

National Child Development Study

User Guide 2006

Prepared to accompany the re-deposit of data

Brian Dodgeon, Jane Elliott, Jon
Johnson, and Peter Shepherd

First Edition

November 2006

Centre for Longitudinal Studies

Bedford Group for Lifecourse & Statistical Studies

Institute of Education, University of London



E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

Table of Contents

Introduction.....	3
What has changed?	3
Re-organisation of datasets	4
Number of cases.....	4
Sex.....	5
Occupation (CASOC) coding at NCDS5.....	5
Mother and Child Survey (NCDS5)	5
Geography.....	5
Renaming	6
Variable labels	6
Value labels.....	7
Derived variables	7
Dropped variables	7

Introduction

This user guide covers the re-deposit with the UK Data Archive in November 2006 of the datasets corresponding to the National Child Development Study surveys from 1958 to 1999/2000. It does not attempt to reproduce earlier documentation.- this may be consulted at the CLS or UK Data Archive websites (<http://www.cls.ioe.ac.uk/ncds> or <http://www.data-archive.ac.uk/findingData/ncdsTitles.asp>). For a general introduction to the data collections up to NCDS5 (1991) consult *NCDS User Support Group Working Paper 1: The National Child Development Study: An introduction, its origins and the methods of data collection*. P. Shepherd (1997) and for the NCDS6 (1999/200) data collection consult *CLS Cohort Studies Working Paper 1: The design and conduct of the 1999-2000 surveys of the National Child Development Study and 1970 British Birth Cohort Study*. J Bynner, N Butler, E Ferri, P Shepherd and K Smith (2000). both available from the CLS website (<http://www.cls.ioe.ac.uk/ncds>).

The National Child Development Study (NCDS) is a continuing, multi-disciplinary longitudinal study. It began when data were collected about 17,415 babies born in Great Britain (England, Scotland and Wales) in one week in 1958. To date, there have been seven attempts to trace all members of the birth cohort in order to monitor their development in key domains – health, education, social and economic. The follow-ups were undertaken when the cohort members were aged 7, 11, 16, 23, 33, 42 and 46 years and, for the first three of these, the birth cohort was also augmented by including in the target sample immigrants born in the relevant week and identified from school registers during tracing. Over the years, information has been gathered from a number of sources (eg: parents, schools, doctors, medical records, cohort members) and in a variety of ways (eg: interview, self-completion, assessments, medical records). It should be noted that details of public examination entry and performance were obtained from schools prior to the 23 year follow-up. Also, during the age 33 follow-up, a special study was also undertaken of the children of one third of the cohort members, including assessments of their behaviour and cognitive development. In addition to these main data collections, a number of special sub-studies where additional data has been gathered for samples of cohort members selected for their special characteristics or circumstances.

Further information about NCDS is available from the CLS website (<http://www.cls.ioe.ac.uk/ncds>)

What has changed?

In re-depositing the NCDS data we have had to strike a balance between tidying up and relabelling data to make it more accessible for new users while ensuring that long-term users with existing syntax and coding files are not unduly affected. We have taken the opportunity to clear up anomalies and inconsistencies, primarily in the variable labels and value labels.

A summary of the main improvements included within the re-deposited data are listed here and further details of each provided below

- Data have been reorganized to reduce the total number of datasets and to reflect more accurately the way that data was collected
- Provision of a response dataset to indicate which sweeps of data are available for each case
- Mis-assigned serial numbers for ten cases at NCDS 5 have been corrected
- A variable for cohort member's sex has been checked and provided for each dataset
- Occupational coding using CASOC at NCDS5 has been corrected
- Additional cases added for NCDS5 Mother and child survey
- New region variables have been developed for each sweep of the study which are more comparable with other studies
- Variable labels and value labels have been improved to reflect as accurately as possible the questions asked and the response categories in each sweep
- Additional derived variables for NCDS5 have been provided
- A few variables that were redundant have been removed

Re-organisation of datasets

The data was originally deposited in the form of 25 datasets (see column 'Source Dataset(s)' in table below). It has now been simplified to eight, with surveys up to age 16 (NCDS0 to 3) in one dataset, NCDS4 in one, and NCDS5 in three. NCDS6 (1999/2000) has been separated from the combined dataset with BCS70 for ease of use. Also a response dataset has been added to complement the Technical Report on Populations and Samples¹.

