

Millennium Cohort Study

Linked education administrative
datasets Ofsted

User Guide (Version 1)

August 2025

Contact

Data queries: help@ukdataservice.ac.uk

Questions and feedback about this user guide should be sent to clsdata@ucl.ac.uk.

Authors

Andrew Peters, Aida Galvez-Sanchez, Emla Fitzsimons, Danielle Gomes.

How to cite this guide

Peters, A., Sanchez-Galvez, A., Fitzsimons, E., Gomes, D. (2025) *Millennium Cohort Study: Linked education administrative datasets-Ofsted User Guide (Version 1)*. London: UCL Centre for Longitudinal Studies.

This guide was published in August 2025 by the UCL Centre for Longitudinal Studies.

Data citation and acknowledgements

You should also cite the data and acknowledge CLS following the guidance from <https://cls.ucl.ac.uk/data-access-training/citing-our-data/>

CLS would like to thank Louise Butler from Ofsted for providing information about the Ofsted data currently available.

Centre for Longitudinal Studies

Centre for Longitudinal Studies (CLS)

UCL Social Research Institute

University College London

20 Bedford Way, London WC1H 0AL

cls.ucl.ac.uk

The UCL Centre for Longitudinal Studies (CLS) is an Economic and Social Research Council (ESRC) Resource Centre. It is home to a unique series of UK national cohort studies. It is part of the [UCL Social Research Institute](#), based at the [IOE, UCL's Faculty of Education and Society](#).

This document is available in alternative formats. Please email the Centre for Longitudinal Studies at clsdata@ucl.ac.uk

Contents

About the Millennium Cohort Study	4
1. Introduction	5
2. Consent to linkage	5
3. Data linkage	5
3.1 External Data Provider datasets	5
3.2 Matching strategy.....	6
4. Linked Research data	7
4.1 Licensing and data access.....	7
4.2 List of datasets.....	7
4.3 Identifiers	8
4.4 Dataset description	9
4.5 Data Processing.....	14
4.6 Missing values	15
4.7 Data de-identification	15
4.8 Output Disclosure Control (for controlled data only)	15
Appendix A: Data panel of schools	16

About the Millennium Cohort Study

The Millennium Cohort Study (MCS) is a longitudinal birth cohort study, following a nationally representative sample of approximately 19,000 people born in the UK at the turn of the century.

Through the study, we have captured rich information about the different aspects of cohort members' lives, from birth to childhood and adolescence, and we are continuing to keep up with them now they are adults.

As a multidisciplinary study, MCS is used by researchers working in a wide range of fields. Findings from MCS have influenced policy at the highest level, and today the study remains a vital source of evidence on the major issues affecting young people's lives.

Further details of the data available from the main surveys are available from the CLS website and, in particular, the MCS Guide to the Datasets at

www.cls.ucl.ac.uk/cls-studies/millennium-cohort-study/.

1. Introduction

This guide describes the data linkage of education administrative records from the publicly available Ofsted reporting of school inspections in England to survey data for cohort members in the Millennium Cohort Study (MCS). The main aim of this data linkage is to enhance the research potential of the study, by combining administrative records with the rich information collected in the surveys.

2. Consent to linkage

This data deposit did not require participant consent, as the linkage is performed at the school level rather than the individual level.

3. Data linkage

3.1 External Data Provider datasets

Ofsted (Office for Standards in Education, Children's Services and Skills) is a non-ministerial government department in the United Kingdom that inspects and regulates education and children's services. Its main role is to assess the quality of education provided by schools, colleges, and other educational institutions, as well as the care services for children. Ofsted produces reports that evaluate institutions on various criteria, including teaching quality, leadership, student outcomes, and safeguarding.

