National Child Development Study

Activity Histories (1974-2013)

Maggie Hancock

A guide to the dataset (Second Edition)

March 2016

The Centre for Longitudinal Studies (CLS) is an ESRC Resource Centre based at the Institution of Education. It provides support and facilities for those using the three internationally-renowned birth cohort studies: the National Child Development Study (1958), the 1970 British Cohort Study and the Millennium Cohort Study (2000). CLS conducts research using the birth cohort study data, with a special interest in family life and parenting, family economics, youth life course transitions and basic skills. The views expressed in this work are those of the author(s) (amend as necessary) and do not necessarily reflect the views of the Economic and Social Research Council. All errors and omissions remain those of the author(s).

This document is available in alternative formats.
Please contact the Centre for Longitudinal Studies.
tel: +44 (0)20 7612 6875
email: clsfeedback@ioe.ac.uk

Table of Contents

Preface

This document has been prepared to accompany the deposit, with the UK Data Archive at the University of Essex, of the derived dataset of activity histories from the 1981[1],1991, 2000, 2004, 2008/9 and 2013 follow-ups of the 1958 National Child Cohort Study (NCDS), a continuing, multidisciplinary, national, longitudinal study.

The elements of the deposit, to which reference will be made throughout this document, are identified below. Users are advised that they will need to consult all elements of the documentation to gain a full understanding of the data.

The Centre for Longitudinal Studies (CLS) request that any work which is based wholly or in part on analysis of National Child Development Study data includes the following acknowledgement:

"The analyses in this work are based wholly or in part on analysis of data from the 1958 National Child Development Study (NCDS). The data was deposited at the UK Data Archive by the Centre for Longitudinal Studies at the UCL Institute of Education, University of London. NCDS is funded by the Economic and Social Research Council (ESRC)."

CLS also request that the term 'NCDS' be included when compiling keywords to accompany any publications based on analysis of the study's data.

Acknowledgements

We would like to thank Jane Elliott, Jon Johnson and Erzsebet Bukodi for their contributions to the construction of the National Child Development Study Activity Histories.

We would also like to thank all the cohort members and their children, who generously gave their time to participate in this project and without whom this dataset would not have been possible.

---

[1] if cohort member not interviewed in 1991

Previous NCDS Activity Histories Dataset

An employment histories dataset was previously created (Ward, 2007). This work was undertaken as part of the Gender Network Project (http://www.genet.ac.uk/Publications/index.html. The current work on NCDS activity histories builds on this previous activity history and incorporates the cleaning included therein. This previous employment history included data up to sweep 7 (2004) only, did not deal with any non-work activities and did not identify duplicate activities (i.e. where an activity was reported again in a later sweep).

NCDS Activity Histories (1974-2013)

Data on work and non-work activities lasting one month or more have been collected in all NCDS sweeps from sweep4 (age 23) onwards.   The purpose of the NCDS activity histories dataset is to merge all data on work and non-work activities in successive sweeps into one longitudinal dataset.

The focus of the questions asked at each sweep vary from: work activities engaged in since leaving school (sweep 4 aged 23); work and non-work activities engaged in since leaving school (sweep 5 aged 33); work and non-work activities engaged in since the last sweep (sweep6 aged 42); work and non-work activities engaged in since the last sweep or aged 16 (sweep7 aged 46) work and non-work activities engaged in since 2000, or 2004 if included in sweep7 (sweep8 aged 50) and work and non-work activities engaged in since 2004 or 2008 if included in sweep8 (sweep9 aged 55).   Therefore the activity histories will start from the time that the cohort member left school and continue until the interview date of the latest data sweep that each cohort member participated in.

The lengths of the activity histories vary depending on the latest sweep that a cohort member was present at. The minimum activity history length recorded is 1 month and the maximum is 480 months (40 years).   Gaps in the activity histories occur where a cohort member has not been present at all sweeps and/or where full activity data were not reported.

The NCDS activity histories (1974-2008) were first deposited at the UK Data Archive in May 2011. This document accompanies the deposit of the second version of the NCDS activity histories (1974-2013) and is based on the documentation that accompanied the first deposit.

Amendments to the NCDS activity histories since version one

The first version of the NCDS activity histories, deposited at the UK Data Archive in May2011 included activities reported by cohort members up to and including the 2008 data sweep. This second version of the activity histories data includes data on activities reported up to and including the 2013 data sweep

 There are also some edits to the pre 2013 sweep activity history data as follows:

Where the cohort member reported at 2008 sweep doing the same activity as the previous data sweep that they participated in (2000 or 2004), the previous version of the NCDS activity histories updated the 2000/2004 current activity to indicate also current activity at sweep8. However further examination of the data revealed that cohort members may report doing the same activity when in fact their employment status or the type of organisation that they worked for had changed between sweeps. Where this is the case, version 2 of the NCDS activity histories includes a new activity with a start date of the 2008 data sweep interview date (N=1475 for cases with lastsweep=sweep7 and N=85 for cases with

lastsweep=sweep6). The end date of the previous sweeps current activity is also set as the 2008 data sweep interview date.

The first version of the NCDS activity histories excluded sweep7 and sweep8 activities with unknown start dates, and imputed end dates (using the start dates of following activities) where possible In version 2 of the NCDS activity histories, any unknown start or end dates for activities current at a sweep are computed to be that sweep's interview date. For activities not still current at a sweep; unknown start dates are set to that activities known end date if available and vice versa.

After version 1 of the NCDS activity histories had been deposited, an issue was identified with sweep5 out of work slots where the activity was coded as F/T education when it should have been coded as 'Other'.   This has now been corrected. The number of cohort members whose data was affected was 808 (4.3%). The number of activity slots affected was 1081 (0.9%). The number of activity months affected was 11,117 (0.2%).

Upon inspection of the data for cohort member with high values for the OVERLAP variable, it was noticed that some cohort members were reporting sweep' in work' activities again in the sweep5 'not in work' activities with activity as 'Other'. Where this is the case the duplicate activities have been dropped (see appendix 2). There were also cases where the sweep5 'out of work' activities were reported with dates which overlapped the dates of the cohort members other activities, these dates were checked against the data for sweep4 and for the sweep5 cohort member interview and corrected if possible (see appendix2).