Content	Cases	New Dataset	Source Dataset(s)
Childhood data (PMS & NCDS 1-3)	18,558	ncsd0123	rec1.por; rec2.por; yyy3.por; dervars.por
Exams	14,331	exams	exams.por
NCDS4 Survey Data	12,537	ncds4	ncds4.por; dervars4.por
NCDS5 Cohort Member Interview and booklets	11,469	ncds5cmi	cmi.por; cmiopen1.por; cmiopen2.por; cmlopen.por; cmilife.por; n5fringe.por; n5casoc2.por; wdyt.por;
NCDS5 Mother and Child Survey	4,287	ncds5mc	mother.por; mip23.por; mopen.por; child.por; hoenv.por; ychild.por
NCDS5 Partner Survey	7,719	ncds5p	plife.por; plopen.por
NCDS6 Cohort Member Survey	11,419	ncds6	ncdsbcs70c1_2.por
Response at each survey	18,558	response	conout.por n4noncon.por

Number of cases

The number of cases in both the summary response file and the childhood data file is 18558. This represents every cohort member who has been present at any of the surveys (N.B. there were a total of 17415 cases in the original Perinatal Mortality Study (NCDS0) in 1958 and it should be noted that immigrants born in the appropriate week were added subsequent to the birth survey, up to age 16).

¹ PLEWIS, I., CALDERWOOD, L., HAWKES, D. and NATHAN, G. (2004) Changes in the NCDS and BCS70 populations and samples over time. CLS Technical Report. London: CLS, Institute of Education

Four cases were 'swapped over' at NCDS5 and have been restored to the correct individual.
 085008X and 085007V
 092206S and 092207U
 383025K and 383023E
 423023Y and 424023D

Two cases were given invalid serial numbers at NCDS5 and NCDS6.
 X82178S reassigned to X28178L
 X67010Q reassigned to X67070K

Sex

The variable N622 (and its companion variables, n622_4, N622_5, N622_6 and n622_7) should be used for the Cohort Member.

Occupation (CASOC) coding at NCDS5

For a number of years the revised CASOC variables, which provide occupation groupings and socio-economic variables at NCDS5, have only been available via CLS. These variables supercede those included with the original deposit and are now included in the re-deposit.

Mother and Child Survey (NCDS5)

11 mother cases and 8 child cases not previously available have been added..

Geography

As a consequence of the Boyle Report² into disclosure control in longitudinal datasets, we are withdrawing from the publicly-available datasets under the standard End User Licence any geographical information at a finer level than standard region. This more detailed data will still be made available to users under Special Licence (see <http://www.esds.ac.uk/orderingdata/specialLicence.asp>).

Description	Variables	Dataset
LEA code no	n4	rec2.por
1969 Local Authority code number 1969 Local Authority - Medical Examination 1969 Local Authority code number	N1101 N1461 N801	rec1.por
Local Authority-Parental Questionnaire Local Authority - Medical Examination Local Authority – Audiogram Local Authority - Schools Questionnaire Local Authority & voluntary schools Local Authority - School Questionnaire Local Authority – Individual Local Authority - Test Booklet	N2349 N1879 N2048 N2092 N2102 N2268 N2702 N2918	rec1.por
NCDS4 1971 Census Small Area Statistics	All	caci71a.por
NCDS4 1981 Census Small Area Statistics	All	county81.por
NCDS4 1981 Census Small Area Statistics	All	dist81.por
NCDS4 1981 Census Small Area Statistics	All	pc81ward.por
NCDS4 1981 Census Small Area Statistics	All	ward81.por

² Boyle, P.J., Cullis, A, Feng, Z., Flowerdew, R. and Gayle, V. 2004 Adding Geographical Variables and Identifiers to Longitudinal Datasets, Report to the National Longitudinal Strategy Committee and ESRC

In line with the Boyle Report recommendations we have also re-calculated region to be on a consistent basis (SSR and GOR) and one that is comparable with other studies (for further information see **NCDS revised region variables.pdf**, included with this re-deposit) that accompanies this release. The following variables have thus been withdrawn as they have been used to derive the revised region variables.

