British Cohort Studies Teaching Dataset

A Guide to the Datasets and Resources

June 2008 (2nd Edition)

Samantha Parsons, Jane Elliott

Centre for Longitudinal Studies
Institute of Education, University of London

PREFACE

This document has been prepared to accompany the deposit, with the UK Data Archive at the University of Essex, a set of teaching datasets and associated resources for lecturers in higher education.

Deposit: Elements

Title	Format
British Cohort Studies Teaching Dataset : A Guide to the Datasets and Resources	PDF
KeyVariables	PDF
OpeningData_SPSS	PDF
MalaiseSummary	PDF
MalaiseVariables	PDF
MalaiseExercises	PDF
MalaiseQuestions	PDF
MalaiseReferences	PDF
MotherMalaiseQuestions	PDF
MotherMalaiseVariables	PDF
RutterBehaviourSummary	PDF
RutterBehaviourVariables	PDF
RutterBehaviourExercises	PDF
RutterBehaviourQuestions	PDF
RutterBehaviourReferences	PDF
CognitiveAssessmentSummary	PDF
CognitiveAssessmentVariables	PDF
CognitiveAssessmentExercises	PDF
CognitiveAssessmentQuestions	PDF
CognitiveAssessmentReferences	PDF
bcs70_keyvariables	SPSS
ncds_keyvariables	SPSS
b034malaise	SPSS
n042malaise	SPSS
b016MotherMalaise	SPSS
b016rutter	SPSS
n016rutter	SPSS
b516reading.sav	SPSS
n716maths.sav	SPSS
n716reading.sav	SPSS

CONTENTS

INTRODUCTION

In January 2006 CLS started a two-year project to create a set of teaching datasets and associated resources for lecturers in higher education. They will be based on the 1958 National Child Development Study, the 1970 British Cohort Study and the Millennium Cohort Study. The project is funded under the ESRC Researcher Development Initiative (RDI).

The aim was to provide a series of datasets that contain fully cleaned and documented subsets of the variables from the National Child Development Study, the British Cohort Study (1970) and the Millennium Cohort Study.

The datasets will facilitate teaching of quantitative data analysis, in particular teaching longitudinal analysis to postgraduate and undergraduate students. The datasets will also be of value to the wider research community as they will include many of the most commonly used variables from the studies, such as data on qualifications, social class, marital status, income, number and ages of children etc.

NCDS

NCDS started life as the Perinatal Mortality Survey, which was designed to examine the social and obstetric factors associated with stillbirth and infant mortality. In the first survey, data were collected about the births and families of 17,638 babies born in Great Britain during one week in March 1958. Since then, there have been seven surveys gathering information from respondents living in England, Scotland and Wales, in order to monitor their health, education, social and economic circumstances. These surveys were carried out in 1965 (age seven), 1969 (age eleven), 1974 (age sixteen), 1981 (age 23), 1991 (age 33), 1999/2000 (age 42) and 2004/2005 (age 46). As part of the 1991 survey, information was additionally collected on the children of one in three cohort members; this included assessments of the behaviour and cognitive development of around 5,000 co-resident children. There have also been surveys of sub-samples of the cohort, the most recent occurring in 1995 (age 37), when a 10% representative sub-sample was assessed for difficulties with basic skills. Finally, during 2002-2004, 9,340 NCDS cohort members participated in a bio-medical survey, carried out by qualified nurses; the bio-medical survey did not cover in detail any of the topics included in the 2004/2005 survey. The 2004/2005 follow-up aimed to extend the data collection of the previous surveys. This latest wave of the NCDS was conducted for the first time as a telephone interview (CATI)[1] when the cohort members were 46 years of age.

Data for NCDS have so far been collected from a number of different sources; the midwife present at birth, the cohort members' parents, the head and class teachers, school health service personnel, the cohort members themselves, their spouses, cohabitees and children, and the 1971 and 1981 censuses. Data has also been collected using a variety of methods; paper and electronic questionnaires and self-completion questionnaires, clinical records, medical examinations, physical measurements, ability tests, and educational assessments.[2]

BCS70

BCS70 began when data were collected about the births and families of over 17,000 babies born in England, Scotland, Wales and Northern Ireland in during one week in 1970. Since the birth survey there have been six other major data collection exercises in order to monitor their health, education, social and economic circumstances. These were carried out in 1975 (age 5), 1980 (age 10), 1986 (age 16), 1996 (age 26), 2000 (age 30) and 2004 (age 34). Samples selected from the birth cohort have also been studied at various ages: for example at age 21, a 10 per cent representative sample was assessed for basic skills difficulties.

From the original focus on the circumstances and outcomes of birth, the study has broadened in scope to map all aspects of health, education and social development of their subjects as they passed

---

[1] CATI stands for Computer Assisted Telephone Interview

[2] For further information on NCDS sweeps see: Power, C. and Elliott, J. (2005) Cohort Profile: 1958 British birth cohort (National Child Development Study). *International Journal of Epidemiology*, 2005, Information can also be found on the CLS website http://www.cls.ioe.ac.uk.

through childhood and adolescence. In latter sweeps, the information collected has covered their transitions into adult life, including leaving full-time education, entering the labour market, setting up independent homes, forming partnerships and becoming parents.