Details of the data used in the activity histories

Sweep	Activity data collected at sweep and used in activity histories
Sweep 4 (1981, aged 23)[2] N.B. Sweep4 activity data is included only if CM not interviewed in sweep5 (2711 cases)	For up to 4 work activities (first 3 and latest 1) since leaving school: • Start and end dates of activity • Whether f/t or p/t • Whether employed/temping or self employed • Socio-economic codes (first and current work activities)
Sweep 5 (1991, aged 33)	Data on current and previous activities (work and non-work) since leaving school as reported in the 'Your life since 1974' self completion questionnaire • Start and end dates of activity • Whether f/t or p/t (work activities) • Socio-economic codes (work activities)
Sweep 537 (1995, aged 37) N.B. Sweep 537 activity data is included only if CM not interviewed in sweep 6 or sweep 7 (109 cases)	Data on work activities since 1991 (when aged 33). • Start and end dates of activity • Whether f/t or p/t (work activities) • Reason left (work activities) • Socio-economic codes (for current/most recent job only)
Sweep 6 (2000, aged 42):	Data on current and previous activities (work and non-work) since CM's latest previous sweep. • Start and end dates of activity • Whether f/t or p/t (work activities) • Reason left (work activities) • Socio-economic codes (work activities)
Sweep 7 (2004, aged 46):	Data on current and previous activities (work and non-work) since CM's latest previous sweep or aged 16: • Start and end dates of activity • Whether f/t or p/t (work activities) • Reason left (work activities) • Socio-economic codes (work activities)
Sweep8 (2008, aged 50):	Data on current and previous activities (work and non-work) since age 42 or age 46 if included at sweep7 • Start and end dates of activity • Whether f/t or p/t (work activities) • Reason left (work activities) • Socio-economic codes (work activities)
Sweep9 (2013, age 55):	Data on current and previous activities (work and non-work) since age 46 or age 50 if included at sweeps 7 or 8 or January 2004 if not • Start and end dates of activity • Whether f/t or p/t (work activities) • Reason left (work activities) • Socio-economic codes (work activities)

[2] In the NCDS4 derived variables dataset there are variables describing the economic activity for each month from age 16 to age 23. These variables are derived from the same employment, looking for work and out of the labour force variables used in the activity history dataset construction and do not offer any added information. They are not used in the activity history dataset construction.

NCDS Activity Histories Dataset

The NCDS activity histories dataset, ncds_activity_histories.sav, contains 1 record for each work or non work activity. In the dataset there are 121,456 records for 18,558 cohort members. 956 cohort members were known to have died before the age of 16 (one dummy record with variable ACTNUM=-1) and of the rest, 2862 cohort members, including those not present at sweeps 4-9 or have no activities reported (one dummy record with variable ACTNUM=0). All activity variables are set to -1 (n/a) on these dummy records. 14,740 cohort members have at least 1 activity reported (variable ACTNUM=1). There is therefore an average of 8 activities for cohort members with at least one activity.

A sample of 4 cases with a subset of the variables is shown below:

NCDSID	.....	NUMACTS	ACTNUM	LSTSWCMC	JSTCMC	JENDCMC	JDUR	JACTIV	.....
N01234B	.....	0	-1	-1	-1	-1	-1	-1	.....
....									
N12345N	.....	6	1	1311	1036	1066	30	1	.....
N12345N	.....	6	2	1311	1066	1131	65	1	.....
N12345N	.....	6	3	1311	1131	1182	51	1	.....
N12345N	.....	6	4	1311	1182	1245	63	18	.....
N12345N	.....	6	5	1311	1245	1266	21	12	.....
N12345N	.....	6	6	1311	1266	-6	45	18	.....
....									
N23456P	.....	0	0	-1	-1	-1	-1	-1	.....
....									
N34567S	.....	4	1	1309	1036	1039	3	12	.....
N34567S	.....	4	2	1309	1039	1063	24	14	.....
N34567S	.....	4	3	1309	1063	1126	63	1	.....
N34567S	.....	4	4	1309	1126	-6	183	1	.....

Case N01234B is known to have died before age 16. Case N12345N has 6 activities reported. Case N23456P was not included in sweeps 4-9 or did not have any activities reported and therefore has 1 dummy activity included in the dataset. Case N34567S has 4 activities reported.

The description of the variables included in the above sample is as follows:

Variable name	Variable description
NUMACTS	Total number of activities recorded for the cohort member (constant across all activities for a cohort member
ACTNUM	Activity sequence number (within cohort member)
LSTSWCMC	Century month[3] of cohort member's latest interview
JSTCMC	Century month[3] of the start date of the activity
JENDCMC	Century month[3] of the end date of the activity
JDUR	Duration of the activity in months
JACTIV	Activity code for the activity (see table below for activity codes)

Data variables included in the activity histories

A summary of the data variables included in the activity histories is as follows:

Participation in each sweep and if so whether any activity data at that sweep
CM's sex
Activity number and total number of activities for CM
Century months[3] of latest interview and activity start and end dates
Activity duration in months (truncated at latest sweep)
Activity type
Activity full-time or part-time (work activities only)
Activity start and end dates (month and year)
Data on length of known activity history / overlaps and gaps (months)
Case and activity errors
Reason left job
Activity employment status and Activity socio-economic codes
Latest sweep activity mentioned at and date of CMS interview at each sweep
Sweep and slot number that activity was reported in

A detailed list of the variables included in the activity histories is found in appendix 1.

---

[3] Century month is calculated as the number of month since 1900

Types of activity included in the activity histories

Variable JACTIV holds the activity type (work and non-work), the activity categories included are:

1	F/t paid employee 30+ hrs
2	P/t paid employee lt 30 hrs
3	F/t self-employed
4	P/t self-employed
5	Employed work but not known if ft/pt
6	Self-emp work but not know if ft/pt
7	Employed, but unpaid
8	F/t work not know if emp or s/e
9	P/t work not know if emp or s/e
10	Work but not known if ft/pt or emp/se
11	Unemployed seeking work
12	F/t education
13	Part-time education
14	Government training scheme
15	Temporarily sick/disabled
16	Permanently sick/disabled
17	Sick/ disabled not known if perm/temp
18	Looking after home/family
19	Wholly retired
20	Voluntary work
21	Maternity leave
22	Travelling/ extended holiday
97	Other

Socioeconomic variables

For work activities, socioeconomic status is coded at each sweep (e.g. NSSEC, RGSC). No coding frame is used across all the sweeps. For completeness all socioeconomic variables coded are included in the NCDS activity history dataset. Where a socioeconomic variable has not been coded at a sweep, the value is set to -8 (don't know/not enough information) for all work activities.