Ofsted produces what they refer to as 'state of the nation' data which is a dataset of the latest grade for all 22,000 open schools, whenever their last full inspection was. Some grades are very recent, but some (for schools that are exempt from routine inspections) can be up to ten years ago. For the many local authority schools that have closed to become academies, they publish the details of the new school alongside the details of the last inspection of the old school, until the new one has been inspected. This is because new schools are not inspected for up to three years, so the grade of the predecessor school remains the best view they have on how good the new school may be (In some cases the school will have the same head/teachers/pupils as the old school anyway).

3.2 Matching strategy

The linkage with Ofsted data was carried out at school level by CLS.

The Ofsted results dataset is merged from results data across years between 2005 and 2019. As such, numerous categories are only available for a certain range of years due to changes in reporting of results. Where result categories are confirmed to be the same across different years they have been combined into one variable. Appropriate missing values have been set for variables which do not apply to certain schools/years. Certain categories of result include both 4-point rating scale and Yes/No values. These have been separated and provided as two distinct variables in the available data.

Ofsted makes the data available online. There is a one-off publication that combines the state of the nation data for 31 August 2010 through to 31 August 2017.

<https://www.gov.uk/government/statistics/changes-to-ofsteds-statistical-reporting-of-inspection-outcomes-for-state-funded-schools-an-analysis-of-the-changes>

See the excel file 'Changes to Ofsted's statistical reporting of inspection outcomes for state-funded schools: historical datasets'. The links to the years 2017 and onwards are provided below:

<https://www.gov.uk/government/statistics/state-funded-schools-inspections-and-outcomes-as-at-31-august-2018#full-publication-update-history>

Between November 2017 and January 2018 Ofsted consulted on changes to their statistical reporting of school inspection outcomes. The first release based on this new methodology was in June 2018. For further information on the original consultation and an analysis of the responses see:

<https://www.gov.uk/government/consultations/changes-to-ofsteds-statistical-reporting-of-inspection-outcomes-for-maintained-schools-and-academies>

<https://www.gov.uk/government/statistics/state-funded-schools-inspections-and-outcomes-as-at-31-august-2019>

4. Linked Research data

4.1 Licensing and data access

The linked Ofsted data have been processed by CLS and supplied to the UK Data Service. All data users need to be registered with the UK Data Service and sign the UKDS End User Licence. Details of how to do this are available at ukdataservice.ac.uk/get-data/how-to-access/registration.

Controlled data (Secure Access Licence)

These data are available as controlled data due to their potentially disclosive and/or sensitive nature and must be accessed as controlled data from the UK Data Service SecureLab. Applicants wishing to access this data need to abide by the terms and conditions of the UK Data Service Secure Access licence.

Before gaining access, researchers must make an application outlining the intended analysis and provide a justification as to why this data is requested. Application guidance can be found at ukdataservice.ac.uk/find-data/access-conditions/secure-application-requirements/apply-to-access-non-ons-data/.

Data access will be granted once the form has been reviewed by UK Data Service and approved by the CLS Data Access Committee.

4.2 List of datasets

The Ofsted results across all available years between 2005 and 2019 are contained in the dataset named:

mcs_education_england_ofsted_2005_to_2019_longf_restricted

Result categories differ across years, with certain categories only applicable to specific years.

The school change dataset

(mcs_education_school_change_2004_to_2017_restricted) was constructed to account for changing identifiers of schools between years. While not specifically developed for the Ofsted data, it is available here due to its utility in linking schools to pupils. The dataset allows schools to be linked both forwards and backwards in time

for individual cohort members. School identifiers were derived using a consistent data panel of schools merged to the school database (for years 2004 to 2017) to identify schools that had changed schools between years and the characteristics related to that change i.e. mergers, splits, openings, closures, changes in school status. A detailed description of the data panel is given in Appendix A.

The deposited datasets are described in Table 1.

Table 1: List of deposited datasets

Name of the dataset	Data source	Coverage	Contents
mcs_education_england_ofsted_2005_to_2019_longf_restricted	Ofsted	2005-2019	Report of schools inspections and outcomes
mcs_education_school_change_2004_to_2017_restricted	CLS	2004-2017	Changes in school URNs across years for individual Cohort members

4.3 Identifiers

The linked Ofsted data are identified via the anonymised URN variable URN_ANON, which can currently be merged with the anonymised URN in other linked education data.

