The Age 17 Sweep – Millennium Cohort Study Webinar
3 December 2020

Centre for Longitudinal Studies, UCL Social Research Institute
MCS7 - Data
Vilma Agalioti-Sgompou, Centre for Longitudinal Studies
Resources

Millennium Cohort Study: Age 17 Sweep (MCS7)
User guide
1st edition, August 2020

Millennium Cohort Study: Seventh Sweep (MCS7)
Technical Report
Prepared for the Centre for Longitudinal Studies, UCL Institute of Education
1st edition, August 2020

Millennium Cohort Study: Age 17 Sweep (MCS7)
Derived variables user guide
1st edition, August 2020
Resources at cls.ucl.ac.uk

MCS Age 17 Sweep

Documentation

User guides

Questionnaires

Technical reports

MCS Age 17 Young Person Online Questionnaire

Young person online (CAWI) questionnaire for MCS 7.
Date published: 04/02/2020
PDF: 938.1 KB

MCS Age 17 Young Person Interview

Young person interview for MCS 7.
Date published: 30/01/2020
PDF: 1.17 MB

MCS Age 17 Parent Online Questionnaire

Paper representation of the parent online questionnaire (CAWI) for MCS 7.
Date published: 30/01/2020
PDF: 1.1 MB
### Key identifiers in MCS

<table>
<thead>
<tr>
<th>MCSID</th>
<th>Family identifier</th>
<th>every family has the same ID across sweeps</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>everyone in the family has the same family ID (Cohort Members, parents, relatives and other people living in the household)</td>
</tr>
<tr>
<td>CNUM</td>
<td>Cohort members, 1, 2 (twins) or 3 (triplets)</td>
<td></td>
</tr>
<tr>
<td>PNUM</td>
<td>Person number, for everyone else in the family apart from cohort members: parents, siblings, grandparents, etc</td>
<td></td>
</tr>
<tr>
<td>ELIG</td>
<td>Eligibility for parent interview: a person selected for <strong>Main</strong> respondent and a person selected for <strong>Partner</strong> respondent</td>
<td></td>
</tr>
<tr>
<td>RESP</td>
<td>Response outcome for the person eligible for parent interview (ELIG): <strong>Main</strong>, <strong>Partner</strong> respondent or No interview (in cases of refusal).</td>
<td></td>
</tr>
</tbody>
</table>
Variable naming conventions

Section of the questionnaire / respondent

Sweep

G

X

X

X

X

X

X

XX

Question code in the questionnaire (usually 4 characters, e.g. ADSA)

For variables from multi-coded questions or loops

Variable naming conventions:

H for Household module
P for Parent interview/Parent respondent
C for Cohort Member responds/ Child provides information about him/herself
D for Derived variables
The samples of MCS7 Age 17

Core (face-to-face survey)
Household Interview

CM CAPI
CM CASI
CM CAWI

Parent-reported SDQ
Parent CAWI
Cognitive Assessment
Physical Measurements

Participants in both: Core & Boost

Boost (online follow-up, web survey)

CM CAPI (some questions)
CM CAWI (some questions)
CM CASI (some questions)

The G_OUT_* variables of the mcs7_hhgrid cover the different questionnaire sections and sample.
mcs7_hhgrid outcome variables (G_OUT_*)

- The outcome variables indicate whether a specific respondent (CM or Parent/Carer) is productive in different interview sections.
- Productive in an outcome means that there is a row in the respective dataset.

### mcs7_hhgrid

<table>
<thead>
<tr>
<th>MCSID</th>
<th>GNUM00 (Parent)</th>
<th>GCNUM00 (Cohort Member)</th>
<th>G_OUT_CM</th>
<th>G_OUT_PAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family 1</td>
<td>1</td>
<td></td>
<td></td>
<td>Productive</td>
</tr>
<tr>
<td>Family 1</td>
<td>2</td>
<td></td>
<td></td>
<td>Unproductive</td>
</tr>
<tr>
<td>Family 1</td>
<td>1</td>
<td></td>
<td></td>
<td>Productive</td>
</tr>
<tr>
<td>Family 1</td>
<td>2</td>
<td></td>
<td></td>
<td>Productive</td>
</tr>
<tr>
<td>Family 2</td>
<td>2</td>
<td></td>
<td></td>
<td>Productive</td>
</tr>
<tr>
<td>Family 2</td>
<td>1</td>
<td></td>
<td></td>
<td>Unproductive</td>
</tr>
<tr>
<td>Family 3</td>
<td>1</td>
<td></td>
<td></td>
<td>Productive</td>
</tr>
</tbody>
</table>