Details of these socioeconomic variables and the sweeps in which they are included are as below:

variable	variable label	Whether included for work activities in sweep (current implies current job only):						
		4	5	537	6	7	8	9
JYLEFT	Reason left job (to match swp4)	1st/current		yes	yes	yes	yes	yes
JYLEFTX	Reason left job (detailed swps6+)			yes	yes	yes	yes	yes
JYLEFTO	Reason left job (detailed swps6+) other specified				yes	yes	yes	yes
JEMPST	Activity employment status	1st/current	yes		yes	yes	yes	yes
J91SEG	Activity SEG '91		yes	most recent	yes	yes	yes	yes
J90SOC	Activity soc90		yes	most recent	yes	yes	yes	
J91SC	Activity social class '91		yes	most recent	yes	yes	yes	yes
J2KSOC	Activity soc2000					yes	yes	yes
J10SOC	Activity soc2010							yes
JNSSEC	Activity NS-SEC (long version)					yes	yes	yes
J8NSSEC	Activity NS-SEC8 (analytic version)					yes	yes	yes
J07SIC	Activity S.I.C. 2003 (swp9)							current
J03SIC	Activity S.I.C. 2003 (swp8)						yes	
J92SIC	Activity S.I.C. 1992 (swp7)					current		
JSIC	Activity S.I.C. (swp6)				current			
JTORG	Activity Type of organisation CM works for (swps6&7)			most recent	current	current		
JCAMGP	Activity Cambridge Sub-Group & Title Ident. (swp5)		yes					
JEMPCC	Activity Employment Status - check code (swp5)		yes					
JCAMSM	Activity Cambridge Scale Score - Males (swp5)		yes					
JCAMSF	Activity Cambridge Scale Score - Females (swp5)		yes					
J88ISC	Activity ISCO '88-Int Stan Class of Occupns (swp5)		yes					
JKOS	Activity KOS - Key Occupns for Stats (swp5)		yes	most recent				
JCODOT	Activity CODOT - 3 digit classification (swp5)		yes	most recent				

variable	variable label	Whether included for work activities in sweep (current implies current job only):						
		4	5	537	6	7	8	9
JOP80C	Activity OPCS 1980 - Operational Code (swp5)		yes					
JOP80G	Activity OPCS 1980 - Occupation Unit Group (swp5)		yes					
J81SEG	Activity Socio-Economic Group 1981 (swp5)		yes					
J81RGS	Activity Social Class RG"s 1981 (swp5)		yes					
JOP70G	Activity OPCS 1970 - Occupation Unit Group (swp5)		yes					
J71SEG	Activity Socio-Economic Group 1971 (swp5)		yes					
J71RGS	Activity Social Class RG"s 1971 (swp5)		yes					
JRG66G	Activity RGs 1966 - Occupation Unit Group (swp5)		yes					
J66SEG	Activity Socio-Economic Group 1966 (swp5)		yes					
J66RGS	Activity Social Class RG"s 1966 (swp5)		yes					
JRG61G	Activity RGs 1961 - Occupation Unit Group (swp5)		yes					
J61SEG	Activity Socio-Economic Group 1961 (swp5)		yes					
J61RGS	Activity Social Class RG"s 1961 (swp5)		yes					
J90GSC	Activity Goldthorpe Social Class 1990 (swp5)		yes					
J80GSC	Activity Goldthorpe Social Class 1980 (swp5)		yes					
J70GSC	Activity Goldthorpe Social Class 1970 (swp5)		yes					
JHGCG	Activity Hope-Goldthorpe Collapsed Group (swp5)		yes					
JHGSS	Activity Hope-Goldthorpe Scale Score (swp5)		yes					
J70SEG	Activity Socio-economic Group 1970 (swp4)	1st/ current						
J70SC	Activity Social Class 1970 (swp4)	1st/ current						
JINDUS	Activity industry (swp4)	1st /current						
J80CO	Activity co80	1st/ current						
J80SC	Activity social class '80	1st/ current						
J80SEG	Activity SEG '80	1st /current						

Construction of the NCDS Activity Histories

The NCDS activity histories dataset is constructed using the SIR programming language. Activity data is accessed from sweeps 4, 5, 6, 7 and 8 and stored in arrays. These arrays are then examined for data issues that require editing; duplication (where activities are reported again in a later sweep) activity overlap and missing end dates (see below).

Cleaning of the NCDS activity data

Activity data from sweeps 4 and 5 has been identified as requiring cleaning in data note 2007/1 (Employment Histories in NCDS4, 5 and 6 – notes on data consolidation and cleaning). This cleaning constituted small amounts of editing to some activity data, mainly adjustments to start and end dates and whether the activity is full-time or part-time. The NCDS activity history construction program includes this cleaning

Other data issues that came to light during the construction of the activity histories has also been incorporated into the NCDS activity history construction program.

Editing of the NCDS activity data due to issues with the data

Where an activity in a sweep is reported to have dates which overlap with activity dates in a previous sweep:

- If the activities in both sweeps are the same, it is assumed that they are duplicated and the start date of the activity in the later sweep is edited to match that of the previous sweep
- Else if the activities are different and if an end date is reported for the activity in the previous sweep the start date of the activity reported in the later sweep is set to this end date.
- Else if the activities are different and the activity in the previous sweep has no end date reported, the end of the activity reported in the previous sweep and the start of the activity reported in the later sweep are set to be the interview date of the later sweep (the 'yleft' variables are set to -8 don't know for the activity in the previous sweep).

Activities current at the time of a sweep interview

Activities current at the time of a sweep interview should be included in the next sweep as either the current activity or a previous activity since the last sweep. In these cases the activities reported in the different sweeps are treated as duplicates see below.

Duplicate activities

All activities in a sweep are compared with all activities in previous sweeps to identify duplicates. Activities are considered to be duplicates if:

They are reported in different sweeps, appear to be the same activity[4] and:

- activity start and end dates match or both end dates missing or end date missing in the earlier sweep (duplicate type 1)   or
- activity start dates are within 6 months of each other and dates overlap, both end dates are missing or end date missing in the earlier sweep (duplicate type2)   or
- activities start within 14 months[5] of each other and no end dates reported or no end date for the activity with the earlier start date (duplicate type 3)

In the case of duplicates, dates reported in the earlier sweep are used as these are assumed to be the more reliable. Non-date data reported in the earlier sweep are used where missing in the later version. Data is only taken from the earlier duplicated activities if this is consistent with the later data.   Duplicated activities are not included in the NCDS activity history dataset.