The School change dataset is linked using MCSID and the only way to match Ofsted results to survey data at the time of deposit is to match the URN_ANON via MCSID by year. Individual schools in the school change dataset are identified by the URN_ANON_XX, where XX is the year the pupil was at that school between 2004 and 2017. This can be matched to the URN_ANON of an appropriate academic year of Ofsted results.

Using Ofsted data alongside other MCS linked education data

The linked Ofsted dataset is provided at school level. Each school is identified with an anonymised version of the school Unique reference Number (URN). This anonymised URN is also used in other MCS linked education datasets, such as linkage to NPD [data](#)¹, thus enabling researchers to combine the different education

¹ Millennium Cohort Study: Linked Education Administrative Datasets (National Pupil Database), England: Secure Access <https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=8481>

data. For example, the Ofsted dataset can be used alongside the MCS NPD linked data by using the anonymised URN as the matching key.

Using Ofsted data alongside other MCS survey data

At the moment, the linked Ofsted school data can only be matched to the survey data when the participant was previously matched as part of the NPD linkages. CLS will soon be updating the school URNs collected from participants during the different MCS survey sweeps, to ensure they are fully compatible with the URNs obtained from DfE/NPD or Ofsted.

The school change dataset can be used to match the school of participants not presented in the linked education dataset to the Ofsted school by identifying the cohort member in the school change dataset and then their school. In order to use Ofsted results alongside survey data, school URNs would need to be matched to an MCSID relative to the date of interview in a sweep first, then to an appropriate Ofsted inspection year for that school. For example, schools attended during the age 17 sweep can be found under URN_ANON_17 in the school change data. From this, the URN can be used to find the school level results in the Ofsted dataset, and the MCSID/CNUM can be linked to the age 17 survey data. This is for linkage to survey data where a participant has given their consent to education linkages but have not been linked via the DfE.

4.4 Dataset description

Variable Order

The full dataset of Ofsted results contains the following variables listed in table 2.

Table 2: Variables in Ofsted results dataset

Variable Name	Variable Label
urn_anon	Anonymous URN
academicyear	The academic year in which the inspection took place
ofsted_region	The Ofsted region in which a school resides (North East, Yorkshire and the Humber are combined into a single region)
school_type	How the provider is categorised in the Department for Education's Get Information About Schools database
ofsted_phase	Ofsted phase of education
overall_eff	Overall effectiveness
inadequate_judgement	Inspection judgements for inadequate schools

Variable Name	Variable Label
teaching_quality	Quality of teaching, learning and assessment
leadership_eff	Effectiveness of leadership and management
earlyyears_prov	Early years provision
sixthform_prov	Sixth form provision (2005-2015)
pupil_achieve_agg	Pupils' achievement (aggregated) (2005-2015)
pupil_achieve	Achievement of pupils (2005-2015)
pupil_achieve_how	How well do pupils achieve? (2005-2015)
pupil_behaviour	Behaviour and safety of pupils (2005-2015)
overall_eff_res	Overall effectiveness of residential experience (aggregated) (2005-2015)
boarding_prov	The effectiveness of the boarding provision (2005-2012)
resident_experience	Overall effectiveness of residential experience (2012-2015)
resident_pupil_outcome	Outcomes for residential pupils (2012-2015)
resident_prov	Quality of residential provision and care (2012-2015)
resident_pupil_safety	Residential pupils safety (2012-2015)
resident_leadership	Leadership and management of the residential provision (2012-2015)
partnerships_wellbeing	The effectiveness of partnerships in promoting learning and well-being (2005-2015)
sustained_improvement_yn	The school's capacity for sustained improvement (2005-2015)
sustained_improvement	The school's capacity for sustained improvement (2005-2015)
pupil_ind_outcome	Outcomes for individuals and groups of pupils (2005-2015)
pupil_attainment	Pupils' attainment (2005-2015)
pupil_progress	The quality of pupils' learning and their progress (2005-2015)
earlyyears_outcome	Outcomes for children in the Early Years Foundation Stage (2005-2015)
sixthform_outcome	Outcomes for students in the sixth form (2005-2015)
sixthform_prov_quality	The quality of provision in the sixth form (2005-2015)
special_needs	The quality of learning for pupils with special educational needs and/or disabilities and their progress (2005-2015)
pupil_cultural_ext	The extent of pupils' spiritual, moral, social and cultural development (2005-2015)
pupil_lifestyle_ext	The extent to which pupils adopt healthy lifestyles (2005-2015)
pupil_feelsafe_ext	The extent to which pupils feel safe (2005-2015)
pupil_attendance	Pupils' attendance (2005-2015)
pupil_contribute_ext	The extent to which pupils contribute to the school and wider community (2005-2015)
pupil_develop_ext	The extent to which pupils develop workplace and other skills that will contribute to their future economic well-being (2005-2015)
curriculum_ext	The extent to which the curriculum meets pupils' needs, including, where relevant, through partnerships (2005-2015)
guidance_eff	The effectiveness of care, guidance and support (2005-2015)

