Millennium Cohort Study

A Guide to the Linked Education Administrative Datasets

Jon Johnson & Rachel Rosenberg

User guide to the data (Second Edition)

July 2013

Table of Contents

ACKNOWLEDGEMENTS

PREFACE

This document has been prepared to accompany the deposit of data from the fourth survey of the Millennium Cohort Study at age 7 (MCS4) with the UK Data Service at the University of Essex.

The first edition of the Guide to the Linked Education data covers linkage for schools in England.

The second edition also covers linkage in Scotland and Wales.

BACKGROUND

1. Introduction

The Millennium Cohort Study (MCS) is the fourth of Britain's world-renowned national longitudinal birth cohort studies. Each follows a large sample of individuals born over a limited period of time through the course of their lives, charting the effects of events and circumstances in early life on outcomes and achievements later on. They show how histories of health, wealth, education, family and employment are interwoven for individuals and vary between them.

The MCS Study offers large-scale information about the New Century's children, and the families who are bringing them up, for the four countries of the United Kingdom. Its first sweep, carried out during 2001-2002, laid the foundations for a major new longitudinal research resource, taking a new "year long" cohort of around 19,000 babies. In England and Wales they were born over the 12 months starting in September 2000, and over 13½ months from late November 2000 in Scotland and Northern Ireland. Information was collected from parents when the babies were aged nine months. The sample design allowed for disproportionate representation of families living in areas of child poverty, and in areas of England with high ethnic minority populations. The first survey recorded the circumstances of pregnancy and birth, as well as those of the all-important early months of life, and the social and economic background of the family into which the children have been born.

This multidisciplinary baseline data reveal the diversity of starting points from which these 'Children of the New Century' have set out. The second survey data add a longitudinal element to the dataset and will allow researchers for the first time to chart the changing circumstances of children and their families, and relate age 3 outcomes to earlier circumstances and experiences. The third survey visited the families when the children reached age 5 and had started primary school.

The fourth returned to the families in 2008 when the children were age 7 and collected information from parents and carer's about schooling, the child's development and parental assessments of cognitive ability and behaviour. This data also included self reported information about the school.

Further details of the data available from the main survey are available from the CLS website and in particular the *MCS Guide to the Datasets* at www.cls.ioe.ac.uk/mcs.

2. Consent to Education Data Linkage

During the fourth sweep, the parents / carer's were asked for consent to link to the child's education records. Details of the form and the wording of the consent is reproduced in the MCS4 Technical Report Appendices Section A16, consent form 3.

Table 1: Consent Outcomes

Country of interview	Yes	No	No form	Total
England	8447	293	259	8999
Wales	1898	67	30	1995
Scotland	1536	62	51	1649
Northern Ireland	1288	73	38	1399
Total	13169	495	378	14042

3. Education Data Linkage

During the fourth sweep, the parents / carers were asked the name of the school the cohort child was attending and a lookup table was used by the field work agency to confirm the name and location of the school. This produced a dataset which included the URN of the school, the Local Education Authority and the school number used by the Department of Education in England, the Welsh Assembly, the Scottish Parliament and NISRA in Northern Ireland.

a. England

After some cleaning at CLS to correct incorrectly assigned schools (for example, previous school which has since closed and been renamed), CLS prepared a file of 8444 cases to send to the Department for Education in England, (this excluded 3 cases where the interview was in England but the school reported was in another country.

The file contained the following information on the cohort child: Name (first and last), Sex, Date of Birth, Postcode of residence at interview, name of school,
URN of school of current school, Local Education Authority of current school,School Number of current school

The NPD database team matched the data on the following basis.
a.   Direct match on all variables
b.   Direct match of 4 of the 5 variables or 3 of 4 etc
c.   Fuzzy matches on names – usually that the first 3 or 4 characters are the same – again all manually checked.
d.   Any other combinations we think are safe based on the data supplied

b. Scotland

After some cleaning at CLS to correct incorrectly assigned schools (for example, previous school which has since closed and been renamed), CLS prepared a file of 1536 cases to send to the Education Analytical Services division within the Learning and Justice Directorate of the Scottish Government,

c. Wales

After some cleaning at CLS to correct incorrectly assigned schools (for example, previous school which has since closed and been renamed), CLS prepared a file **of** 1890 cases to send to the NHS Wales Informatics Centre (NWIS). (this excluded 8 cases where the interview was in Wales but the school reported was in another country. They matched the cohort members to NHS number and this was then sent to SAIL as an anonymised number which was then used to linked to the Welsh National Pupil Database.

4. Linkage outcome

Table 2 shows the outcome of the linkages carried out.

Table 2: Linkage Outcome

	England	Wales	Scotland	N. Ireland
No of consents	8447	1898	1536	1288
No of cases sent for linkage	8444	1890	1536	n/a
No. of cases successfully linked	6841	1890 (1735)	1407	n/a
% of cases successfully linked	81%	100% (92%)	92%	n/a

NB The number of cases in Wales linked to NHS number was 100%, however, due to non-response by schools to the Census, only 1735 were able to provide individual data.