Activity start and end dates and durations

Activities that start and end before the age of 16 are omitted from the activity history dataset. For activities that start before the age of 16 and end after the age of 16, the start date is edited to be March 1974 when the cohort members were 16.

Where an activity is reported to overlap with a subsequent activity, and both are full time activities or either activity is one that cannot be combined with any other full time or part time activity, the activity duration is truncated to end before the subsequent activity starts.   In order that such truncation does not result in a substantial gap in the activity history; it only occurs if the activity end date is less than 6 months after the end date of the subsequent activity or another subsequent activity starts before the end date of the activity. Activity overlaps are allowed where a valid overlap of activities occurs, for example full time education and part time work.

Where an activity has no end date reported, the end date is computed as the start date of the next chronological activity or, if there is no following activity, the end date is computed as the interview date of the latest sweep that the activity was reported at.

Date century months

In order to facilitate date computations, century months, the number of months since 1900, are calculated for all dates. Where the date year is missing, the century month variable is set to -8 (don't know). Where the date month is missing, a value of 6 (June) is used in the century month calculations.

---

[4] Activities are considered to be the same if the same non-work activity or the same work activity with matching (or very similar) soc codes
[5] Examination of the data showed that a common discrepancy in start dates of duplicate activities was differing by 1 year and up to 2 months

Missing data

Variables are set to missing value -1 (not applicable) where no data is expected as in the case of dummy activity records (which are added where a cohort member was reported to have died by age 16 or has had no activity data reported) or where a variable is not applicable for a particular activity, for example socio-economic variables for a non-work activity.

Variables are set to -8 (don't know/ insufficient information) where data is expected but has not been collected.

End of activity variables are set to -6 where an activity is still current at the cohort member's most recent sweep.

Zero duration activities

Activities have duration of zero where:

- An activity is reported to be still current at the cohort members' most recent interview and start in the same month as the month of the cohort members most recent interview
- The cohort member has reported an activity to start and end in the same month
- The start date reported for a succeeding activity causes the end date of an activity to be forced to be the same as the start date for that activity.
- Editing of the data due to date overlap as described above has caused the start date of an activity to be forced to be the same as the end data of that activity.

For completeness these zero duration activities are included in the dataset.

Adding in sweep9 (2013) activity data

As with the 2008 data sweep, the 2013 data sweep used dependent interviewing for the cohort members activity data, allowing for more accurate linking of sweep9 activity data to that of previous sweeps . The cohort members can be grouped according to the responses that they made to the economic activity section of the 2013 data sweep questionnaire as detailed below. The sweep9 data is added to the NCDS activity histories with variable JSLOT in the range 80-96.

Group	Sweep 9 response to economic activty questions	Number of cases
1	CM reports doing same activity as last sweep continuously ECONLAST/N9ECLAST=1 AND CJCONT/N9CJCONT=1	5579
2	CM last interviewed In sweeps 7 or 8. Not doing same activity as last swp or not contiously ECONLAST/N9ECLAST=2 OR CjCONT/N9CJCONT=2 and activity history is asked as expected.	2642
3	CM reports that activity at lastsweep incorrectly recorded (ECONLAST/N9ECLAST=3) OR CM not interviwed in last 2 sweeps (FFSWEEP2=2) and reports still doing activity from lastsweep/January 2004 and continously (N9CURACT=1 and N9CURACS=1)	315
4	CM reports that activity at lastsweep incorrectly recorded (ECONLAST/N9ECLAST=3) OR CM not interviwed in last 2 sweeps (FFSWEEP2=2) and reports not still doing activity from lastsweep/January 2004 or not continously (N9CURACT<>1 or N9CURACS<>1)	173
5	CM reports that activity at lastsweep incorrectly recorded (ECONLAST/N9ECLAST=3) OR CM not interviwed in last 2 sweeps (FFSWEEP2=2) and script error (-2) at variables N9ACTVT2 to N9CURACT.	93
6	CM reports not doing same activity as sweep 7 or 8 or no longer doing corrected activity from lastsweep/Jan2004 but did not enter activity history loop (n9ahistf=1). Asked about sweep9 current activity in variables N9CURAC2 to N9CURSTY.	213
7	CM refuses at n9eclast(whether doing same activity as sweep 7 or 8 N=12) or n9curact (whether still doing corrected activity from lastsweep/Jan2004 N=17). Asked about sweep9 current activity in variables N9CURAC2 to N9CURSTY.	29
8	CM entered activity history records but did not report last activity history activity as still current or not continously	88
9	CM refused when asked if still doing same activity as last sweep or corrected activity from lastsweep/Jan2004 continously (n9cjcont/ N9curacs)	5

See appendix3 for treatment of the sweep9 activity data for each of these groups.

Some general points about the inclusion of the sweep9 activity data is as follows:

Where an activity starts at the same time as a previous sweep current activity, and is not a duplicate, the previous activity end date is computed as the same as its start (0 duration) in order to retain that activity data for completeness.

At sweep9, where the cohort member reports still doing the same job as the previous sweep, if the swp9 soc2010 code does not correspond to the previous sweeps soc2000 code according to the ONS tables (http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/index.html link 'Relationship between SOC2010 and SOC2000') the activity is split into two at the sweep9 interview date in order to preserve the soc2000 code as coded at both the previous sweep and sweep9

Where socio-economic codes are repeated in the sweep9 data (i.e. where the activity at 2004/last sweep or latest activity history activity is still current), the socio economic data, including employment status used in the NCDS activity histories are taken from the N9C variables even if these differ from the N9J /N9A variables.
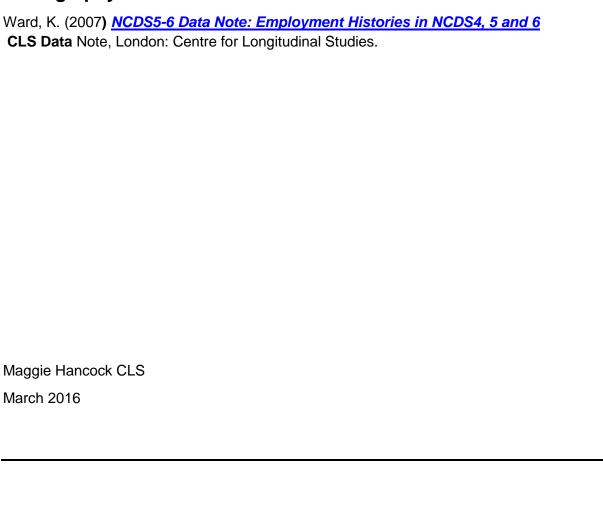
Completeness of NCDS Activity Histories

In order to judge how complete a cohort member's activity history is, the following variables are computed: variable HISTLEN reports the total number of months in the cohort member's activity history; variable OVERLAP reports the number of months where two or more full time activities overlap, that is where the truncation procedure described above was not operated; variable GAP reports the number of months where no activity is recorded for a cohort member. Variables POVERLAP and PGAP report the number of months with overlapping full time activities or no activity reported, as a percentage of the total activity history.

