17.38137   1.294935    13.42   0.000     14.84192    19.92082
EVERFSM_6_SPR10 |  -29.54015   1.455058   -20.30   0.000    -32.39396   -26.68634
                |
      ethnicity |
 Dual/multiple  |   10.02704    3.47769     2.88   0.004     3.199015    16.85506
        Indian  |   24.57435   4.614246     5.33   0.000     15.52157    33.62713
     Pakistani  |   18.02254   4.197636     4.29   0.000     9.749608    26.29547
   Bangladeshi  |   38.19725   4.700555     8.13   0.000     28.95364    47.44086
         Black  |   18.30163   2.523505     7.25   0.000     13.35173    23.25152
 Chinese/Other  |   40.43219   4.622813     8.75   0.000     31.36519    49.49919
                |
          _cons |  -75.79983   4.452243   -17.03   0.000    -84.53266   -67.06701
--------------------------------------------------------------------------------
```

Inverse probability weighting

```
. regress KS4_VAPTSC_PTQ_EE KS2_CVAAPS female EVERFSM_6_SPR10 i.ethnicity [pweight=ipw_mi]

Linear regression                               Number of obs     =      6,328
                                                F(9, 6318)        =     650.04
                                                Prob > F          =     0.0000
                                                R-squared         =     0.6000
                                                Root MSE          =      56.35

--------------------------------------------------------------------------------
                |               Robust
KS4_VAPTSC_PT~E |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------+---------------------------------------------------------------
     KS2_CVAAPS |   14.54177   .2255889    64.46   0.000     14.09954      14.984
         female |   16.70772   1.649057    10.13   0.000       13.475    19.94043
EVERFSM_6_SPR10 |  -29.51776   2.108577   -14.00   0.000    -33.65129   -25.38424
                |
      ethnicity |
 Dual/multiple  |   8.632638   4.552671     1.90   0.058    -.2921438    17.55742
        Indian  |   23.30247   4.496369     5.18   0.000     14.48806    32.11689
     Pakistani  |   18.68102    5.54366     3.37   0.001     7.813565    29.54848
   Bangladeshi  |   36.37456   4.789797     7.59   0.000     26.98493    45.76419
         Black  |   18.47397   3.479502     5.31   0.000     11.65296    25.29497
 Chinese/Other  |   41.88867   5.247907     7.98   0.000     31.60099    52.17635
                |
          _cons |  -82.33231   6.699903   -12.29   0.000    -95.46639   -69.19822
--------------------------------------------------------------------------------
```

Complete-case analysis

```
. regress KS4_VAPTSC_PTQ_EE KS2_CVAAPS female EVERFSM_6_SPR10 i.ethnicity

      Source |       SS           df       MS      Number of obs   =     6,434
-------------+----------------------------------   F(9, 6424)      =   1095.19
       Model |  31650227.8         9  3516691.98   Prob > F        =    0.0000
    Residual |  20627733.4     6,424  3211.04193   R-squared       =    0.6054
-------------+----------------------------------   Adj R-squared   =    0.6049
       Total |  52277961.2     6,433  8126.52903   Root MSE        =    56.666

--------------------------------------------------------------------------------
KS4_VAPTSC_PT~E |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------+---------------------------------------------------------------
     KS2_CVAAPS |   14.38637   .1664524    86.43   0.000     14.06006    14.71267
         female |   17.32595   1.416475    12.23   0.000     14.54919    20.10271
EVERFSM_6_SPR10 |  -29.02829    1.59709   -18.18   0.000    -32.15912   -25.89746
                |
      ethnicity |
 Dual/multiple  |   12.12312   3.741446     3.24   0.001      4.78864     19.4576
        Indian  |   19.71341    5.11134     3.86   0.000     9.693478    29.73334
     Pakistani  |   15.71236    4.51935     3.48   0.001     6.852929    24.57179
   Bangladeshi  |   38.64049   5.088516     7.59   0.000     28.66531    48.61568
         Black  |   18.29969   2.785853     6.57   0.000     12.83849    23.76089
 Chinese/Other  |   39.44225   4.981985     7.92   0.000     29.67589     49.2086
                |
          _cons |  -77.15237   4.808076   -16.05   0.000     -86.5778   -67.72694
--------------------------------------------------------------------------------
```

Conclusion and final note for analysts

Boycotts of national tests leave gaps in pupils' attainment records and, in the case of LSYPE2, threaten to undermine a large-scale longitudinal study with substantial policy relevance. We believe that this project will allow analysts to make more use of the LSYPE2 dataset, as prior attainment data is something that will be incorporated into even basic analyses of LSYPE2 data. Similarly, our work provides a platform for other researchers to extend the work on missing data presented here.

It is for analysts to decide whether the missing data arising from the boycott will cause difficulties regarding inferences and conclusions for their particular analysis – and to take appropriate steps to deal with these difficulties. Using multiple imputation and inverse probability weighting, we have been able to produce plausible values for KS2 scores (via MI) and analytical weights (via IPW) for pupils missing data due to the boycott, thus giving analysts two options when deciding how to deal with missingness.

References

BARTLETT, J. W. & CARPENTER, J. 2013a. Missing Data - Concepts. LEMMA VLE Module 14, 1-31. URL: http://www.bristol.ac.uk/cmm/learning/online-course/.

BARTLETT, J. W. & CARPENTER, J. 2013b. Missing Data - Stata Pratical. LEMMA VLE Module 14, 1-25. URL: http://www.bristol.ac.uk/cmm/learning/online-course/. .

BEW, P. 2011. Review of Key Stage 2 testing, assessment and accountability https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/180401/DFE-00035-2011.pdf.

DEPARTMENT FOR EDUCATION 2010a. National curriculum assessments at key stage 2 and 3 in England - academic year 2009 to 2010 (provisional) https://www.gov.uk/government/statistics/national-curriculum-assessments-at-key-stage-2-and-3-in-england-academic-year-2009-to-2010-provisional.

DEPARTMENT FOR EDUCATION 2010b. National curriculum assessments at key stage 2: England - academic year 2009 to 2010 (revised) https://www.gov.uk/government/statistics/national-curriculum-assessments-at-key-stage-2-england-academic-year-2009-to-2010-revised.

OFSTED & DEPARTMENT FOR EDUCATION 2015. RAISEonline newsletter for secondary schools. Reports affected by Key Stage 2 2010 boycott. www.raiseonline.org.

SCHARFSTEIN, D. O., ROTNITZKY, A. & ROBINS, J. M. 1999. Adjusting for nonignorable drop-out using semiparametric nonresponse models. *Journal of the American Statistical Association,* 94**,** 1096-1120.

TNS BMRB 2013. Technical report for the 2nd cohort of the Longitudinal Study of Young People in England (LSYPE2): Wave 1.