5. Data variables

Table 3 lists the variables obtained from the linkage to the National Pupil Database (NPD).

Those labelled KS1_ are from the KS1 data held in the NPD. The remaining variables are from the school census 2010. In addition to these are LEANON and SCHOOLANON, which are anonymised Local Education Authority and School Number, to allow comparison of results across LEA and school.

Table 3: Data Variables (England)

File name: mcs4_ks1_plasc.sav

Variable	Description	Source
MCSID	MCS Research ID	
CNUM	Cohort Child Number	
ACADYR	KS1_ACADYR Academic year	NPD
AGESTART	KS1_AGE_START Age of pupil at start of the academic year (in full years)	NPD
TOECODE	KS1_ToE_CODE Type of establishment code taken from Edubase	NPD
NFTYPE	KS1_NFTYPE Institution type generated by AAT - numeric	NPD
MMSCH	KS1_MMSCH Maintained Mainstream school (including CTCs and Academies)	NPD
MMSCH2	KS1_MMSCH2 Maintained Mainstream school (excluding CTCs and Academies)	NPD
MSCH	KS1_MSCH Maintained school (incluing CTCs and Academies)	NPD
MSCH2	KS1_MSCH2 Maintained school (excluding CTCs and Academies)	NPD
MOB1	KS1_MOB1 Pupil joined in latest academic year (year 2)	NPD
MOB2	KS1_MOB2 Pupil joined in previous academic year (year 1)	NPD
APS	KS1_APS Average attainment point score	NPD
SPKLISTN	KS1_SPEAKANDLISTEN National Curriculum level (Teacher Assessment)	NPD
LEVLSPLI	KS1_LEVLSPLI Achieved Level 1 or below in KS1 Speaking + Listening	NPD
EVXSPLI	KS1_LEVXSPLI Achieved Level 2 or above in KS1 Speaking + Listening	NPD
EVAXSPLI	KS1_LEVAXSPLI Achieved Level 3 or abovein KS1 Speaking + Listening	NPD
READING	KS1_READING National Curriculum level awarded for reading	NPD
READPNTS	KS1_READPOINTS Reading attainment point score	NPD
EVLREAD	KS1_LEVLREAD Achieved Level 1 or below in KS1 Reading	NPD
LEVXREAD	KS1_LEVXREAD Achieved Level 2 or above in KS1 Reading	NPD
LV2BREAD	KS1_LEV2BREAD Achieved Level 2B or above in KS1 Reading	NPD
LVAXREAD	KS1_LEVAXREA Achieved Level 3 or above in KS1 Reading	NPD
WRITING	KS1_WRITING National Curriculum level awarded for writing	NPD
WRITPNTS	KS1_WRITPOINTS Writing attainment point score	NPD
LVLWRIT	KS1_LEVLWRIT Achieved Level 1 or below in KS1 Writing	NPD
LVXWRIT	KS1_LEVXWRIT Achieved Level 2 or above in KS1 Writing	NPD

LV2BWRIT	KS1_LEV2BWRIT Achieved Level 2B or above in KS1 Writing	NPD
LVAXWRIT	KS1_LEVAXWRIT Achieved Level 3 or above in KS1 Writing	NPD
RDWRIT	KS1_READWRIT English Average Level (derived from Reading and Writing)	NPD
RDWRPNTS	KS1_READWRITPOINTS Reading and writing attainment point score	NPD
MATHS	KS1_MATHS National Curriculum level awarded for Maths	NPD
MPOINTS	KS1_MATPOINTS Maths attainment point score	NPD
LVLMAT	KS1_LEVLMAT Achieved Level 1 or below in KS1 Maths	NPD
LVXMAT	KS1_LEVXMAT Achieved Level 2 or above in KS1 Maths	NPD
LV2BMAT	KS1_LEV2BMAT Achieved Level 2B or above in KS1 Maths	NPD
LVAXMAT	KS1_LEVAXMAT Achieved Level 3 or above in KS1 Maths	NPD
SCXPINVT	KS1_SCIEXPINVEST National Curriculum level awarded for scientific enquiry	NPD
SCLIFPRT	KS1_SCILIFPROT National Curriculum level awarded for life and living processes	NPD
SCMPROP	KS1_SCIMATPROP National Curriculum level awarded for materials and properties	NPD
SCIPPROC	KS1_SCIPHYSPROC National Curriculum level awarded for physical processes	NPD
SCIENCE	KS1_SCIENCE National Curriculum level awarded for overall Science	NPD
SCPNTS	KS1_SCIPOINTS Overall Science attainment point score	NPD
LVLSCI	KS1_LEVLSCI Achieved Level 1 or below in KS1 overall Science	NPD
LVXSCI	KS1_LEVXSCI Achieved Level 2 or above in KS1 overall Science	NPD
LVAXSCI	KS1_LEVAXSCI Achieved Level 3 or above in KS1 overall Science	NPD
LEAANON	LEA / Board anonymised	DFE
SCHOOLANON	School no, anonymised	DFE
RECSTAT	Record Status	Census
ACAYEAR	AcademicYear	Census
CENTERM	CensusTerm	Census
ONROLL	OnRoll	Census
ETHNMIN	Pupils minor ethnic group based on ethnic code	Census
ETHNMAJ	Pupils major ethnic group based on ethnic code	Census
ETHNSRC	Source of pupils ethnicity code	Census
FSMELIG	Pupil eligible for free school meals	Census
CONNEX	Do parents consent to pupil data being shared with Connexions	Census