Of the 14,740 cases with at least one activity reported, variable HISTLEN can be computed for 14,225 cases only due to 515 cases having missing dates. A summary of these completeness variables is found below:

Variable	Statistic reported by variable	Number of cases with valid values	% of valid cases with variable value of zero	Statistics of non-zero values for valid cases				
				N	Min	Max	Mean	StdDev
HISTLEN	Number of months in CM's activity history	14225	0	14225	1	480	314.2	141.4
OVERLAP	Number of months where two or more full time activities overlap	14740	98.9	155	1	34	6.9	6.1
POVERLAP	% of activity history where two or more full time activities overlap	14225	99.4	92	1	36	3.5	4.5
GAP	Number of months where no activity is recorded	14740	54.3	6736	1	310	17.36	31.6
PGAP	% of activity history where no activity is recorded	14225	73.8	3733	1	100	12.23	16.8

Bibliography

Ward, K. (2007**) *NCDS5-6 Data Note: Employment Histories in NCDS4, 5 and 6*
 CLS Data Note, London: Centre for Longitudinal Studies.

Maggie Hancock CLS

March 2016

Appendix 1: Variables in NCDS Activity Histories Dataset

Variable	Variable Label
WSWEEP04	Sweep4: Wh participated and wh any activity data
WSWEEP05	Sweep5: Wh participated and wh any activity data
WSWEEP37	Sweep537: Wh participated and wh any activity data
WSWEEP06	Sweep6: Wh participated and wh any activity data
WSWEEP07	Sweep7: Wh participated and wh any activity data
WSWEEP08	Sweep8: Wh participated and wh any activity data
WSWEEP09	Sweep9: Wh participated and wh any activity data
CMSEX	CM's sex
NUMACTS	Total number of activities
ACTNUM	Activity number
LSTSWCMC	Century month[6] of latest interview
JSTCMC	Activity start century month[6]
JENDCMC	Activity end century month[6]
JDUR	Activity duration in months (truncated at latest sweep)
JACTIV	Activity activity type
JFTPT	Activity full-time or part-time (work activities only
JSTMTH	Activity start month
JSTYR	Activity start year
JENDMTH	Activity end month
JENDYR	Activity end year
HISTLEN	Length of known activity history incl. gaps (months)
OVERLAP	Total overlap in activities (months)
POVERLAP	Percentage overlap in activities (of total history)
GAP	Total gap in activities (months)
PGAP	Percentage gap in activities (of total history)
CERRORS	Case has activity history errors
AERRORS	Activity has activity history errors
AERROR1	Act.err: End date overlaps next activity, truncated
AERROR2	Act.err: Start date missing
AERROR3	Act.err: End date missing
AERROR4	Act.err: End date before start date, set to start date, dur set to -8 dk
AERROR5	Act err: Non f/t ed activity<age16, start set to 3 1974
AERROR6	Act err: Act. curr. at latest swp but start<sweep date, dur set to 0
AERROR7	Act err: Forced start of activity
AERROR8	Act err: Forced end of activity
JYLEFT	Activity Reason left job (s4, s6,s7,s8,9 work activities)
JYLEFTX	Activity Reason left job (detailed s537,s6,s7,s8,9)
JYLEFTO	Activity Reason left job (detailed s6,s7,s8,s9) other specified
JEMPST	Activity employment status (s5,s6,s7,s8,s9 work activities)
J91SEG	Activity SEG '91(s5,s537,s6,s7,s8,s9 work activities)
J90SOC	Activity soc90 (s5,s537,s6,s7,s8,s9 work activities)
J91SC	Activity social class '91 (s5,s537,s6,s7,s8,s9 work activities)
J2KSOC	Activity soc2000 (s7,s8,s9 work activities)

---

[6] Century month is calculated as the number of months since 1900

Variable	Variable Label
J10SOC	Activity soc2010 (s9 work activities)
JNSSEC	Activity NS-SEC (long version - s7,s8,s9 work activities)
J8NSSEC	Activity NS-SEC8 (analytic version - s7,s8,s9 work activities)
J07SIC	Activity S.I.C. 2007 (s9 current work activity)
J03SIC	Activity S.I.C. 2003 (s8 work activities)
J92SIC	Activity S.I.C. 1992 (s7 current work activity)
JSIC	Activity S.I.C. (s6 current work activity)
JTORG	Activity Type of organisation CM works for (s537,s6 work activities)
JCAMGP	Activity Cambridge Sub-Group & Title Ident. (s5 work activities)
JEMPCC	Activity Employment Status - check code (s5 work activities)
JCAMSM	Activity Cambridge Scale Score - Males (s5 work activities)
JCAMSF	Activity Cambridge Scale Score - Females (s5 work activities)
J88ISC	Activity ISCO '88-Int Stan Class of Occupns (s5 work activities)
JKOS	Activity KOS - Key Occupns for Stats (s5 work activities)
JCODOT	Activity CODOT - 3 digit classification (s5 work activities)
JOP80C	Activity OPCS 1980 - Operational Code (s5 work activities)
JOP80G	Activity OPCS 1980 - Occupation Unit Group (s5 work activities)
J81SEG	Activity Socio-Economic Group 1981 (s5 work activities)
J81RGS	Activity Social Class RG''s 1981 (s5 work activities)
JOP70G	Activity OPCS 1970 - Occupation Unit Group (s5 work activities)
J71SEG	Activity Socio-Economic Group 1971 (s5 work activities)
J71RGS	Activity Social Class RG''s 1971 (s5 work activities)
JRG66G	Activity RGs 1966 - Occupation Unit Group (s5 work activities)
J66SEG	Activity Socio-Economic Group 1966 (s5 work activities)
J66RGS	Activity Social Class RG''s 1966 (s5 work activities)
JRG61G	Activity RGs 1961 - Occupation Unit Group (s5 work activities)
J61SEG	Activity Socio-Economic Group 1961 (s5 work activities)
J61RGS	Activity Social Class RG''s 1961 (s5 work activities)
J90GSC	Activity Goldthorpe Social Class 1990 (s5 work activities)
J80GSC	Activity Goldthorpe Social Class 1980 (s5 work activities)
J70GSC	Activity Goldthorpe Social Class 1970 (s5 work activities)
JHGCG	Activity Hope-Goldthorpe Collapsed Group (s5 work activities)
JHGSS	Activity Hope-Goldthorpe Scale Score (s5 work activities)
J70SEG	Activity Socio-economic Group 1970 (s4 first & current work activities)
J70SC	Activity Social Class 1970 (s4 first & current work activities)
JINDUS	Activity industry (s4 first & current work activities)
J80CO	Activity co80 (s4 first & current work activities)
J80SC	Activity social class '80 (s4 first & current work activities)
J80SEG	Activity SEG '80 (s4 first & current work activities)
INTMON04	Month of sweep 4 interview
INTYR04	Year of sweep 4 interview
INTCMC04	Century month[7] of sweep 4 interview
INTMON05	Month of sweep 5 interview
INTYR05	Year of sweep 5 interview