Variable Name	Variable Label
assessment_support	The use of assessment to support learning (2005-2015)
earlyyears_prov_quality	The quality of provision in the Early Years Foundation Stage (2005-2015)
earlyyears_leadership_eff	The effectiveness of leadership and management of the Early Years Foundation Stage (2005-2015)
sixthform_leadership_eff	The effectiveness of leadership and management of the sixth form (2005-2015)
equality_eff	The effectiveness with which the school promotes equality of opportunity and tackles discrimination (2005-2015)
community_eff	The effectiveness with which the school promotes community cohesion (2005-2015)
resources_eff	The effectiveness with which the school deploys resources to achieve value for money (2005-2015)
governing_eff	The effectiveness of the governing body in challenging and supporting the school so that weaknesses are tackled decisively (2005-2015)
safeguard	Is safeguarding effective
safeguard_eff	The effectiveness of safeguarding procedures (2005-2015)
engagement_eff	The effectiveness of the school's engagement with parents and carers (2005-2015)
leadership_teaching	The leadership and management of teaching and learning (2005-2015)
category_concern	Category of concern (2015-2019)
school_open_date	Date the school opened (2015-2019)
admissions_policy	The admissions policy of the school (2015-2019)
religious_character	Designated religious character of the school. This is sourced from Get Information About Schools, maintained by the Department for Education (2015-2019)
religious_ethos	Religious ethos declared by the school. This is sourced from Get Information About Schools, maintained by the Department for Education (2015-2019)
faith_grouping	Groups faith schools in to these five categories: Christian, Jewish, Muslim, other faith or non-faith using either religious character or religious ethos (2015-2019)
warnings_issued	Number of warning notices issued in academic year (2015-2019)
income_deprivation	Deprivation is based on the Income Deprivation Affecting Children Index (IDACI) 2015 (2015-2019)
total_pupils	The number of pupils on roll, from the Department for Education's Get Information About Schools database which uses data from the school census in January 2019 (2015-2019)
inspection_type	The type of inspection conducted (2015-2019)
inspection_type_grouping	Inspection type grouping (2015-2019)
event_type_grouping	Event type grouping (2015-2019)
inspection_convert	Did the latest S8 inspection convert to a full inspection? (2015-2019)
notconvert_outcome	Outcomes for S8 inspections that did not convert (2015-2019)

