

The Millennium Cohort Study

CWIPP seminar: 28th January 2014

ACCESSING THE MCS DATA

Getting to know the data folder

- Data
 - Longitudinal family file
 - Derived variables
 - Geographically linked data
 - Household grid
 - Parent Interview – **Main survey dataset**
- Documentation
 - Data dictionaries
 - Conversion from old variable names
 - Technical reports and user guides

The datasets (MCS1)

Dataset	Description
Parent interview	Main survey dataset – survey responses of main and partner respondents and all information about the cohort member
Derived variables	Useful variables constructed by dataset managers but not in main survey dataset (e.g. family structure, ethnicity, socioeconomic status)
Geographically linked data	Mainly Index of Multiple Deprivation (IMD) related variables
Household grid	Sex, DoB and employment status of all household members. Relationship between all household members.
Longitudinal family file	Sampling and weighting variables

Variable names

amfcincb

bhsdoa01

cxkfjwa0


Variable names

amfcincb


[a] [m] [fcin] [c] [b]

[Prefix1][Prefix2][Question name][Suffix1][Suffix2]


MCS wave:
a=1, b=2,
c=3...




Respondent or
instrument:
M=main, p=partner,
h=household, d=derived,
c=cognitive assessment




All questions have
four letter code: see
surveys – repeated
questions mostly
consistent across
waves



Disaggregates
repeated questions:
0=no repeat,
a=iteration 1,
b=iteration 2...



Multi-answer questions:
0=no multi-answer,
a=answer 1, b=answer
2...



Linking the datasets

- Essential for longitudinal analyses
- Linking by family
 - Single link variable: mcsid
- Linking by cohort member
 - Remember: some families have multiple cohort members
 - Link variables: mcsid **and** xhcny0
 - Where x is the survey wave and y is the cohort member of interest
- Can also link by/to
 - Main respondent
 - Partner respondent
 - Older siblings
 - Household grid by individual in main dataset

Additional features and resources

- Feed forward data
 - Data collected at earlier waves and automatically generated at later waves
 - Examples: participant age and gender
- Secure data
 - Available via ONS Secure Data Service
 - Ward level geographic identifiers
 - Educational data from the National Pupil Database
- Key resource websites
 - Centre for Longitudinal Studies website – All documentation
 - National Data Archive – Datasets and documentation
 - Society for Lifecourse and Longitudinal Studies - Researchers