### mcs7_parent_interview

<table>
<thead>
<tr>
<th>MCSID</th>
<th>GNUM00 (parent)</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family 1</td>
<td>1</td>
<td>A</td>
</tr>
<tr>
<td>Family 2</td>
<td>2</td>
<td>C</td>
</tr>
</tbody>
</table>

### mcs7_cm_interview (CM CAPI, CM CASI, CM CAWI)

<table>
<thead>
<tr>
<th>MCSID</th>
<th>GNUM00 (cohort member)</th>
<th>CAPI face-to-face interview variables</th>
<th>CASI (self-completion) variables</th>
<th>CAWI variables (web)</th>
<th>Physical Measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family 1</td>
<td>1</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Family 1</td>
<td>2</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>Family 3</td>
<td>1</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>-1</td>
</tr>
</tbody>
</table>
## The G_OUT_SAMPLE variable

It shows how the Cohort Member was interviewed:
- Face-to-face only (Core)
- Online only (Boost)
- Both (Core & Boost)
- Total in Boost ~250

- Boost cases have data primarily in the mcs7_cm_interview
- Only limited Boost cases have: family information in the household grid, parent interview, cognitive assessment or physical measurements

### mcs7_cm_interview

<table>
<thead>
<tr>
<th>MCSID</th>
<th>GCNUM00 (cohort member)</th>
<th>CAPI (parent) face-to-face interview variables</th>
<th>CASI (self-completion) variables</th>
<th>CAWI variable (web)</th>
<th>Physical Measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family 1</td>
<td>1</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Family 1</td>
<td>2</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>Family 2</td>
<td>1</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Family 3</td>
<td>1</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>-1</td>
</tr>
</tbody>
</table>

### mcs7_hhgrid

<table>
<thead>
<tr>
<th>MCSID</th>
<th>GPNUM00 (parent)</th>
<th>GCNUM00 (cohort member)</th>
<th>Variable</th>
<th>Outcome &amp; Sample variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family 1</td>
<td>1</td>
<td></td>
<td>A</td>
<td>G_OUT_SAMPLE = 1 (Core face-to-face interviewer visit)</td>
</tr>
<tr>
<td>Family 1</td>
<td>2</td>
<td>twins</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>Family 1</td>
<td></td>
<td></td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Family 1</td>
<td></td>
<td></td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Family 2</td>
<td>2</td>
<td></td>
<td>E</td>
<td>G_OUT_SAMPLE = 3 (both: Core face-to-face &amp; online follow-up Boost)</td>
</tr>
<tr>
<td>Family 2</td>
<td></td>
<td></td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>Family 3</td>
<td>1</td>
<td>-1</td>
<td>G_OUT_SAMPLE = 2 (online follow-up Boost)</td>
<td></td>
</tr>
</tbody>
</table>

- The G_OUT_SAMPLE variable
  - Face-to-face only (Core)
  - Online only (Boost)
  - Both (Core & Boost)
  - Total in Boost ~250

- Boost cases have data primarily in the mcs7_cm_interview
- Only limited Boost cases have: family information in the household grid, parent interview, cognitive assessment or physical measurements
Handling Parent/Carer data in MCS7

- Both parents received the same questionnaire in the 7th sweep (Sweeps 1-6 the Main answered more questions than the Partner)
- The ELIG variable exists in the mcs7_hhgrid but it is different to previous sweeps
- Guidance on creating a composite variable that summarises the data (e.g. income) of the parents/carers per household (see Data Handling guide)
MCS Longitudinal Data Dictionary

Question code (CAPI) possible to find a variable between sweeps

Guidance on how to use

Frequently updated with new releases added to the data

In the documentation folders downloaded from the UKDS with the data
The mcs_longitudinal_family_file

- It contains the weights (UKDS SN: 8172)
- The variable DATA_AVAILABILITY shows the cases that have data available under End User Licence or Secure Access
Finding specific variables:
- MCS Longitudinal Data Dictionary
- Questionnaires
- CLOSER platform: discovery.closer.ac.uk/

For any query that is not covered by the user guides: clsfeedback@ucl.ac.uk
Feel free to ask us questions 😊