---

[7] Century month is calculated as the number of months since 1900

Variable	Variable Label
INTCMC05	Century month[8] of sweep 5 interview
INTMON37	Month of age37 basic skill survey interview (estimate)
INTYR37	Year of age37 basic skill survey interview (estimate)
INTCMC37	Century month of age37 basic skill survey interview (estimate)
INTMON06	Month of sweep 6 interview
INTYR06	Year of sweep 6 interview
INTCMC06	Century month[8] of sweep 6 interview
INTMON07	Month of sweep 7 interview
INTYR07	Year of sweep 7 interview
INTCMC07	Century month[8] of sweep 7 interview
INTMON08	Month of sweep 8 interview
INTYR08	Year of sweep 8 interview
INTCMC08	Century month[8] of sweep 8 interview
INTMON09	Month of sweep 9 interview
INTYR09	Year of sweep 9 interview
INTCMC09	Century month[8] of sweep 9 interview
JLSTSWP	Latest sweep activity mentioned at
JLSTMTH	Month of latest sweep activity mentioned at
JLSTYR	Year of latest sweep activity mentioned at
JLSTCMC	Century month[8] of latest sweep activity mentioned at
JSLOT	Sweep and slot number that activity was reported in

---

[8] Century month is calculated as the number of months since 1900

Appendix 2: Cleaning undertaken during construction of activity histories:

Start or end dates edited or activity slots dropped.
Note the number in brackets is the slot number, see slots number descriptions below.

if(ncdsid='N10248K') ajbstyr(17)=1977
if(ncdsid='N10322C') ajbendyr(72)=2004; ajbstyr(73)=2004
if(ncdsid='N10997R') ajbendmth(25)=5 ; ajbendyr(25)=1976
if(ncdsid='N11135F') ajbstyr(69)=1992
if(ncdsid='N11189V') ajbendyr(70)=aswpyr(700); ajbendmth(70)=aswpmth(700);
ajbstyr(71)=aswpyr(700); ajbstmth(71)=aswpmth(700); jswp8=1
if(ncdsid='N11210Z') ajbstmth(13)=3; ajbstmth(58)=3
if(ncdsid='N11818B') ajbendyr(70)=1994 ; ajbstyr(71)=1994 ; ajbendyr(71)=1996; ajbstyr(72)=1996
if(ncdsid='N11835C') ajbendmth(16)=8 ; ajbendyr(16)=1983
if(ncdsid='N12100Y') slot22 dropped
if(ncdsid='N12295A') slot25-36 dropped
if(ncdsid='N12331M') ajbendmth(25)=9 ; ajbendyr(25)=1978
if(ncdsid='N12838K') ajbendmth(13)=10 ; ajbendyr(13)=1983
if(ncdsid='N12936L') ajbendyr(15)=1981
if(ncdsid='N12945M') ajbstmth(15)=7
if(ncdsid='N13344W') slot68 dropped
if(ncdsid='N13440V') ajbstyr(27)=1980
if(ncdsid='N13477J') ajbendmth(14)=9 ; ajbendyr(14)=1986
if(ncdsid='N13596Q') slot26 dropped
if(ncdsid='N13810B') ajbendyr(71)=2008; ajbendmth(71)=2; ajbstyr(72)=2008; ajbstmth(72)=2
if(ncdsid='N14325X') ajbendmth(25)=1
if(ncdsid='N14581M') ajbendyr(70)=1992 ; ajbstyr(71)=1992 ; ajbendyr(71)=1994; ajbstyr(72)=1994 ;
ajbendyr(72)=1996; ajbstyr(73)=1996
if(ncdsid='N15793B') ajbactiv(70)=1
if(ncdsid='N17134E') ajbendyr(25)=1980
if(ncdsid='N17235J') ajbstyr(25)=1977
if(ncdsid='N17246M') ajbstmth(71)=4
if(ncdsid='N17281Q') ajbstmth(71)=8
if(ncdsid='N17329P') ajbstyr(16)=1977
if(ncdsid='N17329P') slot15 dropped
if(ncdsid='N17487C') ajbstmth(3)=9
if(ncdsid='N17790G') ajbstmth(71)=8
if(ncdsid='N18068R') slot25 dropped
if(ncdsid='N19150N') ajbstmth(70)=5; ajbendmth(70)=5; ajbstmth(71)=6
if(ncdsid='N19191X') ajbendmth(27)=6 ; ajbendyr(27)=1979
if(ncdsid='N19284B') ajbstyr(26)=1987
if(ncdsid='N19322Q') ajbendmth(25)=6; ajbendyr(25)=1987; ajbendmth(26)=12; ajbendyr(26)=1990;
if(ncdsid='N19535B') slots66-68 dropped
if(ncdsid='N19917M') ajbendyr(17)=1981
if(ncdsid='N21380X') ajbstmth(37)=8; ajbstyr(37)=1974
if(ncdsid='N22087Z') ajbendyr(25)=1979
if(ncdsid='N22560D') ajbstyr(25)=1981
if(ncdsid='N22577N') ajbstyr(13)=1980
if(ncdsid='N22632C') ajbstyr(69)=aswpyr(700); ajbstmth(69)= aswpmth(700)
if(ncdsid='N23029T') slot25 dropped
if(ncdsid='N23482L') ajbstmth(67)=2 ; ajbstyr(67)=1994
if(ncdsid='N23589W') ajbendmth(70)=7;   ajbstmth(71)=7
if(ncdsid='N24005P') slot68 dropped
if(ncdsid='N24057B') ajbendmth(20)=12 ; ajbendyr(20)=1985
if(ncdsid='N24294N') compute ajbstmth(80)=1; ajbstyr(80)=2004