Variable Name	Variable Label
pupil_outcome	Outcomes for pupils (2015-2019)
development_welfare	Personal development, behaviour and welfare (2015-2019)
sixthform	Whether the schools has sixth form provision (where provided) (2015-2019)
sixthform_study	16 to 19 study programmes (where applicable) (2015-2019)
prev_urn_anon	Anonymous predecessor URN of schools converting to academy
prev_inspection	Does the previous inspection relate to the school in its current form? (2015-2019)
prev_school_type	School type at time of previous inspection (2015-2019)
prev_overall_eff	Previous overall effectiveness (2015-2019)
prev_category_concern	Previous category of concern (2015-2019)
prev_leadership_eff	Previous effectiveness of leadership and management (2015-2019)
prev_teaching_quality	Previous quality of teaching, learning and assessment (2015-2019)
prev_development_welfare	Previous personal development, behaviour and welfare (2015-2019)
prev_pupil_outcome	Previous outcomes for pupils (2015-2019)
prev_earlyyears_prov	Previous early years provision (where applicable) (2015-2019)
prev_sixthform_study	Previous 16 to 19 study programmes (where applicable) (2015-2019)
prev_safeguard	Previous is safeguarding effective? (2015-2019)

The contents of the school change dataset are listed in Table 3. This dataset is derived by merging the school census information with the school panel information and retrieval of information where the URN has changed between successive academic years. These have now been updated using a download of establishment data from 'Get information about schools' (see Appendix A). Please note that the reporting period of the school panel is 1st September to 31st August. This differs from the school census in that the census date for the latter is the third week of January. Hence there may be cases not reported within the period after this date.

Table 3: Variables in School change dataset

Variable Name	Variable Label
mcsid	MCS Research ID - Anonymised Family/Household identifier
cnum	Cohort member number within an MCS family
urn_anon_04	Unique school reference number (anonymised)
in04	Flag: present in year 2004
urn_anon_05	Unique school reference number (anonymised)
in05	Flag: present in year 2005
same_school_2005	Whether at same school since previous year
urn_anon_06	Unique school reference number (anonymised)

Variable Name	Variable Label
in06	Flag: present in year 2006
same_school_2006	Whether at same school since previous year
urn_anon_07	Unique school reference number (anonymised)
in07	Flag: present in year 2007
same_school_2007	Whether at same school since previous year
urn_anon_08	Unique school reference number (anonymised)
in08	Flag: present in year 2008
same_school_2008	Whether at same school since previous year
urn_anon_09	Unique school reference number (anonymised)
in09	Flag: present in year 2009
same_school_2009	Whether at same school since previous year
urn_anon_10	Unique school reference number (anonymised)
in10	Flag: present in year 2010
same_school_2010	Whether at same school since previous year
urn_anon_11	Unique school reference number (anonymised)
in11	Flag: present in year 2011
same_school_2011	Whether at same school since previous year
urn_anon_12	Unique school reference number (anonymised)
in12	Flag: present in year 2012
same_school_2012	Whether at same school since previous year
urn_anon_13	Unique school reference number (anonymised)
in13	Flag: present in year 2013
same_school_2013	Whether at same school since previous year
urn_anon_14	Unique school reference number (anonymised)
in14	Flag: present in year 2014
same_school_2014	Whether at same school since previous year
urn_anon_15	Unique school reference number (anonymised)
in15	Flag: present in year 2015
same_school_2015	Whether at same school since previous year
urn_anon_16	Unique school reference number (anonymised)
in16	Flag: present in year 2016
same_school_2016	Whether at same school since previous year
urn_anon_17	Unique school reference number (anonymised)
in17	Flag: present in year 2017
same_school_2017	Whether at same school since previous year

The variable same_school_20xx indicates whether the school attended is the same as in previous year 20xx-1. URNs are compared for 20xx and 20xx-1 and a flag is

given if URNs exist for both years. This is derived from the consistent data panel of schools (see Appendix A for a breakdown of the derivation process).

Variable Names

Variable Names have been created for the Ofsted results data intended to be as consistent and informative as possible. They are designed so the prefix refers to the area of inspection (e.g. pupil, leadership), while the suffix, where applicable, regards the additional details (e.g. eff(ectiveness), prov(ision)). Variables beginning 'prev_' refer to the previous inspection at that school

Variable Labels

Labels are mostly identical to those used in the results of Ofsted assessments. The only changes made are to make wording more consistent across the data.