The latest rounds of data collection for BCS70 took place in 2004/2005 when BCS70 cohort members were 34/35. The main aim of this most recent survey was to explore the factors central to the formation and maintenance of adult identity in each of the following domains:

- Lifelong learning
- Relationships, parenting and housing
- Employment and income
- Health and health behaviour
- Citizenship and values

USING THE DATASETS AND RESOURCES

There are three sets of resources for teaching students quantitative methods using longitudinal data from the British Birth Cohorts. One is introductory and based on 'key' socio-economic information collected in the two cohort studies from birth to age 34 (BCS70) or age 42 (NCDS). The other two resource packs are more comprehensive and based on the *Rutter Behaviour Scale* and the *Malaise Inventory* (Rutter 1970). Each pack contains two SPSS datasets (one based on BCS70 and one based on NCDS) and five/six accompanying documents. All documents are stored in .pdf format at the data archive. Editable Word documents are available from the CLS website. See
http://www.cls.ioe.ac.uk/text.asp?section=0001000100050017

Key Variables Resource pack

The Key Variables resource pack includes two SPSS data files, b034KeyVariables.sav and n042KeyVariables.sav and an accompanying descriptive document KeyVariables.pdf. Each data file contains socio-economic information from childhood and adult life, including parental education, family social class in childhood, and cohort members' own education, employment and occupation experiences. b034KeyVariables.sav contains a selection of key variables from all sweeps of data collection in BCS70 that have taken place between birth to age 34. n042KeyVariables.sav contains a selection of key variables from all sweeps of data collection in NCDS that have taken place between birth to age 42.

Rutter Behaviour Scale Resource pack

The *Rutter Behaviour Scale* is an established scale to measure signs of behaviour disorders in children and teenagers. The Rutter Behaviour Scale resource pack contains two SPSS data files, *b016rutter.sav* and *n016rutter.sav* and five accompanying pdf documents.

Data

Each data file contains overall Rutter Behaviour scores (an indicator of challenging behaviour in childhood) for cohort members at various ages in childhood up to age 16. Socio-economic information from childhood and adult life are also included along with the individual Rutter behaviour questions that were used to make the overall score at each age.

Documentation

RutterBehaviourSummary: This document summarises the set of resources based on the Rutter Behaviour Scale.

*RutterBehaviourQuestion*s: this document specifies the questions asked at each age, the different answer categories used to answer the questions at each age and in each cohort and examples of how the questions were presented to the cohort members in the questionnaires.

RutterBehaviourVariables: this document lists all the variables in the b016rutter.sav and the n016rutter.sav SPSS data files. The original variable name that was assigned to a question and found in the full dataset and documentation of a particular survey is also included.

RutterBehaviourExercises: this document provides a few step-by-step exercises for students to work through and additional exercises are suggested.

RutterBehaviourReferences: this document contains references and further reading on the Rutter Behaviour Scales, including research based on the Rutter Behaviour Scales using BCS70 and NCDS data.

Malaise Inventory Resource Pack

The Malaise Inventory is an established scale to measure signs of psychological distress or depression in teenagers and adults. The Malaise Inventory resource pack contains two SPSS data files, *b034malaise.sav* and *n042malaise.sav* and six accompanying pdf documents.

Data
The b034malaise.sav and n042malaise.sav data files each contain overall Malaise scores (an indicator of depression) for cohort members at various ages in adulthood. Socio-economic information from childhood and adult life are also included along with the individual malaise questions that were used to make the overall score at each age.

Documentation
MalaiseSummary: This document summarises the set of resources based on the Malaise Inventory.

MalaiseQuestions: this document specifies the questions asked at each age, the different answer categories used to answer the questions at each age and in each cohort and examples of how the questions were presented to the cohort members in the questionnaires.

MalaiseVariables: this document lists all the variables in the b034malaise.sav and n042malaise.sav SPSS data files. The original variable name that was assigned to a question and found in the full dataset and documentation of a particular survey is also included.

MalaiseExercises: this document provides a few step-by-step exercises for students to work through and additional exercises are suggested.

MalaiseReferences: this document contains references and further reading on the Malaise Inventory, including research based on the Malaise Inventory using BCS70 and NCDS data.

OpeningData_SPSS: this document shows students how to open a data file in SPSS, how to save a subset of the variables from the data file into a new data file and how to merge two data files. The examples use the b034malaise.sav and the b016mothermalaise.sav data files.

Mother Malaise Inventory Resource Pack

The Malaise Inventory is an established scale to measure signs of psychological distress or depression in teenagers and adults. The Mother Malaise Inventory resource pack contains one SPSS data files, *b016mothermalaise.sav* and two accompanying pdf documents.

Data
The b016mothermalaise.sav data file contains overall Malaise scores for BCS70 cohort members mothers at various ages in childhood.

Documentation

MotherMalaiseQuestions: this document specifies the questions asked at each sweep of data collection, the different answer categories used to answer the questions at each age and in each cohort and examples of how the questions were presented to the cohort members' mothers in the questionnaires.

MotherMalaiseVariables: this document lists all the variables in the b016mothermalaise.sav SPSS data file. The original variable name that was assigned to a question and found in the full dataset and documentation of a particular survey is also included.

Cognitive Development through Childhood Resource Pack

This set of resources is based on the assessment of cognitive development in children – namely reading and mathematics – for teaching students quantitative methods. The assessments completed by the children differ by age and cohort. Additional supplementary resources are also outlined. The Resource Pack contains three SPSS data files and five accompanying pdf documents.

Data

The b516reading.sav data file based on information provided by the 1970 British Cohort Study. The n716reading.sav and n716maths.sav data files are based on the 1958 National Child Development Study Each data file contains overall assessment scores at three different age points in childhood. For BCS70 this was at age 5, 10 and 16, for NCDS this was age 7, 11 and 16. The exact age of each child at the time they sat the assessment is also included in the dataset.

Documentation

CognitiveAssessmentQuestions: this document includes: a short description of the reading and mathematics assessments sat at each age, and examples of how the questions were presented to the cohort members in the questionnaires.

CognitiveAssessmentVariables: this document lists all the variables in the b516reading.sav, the n716reading.sav and the n716maths.sav SPSS data files. Original variable names from the full dataset and documentation of a particular survey are also included.).