if(ncdsid='N24353G') ajbendmth(14)=7
if(ncdsid='N24437K') ajbendyr(25)=1976
if(ncdsid='N24560M') slot68 dropped
if(ncdsid='N24623K') ajbstmth(57)=5
if(ncdsid='N24712K') ajbendmth(28)=10
if(ncdsid='N24840S') ajbendmth(70)=12; ajbstmth(71)=12
if(ncdsid='N24958F') ajbendmth(25)=2 ; ajbendyr(25)=1983
if(ncdsid='N25345L') ajbstyr(curr8)=aswpyr(800); ajbstmth(curr8)=aswpmth(800)
if(ncdsid='N25828A') slot27 dropped
if(ncdsid='N26015Z') ajbstmth(58)=3 ; ajbstyr(58) =1974
if(ncdsid='N26168R') slots25-28 dropped
if(ncdsid='N26237M') ajbendyr(25)=1979
if(ncdsid='N26422L') ajbendmth(68)=8 ; ajbendyr(68)=1991
if(ncdsid='N26422L') ajbstyr(69)=1998
if(ncdsid='N26567C') ajbendmth(19)=10 ; ajbendyr(19)=1987
if(ncdsid='N27244Q') slot25 dropped
if(ncdsid='N27572D') ajbendmth(25)=8 ; ajbendyr(25)=1980
if(ncdsid='N27724B') ajbstyr(28)=1986
if(ncdsid='N28007H') ajbendyr(26)=1983 ; ajbendmth(26)=11
if(ncdsid='N28135Q') slot27 dropped
if(ncdsid='N28151Q') slot26 dropped
if(ncdsid='N28152R') slot25 dropped
if(ncdsid='N28159Y') ajbendmth(13)=9 ; ajbendyr(13)=1974
if(ncdsid='N28237V') ajbendyr(70)=1985 ; ajbstyr(71)=1985 ; ajbendyr(71)=1990; ajbstyr(72)=1990 ;
ajbendyr(72)=1995; ajbstyr(73)=1995
if(ncdsid='N28273Z') ajbstmth(71)=12

Slot numbers   (1)   "Swp 4: In work 1"
                        (2)   "Swp 4: In work 2"
                        (3)   "Swp 4: In work 3"
                        (4)   "Swp 4: In work 4"
                        (5)   "Swp 4: Out of work 1"
                        (6)   "Swp 4: Out of work 2"
                        (7)   "Swp 4: Out of work 3"
                        (8)   "Swp 4: Out of work 4"
                        (9)   "Swp 4: Looking for work 1"
                        (10)   "Swp 4: Looking for work 2"
                        (11)   "Swp 4: Looking for work 3"
                        (12)   "Swp 4: Looking for work 4"
                        (13)   "Swp 5: In work 1"
                        (14)   "Swp 5: In work 2"
                        (15)   "Swp 5: In work 3"
                        (16)   "Swp 5: In work 4"
                        (17)   "Swp 5: In work 5"
                        (18)   "Swp 5: In work 6"
                        (19)   "Swp 5: In work 7"
                        (20)   "Swp 5: In work 8"
                        (21)   "Swp 5: In work 9"
                        (22)   "Swp 5: In work 10"
                        (23)   "Swp 5: In work 11"
                        (24)   "Swp 5: In work 12"
                        (25)   "Swp 5: Not in work 1"
                        (26)   "Swp 5: Not in work 2"
                        (27)   "Swp 5: Not in work 3"
                        (28)   "Swp 5: Not in work 4"
                        (29)   "Swp 5: Not in work 5"
                        (30)   "Swp 5: Not in work 6"
                        (31)   "Swp 5: Not in work 7"

(32) "Swp 5: Not in work 8"
(33) "Swp 5: Not in work 9"
(34) "Swp 5: Not in work 10"
(35) "Swp 5: Not in work 11"
(36) "Swp 5: Not in work 12"
(37) "Swp 537: In work 1"
(38) "Swp 537: In work 2"
(39) "Swp 537: In work 3"
(40) "Swp 537: In work 4"
(41) "Swp 537: In work 5"
(42) "Swp 537: In work 6"
(43) "Swp 537: In work 7"
(44) "Swp 537: In work 8"
(45) "Swp 537: In work 9"
(46) "Swp 537: In work 10"
(47) "Swp 537: Looking for work/unemployed"
(58) "Swp 6: Curract"
(57) "Swp 6: Prevact 1"
(56) "Swp 6: Prevact 2"
(55) "Swp 6: Prevact 3"
(54) "Swp 6: Prevact 4"
(53) "Swp 6: Prevact 5"
(52) "Swp 6: Prevact 6"
(51) "Swp 6: Prevact 7"
(50) "Swp 6: Prevact 8"
(49) "Swp 6: Prevact 9"
(48) "Swp 6: Prevact 10"
(69) "Swp 7: Curract"
(68) "Swp 7: Prevact 1"
(67) "Swp 7: Prevact 2"
(66) "Swp 7: Prevact 3"
(65) "Swp 7: Prevact 4"
(64) "Swp 7: Prevact 5"
(63) "Swp 7: Prevact 6"
(62) "Swp 7: Prevact 7"
(61) "Swp 7: Prevact 8"
(60) "Swp 7: Prevact 9"
(59) "Swp 7: Prevact 10"
(70) "Swp 8: Activity 1"
(71) "Swp 8: Activity 2"
(72) "Swp 8: Activity 3"
(73) "Swp 8: Activity 4"
(74) "Swp 8: Activity 5"
(75) "Swp 8: Activity 6"
(76) "Swp 8: Activity 7"
(77) "Swp 8: Activity 8"
(78) "Swp 8: Activity 9"
(79) "Swp 8: Activity 10"
(80) " Swp9: 2004/lstswp activity"
(81) "Swp9: activity history activity1"
(82) "Swp9: activity history activity2"
(83) "Swp9: activity history activity3"
(84) "Swp9: activity history activity4"
(85) "Swp9: activity history activity5"
(86) "Swp9: activity history activity6"
(87) "Swp9: activity history activity7"
(88) "Swp9: activity history activity8"
(89) "Swp9: activity history activity9"

(90) "Swp9: activity history activity10"
(91) "Swp9: activity history activity11"
(92) "Swp9: activity history activity12"
(93) "Swp9: activity history activity13"
(94) "Swp9: activity history activity14"
(95) "Swp9: activity history activity15"
(96) "Swp9: current activity if not obtained"

Appendix 3: Adding sweep9 activities to the NCDS activity histories

The sweep9 data to be added to the NCDS activity histories was grouped according to responses to the initial questions in the economic activity section of the 2013 data sweep questionnaire and treated accordingly as detailed beloiw.