The other addition is year ranges contained in variable labels describe the range of years the results are available for. When there are no year ranges, results are available for all years (2005-2019).

Value labels

The main Ofsted results data are on a 4-point scale with the following values:

1. Outstanding
2. Good
3. Requires improvement/Satisfactory
4. Inadequate

Additional Ofsted codes included where there was an inspection but no rating:

8. Does not apply
9. No judgement

4.5 Data Processing

Ofsted results for the full range of years have been combined into a single dataset where each row contains Ofsted results for a school in a single academic year. Where appropriate, missing values have been assigned where no data was available.

4.6 Missing values

The following missing values are used in the Ofsted results data:

- 9. Unknown
- 8. Not recorded
- 1. Not applicable

4.7 Data de-identification

As well as actual URN, specific details about each school such as name is not made available in order to prevent identification of the schools cohort members attended.

4.8 Output Disclosure Control (for controlled data only)

UK Data Service requirements

The linked MCS Ofsted data are available via the UK Data Service Secure Lab. Access to this data is controlled and the UK Data Service will always perform a certain level of disclosure control on the outputs generated by researchers, as outlined in the SDC Handbook which can be downloaded from:

securedatagroup.org/sdc-handbook/

The two UK Data Service Secure Lab rules of thumb that will be applied to all outputs are:

- Threshold rule: No cells should contain less than 10 observations
- Dominance rule: No observation should dominate the data to a huge extent

Appendix A: Data panel of schools

Overview

The school database is designed to link school identifiers over time, to aid researchers. Examples of uses include creating a panel of academic attainment (or other characteristics) at the school-level, tracking schools that convert to academy (or change status in another way) and linking to external sources of school data. The database is from the academic year 1998/1999 to 2016/2017. 2016/2017 is the latest year of school census data currently available. The panel was initially created by Ellen Greaves with Lorraine Dearden (for academic years 2005/2006 to 2013/2014). The three most recent cohorts were added by Ellen Greaves.

Structure of the panel

Each row contains a URN, URN and LAESTAB for each academic year. School names for the most recent years are included. More detail follows.

URN: each URN that ever appears in the school database will have a row in the school database. This is so that researchers can use this URN variable to link to their data. For example, the following school changes from 100222 to 121792 in 2005/2006.

urn	...	urn_2004	urn_2005	urn_2006	urn_2007	urn_2008	...
100222		100222	100222	121792	121792	121792	
121792		100222	100222	121792	121792	121792	

Where a school merges, the urn of the merged school will be repeated to show the two (or more) linked predecessors. For example, the following schools merge in 2001/2002.

urn	...	urn_2000	urn_2001	urn_2002	urn_2003	urn_2004	...
100123		100123	100123	132827	132827	132827	
100124		100124	100124	132827	132827	132827	
132827		100123	100123	132827	132827	132827	

urn	...	urn_2000	urn_2001	urn_2002	urn_2003	urn_2004	...
132827		100124	100124	132827	132827	132827	

Where a school splits, the urn of the school that splits will be repeated to show the linked successor. For example, the following schools split in 2006/2007.

urn	...	urn_2004	urn_2005	urn_2006	urn_2007	urn_2008	...
101150		101150	101150	101150	130912	130912	
101150		101150	101150	101150	131262	131262	
130912		101150	101150	101150	130912	130912	
131262		101150	101150	101150	131262	131262	

Where a school opens without a predecessor or closes without a successor there will be missing values whenever the school is not open. The examples below show one school that opens and one school that closes in 2008/2009.

urn	...	urn_2006	urn_2007	urn_2008	urn_2009	urn_2010	...
102707		102707	102707	102707	.	.	
135148		.	.	.	135148	135148	

Using the school change dataset

The school database can be used in a number of ways.

Look-up

The simplest is to lookup individual or a small number of schools. For example `-br if inlist(urn, 101150)` - would show the URNs and laestabs over time for the school in example three. This might be useful for schools in a small scale pilot, or to check some existing data.