LANGMIN	Pupils minor language group based on language code	Census
LANGMAJ	Pupils major language group based on language code	Census
GANDTIND	Indicates if the pupil is in the Gifted and Talented cohort for the school	Census
MTRAVEL	Indicates the usual mode of travel used by the pupil	Census
ENRLSTAT	Indicates enrolment status of pupil	Census
TYPCLASS	Type of Class	Census
ENTRYDTE	Date of entry to current school	Census
PARTTIME	Indicates if the pupil is part time or not	Census
BOARDER	Indicates if the pupil is a boarder or not	Census
NCYRACT	The year group in which the pupil is taught	Census
SENPROV	Provision types under the SEN Code of Practice	Census
SENPROV2	Pupils major SEN provision group based on SEN provision code	Census
SENUIND	Pupil with SEN in a mainstream school is a member of a SEN unit	Census
RPROVIND	Pupil with SEN in a mainstream school is a member of a resourced provision	Census
MOBIND	Pupils postcode has changed from previous Spring Census (Spring only)	Census
DCURRSCH	Distance in miles from pupils postcode to current school (Spring only)	Census
DNEARSCH	Distance in miles from pupils postcode to nearest suitable school (Spring)	Census

Table 4: Data Variables (Scotland)

File name: mcs4_scoted_final.sav

Variable	Description	Source
MCSID	MCS Research ID	
CMNUM00	Cohort Child Number	
MATCHING	Uniquely identifiable status	
DIFFSCHL	Whether different school listed on the pupil census (1=yes)	
POSSATT	Possible attendance of student in half days	AA
REVATT	Actual attendance of student in half days	AA
TEMPEXCL	Temporary exclusions	AA
STUSTAGE	Stage of the student	PC
ASNIND	Whether additional support needed (1=yes)	PC
ETHNICBG	Ethnic background	PC
FSMREG	Whether registered for free school meals (1=yes)	PC
LEVELENG	Level of English	PC
FTPRP	Percentage of total FTE that are primary teachers	PC
FTSEP	Percentage of total FTE that are secondary teachers	PC
FTSPP	Percentage of total FTE that are special teachers	PC
PPRIMFSM	Percentage primary free school meal entitled	SMS
PSPECFSM	Percentage special free school meal entitled	SMS
PSECOFSM	Percentage secondary free school meal entitled	SMS
ASNPRIMP	Percentage with additional support needs at primary	PC
ASNSPECP	Percentage with additional support needs at special	PC
EALPRIMP	Percentage with English as another language at primary	PC
EALSPECP	Percentage with English as another language at special	PC
MALEP	Percentage of students male	PC
FEMALEP	Percentage of students female	PC
ASNP	Percentage of students with additional support needs	PC

Variable	Description	Source
EALP	Percentage of students with English as another language	PC
SECTOR	Which sector - Primary,special or secondary	PC
PETHNIC	Percentage of Ethnic Minorities	PC
LEAANON	LEA / Board anonymised	
SCHOOLANON	School no, anonymised	

Key to source:
AA (Attendance and Absence for year 2008-2009)
PC (Pupil Census 2008 for year 2008-2009)
SMS (School Meals Survey)

Table 5: Data Variables (Wales)

File name: mcs4_walesed_final.sav

Variable	Description	Source
MCSID	MCS Research ID	
CMNUM00	Cohort Child Number	
ALF_STS	ALF status code in education dataset	PC
ACYEAR	Academic year	PC
FSMELIG	Free school meal eligibility	PC
ENSUB	English teacher assessment level	PC
CYSUB	Welsh teacher assessment level	PC
MASUB	Maths teacher assessment level	PC
SCSUB	Science teacher assessment level	PC
CSI	Did pupil achieve level 2 or above in (ENSUB or CYSUB) and MASUB?	PC
LANG	Language teacher assessment level	PC
LEAANON	LEA / Board anonymised	
SCHOOLANON	School no, anonymised	

Key to source:
PC (Pupil Census 2008 for year 2008-2009)