Description	Variable	Dataset
0 Standard region	n624	rec1.por
1 Standard region	n623	rec1.por
2S Standard region	n802	rec1.por
2S Standard region (BSAG)	n962	rec1.por
2P Standard region	n1102	rec1.por
2M Standard region	n1462	rec1.por
2T Standard region	n902	rec1.por
3M Standard region	n1880	rec1.por
3M Standard region (Audiogram)	n2049	rec2.por
3S Standard region	n2093	rec1.por
3P Standard region	n2350	rec1.por
3I Standard Region	n2703	rec1.por
3T Standard Region	n2919	rec1.por
Region (Exams) (same as n2093)	e232	exams.por

Renaming

A few derived variables at NCDS4 have been renamed to ensure they remain unique within this and future deposits. The variables are:

tenure is now dtenure.
hhsz is now dhhsz.
empstat is now dempstat.

Variable labels

Whilst we have attempted to encapsulate the question asked within the variable label, users are **strongly advised** to check the paper questionnaire and CAPI (computer assisted personal interviewing) documentation (available on the CLS website, and with the relevant datasets) to understand exactly what was asked and the context of the question within the survey.

To facilitate this, an Idealist Data Dictionary is provided and additionally the variable label contains a summary of the location of the question.

For example, variable n204 is labelled "1P Use of bathroom". This means that it was collected at survey 1, Parental Interview. Variable n500527 is labelled "CMI:3,A5a) CURRENT/MOST RECENT JOB: Employee or self-employed". This means that it was collected in the Cohort Member Interview, page 3, question A5a.

A summary of the prefixes used in variable labels is shown below.

Prefix	Survey	Instrument / Questionnaire
0	PMS (1958)	All
1P	NCDS1 (1965)	Parental Questionnaire
1PD	NCDS1 (1965)	Parental Questionnaire (Derived Variable)
1M	NCDS1 (1965)	Medical Questionnaire
1S	NCDS1 (1965)	Educational Assessment
1D	NCDS1 (1965)	Derived Variable
2P	NCDS2 (1969)	Parental Questionnaire

2M	NCDS2 (1969)	Medical Questionnaire
2S	NCDS2 (1969)	Educational Assessment
2S BSAG	NCDS2 (1969)	Educational Assessment (Bristol Social Adjustment Guide)
2T	NCDS2 (1969)	Test Booklets
3P	NCDS3 (1974)	Parental Interview
3S	NCDS3 (1974)	Educational Questionnaire
3M	NCDS3 (1974)	Medical Examination
3I	NCDS3 (1974)	Individual Questionnaire
3T	NCDS3 (1974)	Test Booklet
4I	NCDS4 (1981)	Interview
4D	NCDS4 (1981)	Derived Variables
E	Exams (1978)	Examination Survey
CMI	NCDS5 (1991)	Cohort Member Interview
CASOC	NCDS5 (1991)	Cohort Member Interview (Derived CASOC)
CMIOOPEN1	NCDS5 (1991)	Cohort Member Interview (Coded Open Answers)
CMIOOPEN2	NCDS5 (1991)	Cohort Member Interview (Coded Open Answers)
CML	NCDS5 (1991)	Cohort Member "Your Life"
CMLOPEN	NCDS5 (1991)	Cohort Member "Your Life" (Coded Open Answers)
WDYT	NCDS5 (1991)	Cohort Member "What Do You Think"
DV	NCDS5 (1991)	Derived Variables
SC	NCDS6 (1999/2000)	Self Completion

Value labels

As part of the re-deposit, extensive labelling of values has been undertaken. Where a value existed but was not labelled this has now been rectified and if appropriate the value has been set to missing. Where a missing value was set to system missing (in SPSS parlance), we have not attempted to assign an item-response missing value, so as not to interfere with any existing code.

Derived variables

Derived variables at NCDS5 were not deposited with the data collected in 1991 (although the SPSS syntax was). The relevant syntax file has now been run and the derived variables added to the NCDS5 Cohort Member Interview dataset. Derived variables for height and weight at age 7, 11, 16 and 23 are also now assigned to the relevant survey dataset.

Dropped variables

Some empty, redundant variables have been removed.

Further information

Further information on NCDS is available from the CLS website (<http://www.cls.ioe.ac.uk/ncds>). CLS can also be contacted at the following email address: clsfeedback@ioe.ac.uk

Comments

We would value your comments on this document, or on any aspect of NCDS. Please email: clsfeedback@ioe.ac.uk.