Tracking schools that change status

Changes in URN or laestab can be easily identified in the school panel (see example code below). Whether this change relates to a change in status (for example academy conversion) could be validated by linking in data on school type from the most recent version of Edubase. See following section for linking to other sources.

Code for changes:

```
gen change_2006 = urn_2006 != urn_2005
replace change_2006 = . if urn_2006 == . | urn_2005 == .
label var change_2006 "School changed URN between 2005 and 2006"
```

Linking to external sources of data

Researchers often have to link sources of data to create their desired dataset for analysis. A common example is linking Ofsted grades to school performance data or the National Pupil Database. As school identifiers can change mid-way through the year this can be problematic. For example, the URN in the Ofsted data may be from October when the school was inspected, but the URN in the school performance data may be from June, after the school had changed status. The school database can be used to attempt to link on adjacent URNs, for example.

Code for linking to external sources of data:

1. Create unique URN in the external data, by dropping duplicates (collapsing to school level from pupil level) or reshaping.
2. Link main dataset to the school database. Keep adjacent URNs. For example if the external data are from 2014 keep urn_2013 and urn_2015 in addition to urn_2014.
3. Link main dataset to the external data using urn_2014. If unmerged, link using urn_2013 and repeat using urn_2015.

Creating a school panel

A panel dataset is common in applied research to analyse changes over time or account for time invariant school characteristics (school fixed effects). Creating such a panel is complicated by changes in the URN over time. The following code shows one way to create a panel. The initial data is at the individual level with multiple academic years.

1. collapse pupil-level data to school level URNs

```
keep urn censusyear
duplicates drop
```

```
sort urn censusyear
save "$filepath/urn_censusyear.dta", replace
```

2. link school-level URNs to school database and extract the latest URN. This is done year by year so that can use a -1:m merge-in step 4. Save this link.

```
forvalues year = 2010/2015 {
  preserve
  keep if censusyear == `year'
  merge 1:m urn using "$filepath\data_panel_2017.dta"
  tab urn if _me == 1
  drop if _me == 2
  gen urn_newest = urn_2017
  foreach yearrepeat in 2016 2015 2014 2013 2012 2011 2010 {
    replace urn_newest = urn_`yearrepeat' if urn_newest == .
  }
  keep urn censusyear urn_newest
  duplicates drop
  save "$filepath\urn_newest_`year'.dta", replace
  restore
}
```

3. Append all years together. This is so can use a -1:m merge-in step 4.

```
use "$filepath\urn_newest_2010.dta", clear
forvalues year = 2011/2015 {
  append using "$filepath\urn_newest_`year'.dta"
}
save "$filepath\urn_newest.dta", replace
```

4. merge the link to newest URN back into the existing data to create panel

```
use "$filepath\urn_censusyear.dta", clear

merge 1:m urn censusyear using "$filepath\urn_newest.dta"
assert _me == 3
drop _merge
```

Example of new data

censusyear	urn	urn_newest
2010	108217	133758
2011	108217	133758
2012	133758	133758
2013	133758	133758
2014	133758	133758
2015	133758	133758

5. Check which URN is correct for your context for schools that split.

For example, school 137798 (Bluecoat Academy) splits into 145145 (Bluecoat Primary Academy) and 145146 (Bluecoat Wollaton Academy) in 2017. Whether the panel should track attainment for Bluecoat Primary Academy or Bluecoat Wollaton Academy, or whether to drop this school entirely from the sample is a decision for the researcher.

(Note that Bluecoat Academy had converted from The Nottingham Bluecoat School and Technology College (122894) in 2012.)

censusyear	urn	urn_newest
2010	122894	145145
2011	122894	145145
2012	137798	145145
2013	137798	145145
2014	137798	145145
2015	137798	145145
2010	122894	145146
2011	122894	145146
2012	137798	145146
2013	137798	145146
2014	137798	145146
2015	137798	145146