Swp9 group		Number of cases	sweep9 subgroup		Number of cases	Work history amendment
1	CM reports doing same activity as last sweep continuously ECONLAST/N9ECLAST=1 AND CJCONT/N9CJCONT=1	5579	1.1	CM last interviewed in sweep8, jobtitle has not changed (N9CJTIT=2), N9CEMPST and N9CJORG match employment status and organisation type from sweep8, and N9S2010 corresponds to sweep8 soc2000 code using tables from ONS website (see section 5.6)	3228-597	Sweep8 current activity values J10SOC and J07SIC updated with values for N9CS2010 and N9CSIC (note other socio-economic vars remain as coded at sweep8) and jlstswp and JLSTSWP, JLSTMTH, JLSTYR, JLSTCMC(Latest sweep activity mentioned at and dates) amended to sweep9.
			1.2	CM last interviewed in sweep8, job title reported changed at N9CJTIT or N9CEMPST and N9CJORG don't match employment status and organisation type from sweep8 or N9S2010 does not correspond to sweep8 soc2000 code using tables from ONS website (see section 5.6)	2151+597	Sweep8 current activity end date will be set to sweep9 interview date and a new sweep9 current activity is generated from the sweep9 current activity data (variables ND9ECACT, N9CEMPST to N9CSIC) with start date of sweep9 interview date.
			1.3	CM last interviewed in sweep7, jobtitle has not changed (N9CJTIT=2) , N9CEMPST and N9CJORG match employment status and organisation type from sweep8 and and N9S2010 corresponds to sweep7 soc2000 code using tables from ONS website (see section 5.6)	110-22	Sweep7 current activity values J10SOC and J07SIC updated with values for N9CS2010 and N9CSIC (note other socio-economic vars remain as coded at sweep8) and jlstswp and JLSTSWP, JLSTMTH, JLSTYR, JLSTCMC(Latest sweep activity mentioned at and dates) amended to sweep9.
			1.4	CM last interviewed in sweep7, job title reported changed at N9CJTIT or N9CEMPST and N9CJORG don't match employment status and organisation type from sweep8 or N9S2010 does not correspond to sweep7 soc2000 code using tables from ONS website (see section 5.6)	90+22	Sweep7 current activity end date will be set to sweep9 interview date and a new sweep9 current activity is generated from the sweep9 current activity data (variables ND9ECACT, N9CEMPST to N9CSIC) with start date of sweep9 interview date.

Swp9 group		Number of cases	sweep9 subgroup		Number of cases	Work history amendment
2	CM last interviewed In sweeps 7 or 8. Not doing same activity as last swp or not continuously ECONLAST/N9ECLAST=2 OR CjCONT/N9CJCONT=2 and activity history is asked as expected.	2642	2	No subgroups for sweep9 group2 cases	2642	The end of the previous sweeps current activity is set as reported in variables N9AENDY N9AENDM in first activity history record (or set to that sweeps interview date if N9AENDY in first activity history record is missing, if just N9AENDM missing, endmonth set to June) . The end of each activity in the activity histories is set as reported in variables N9AENDY N9AENDM in the following activity (or set to the start of that ac tivity if variable N9AENDY in the following activity is missing). Activities added as reported in activity history records with start dates set to the end of the previous activity as reported in variables N9AENDY N9AENDM (or sweep 9 interview date if variable N9AENDY mssing and CM reports still doing activty at sweep9) . Sweep9 current activity is the last activity history record which CM reports as still doing and continuously (N9ACURAC=1 and N9ACURAS=1). Data for this activity is taken from sweep9 current activity data (variables ND9ECACT, N9CEMPST to N9CSIC)
3	CM reports that activity at lastsweep incorrectly recorded (ECONLAST/N9ECLAST=3) OR CM not interviwed in last 2 sweeps (FFSWEEP2=2) and reports still doing activity from lastsweep/January 2004 and continuously (N9CURACT=1 and N9CURACS=1)	315	3.1	No unfinished previous activity	31	Sweep 9 current activity data (corrected activity from lastsweep/Jan2004, jslot=80) is taken from sweep9 current activity data (variables ND9ECACT, N9CEMPST to N9CSIC).
			3.2	Start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) > start date for most recent previous activity.	135	As for sweep9 subgroup 3.1 above plus.... End dates for most recent previous activity set to the earliest of that sweeps interview date/ start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth).
			3.3	Start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) <= start date for most recent previous activity and CM last interviewed in sweeps 7 or 8.	87	As for sweep9 subgroup 3.1 above plus.... End date of most recent previous activity is set to its start date. Start date of corrected activity from lastsweep/Jan2004 (jslot=80) set to start of most recent previous activity.
			3.4	Start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) <= start date for most recent previous activity and CM not interviewed in sweeps 7 or 8.	37	As for sweep9 subgroup 3.1 above plus.... End date of most recent previous activity is set to that sweeps interview date. Start date of corrected activity from lastsweep/Jan2004 (jslot=80) is set to most recent previous activity sweeps interview date.
			3.5	Start year for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) missing and CM last interviewed in sweeps 7 or 8.	9	As for sweep9 subgroup 3.1 above plus.... End date of most recent previous activity is set to that sweeps interview date. Start date of corrected activity from lastsweep/Jan2004 (jslot=80) is set to sweep7/sweep8 interview date
			3.6	Start year for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) missing and CM not interviewed in sweeps 7 or 8.	16	As for sweep9 subgroup 3.1 above plus.... End date of most recent previous activity is set to that sweeps interview date. Start date of corrected activity from lastsweep/Jan2004 (jslot=80 amd sweep9 current activity) is set to January 2004.

Swp9 group		Number of cases	sweep9 subgroup		Number of cases	Work history amendment
4	CM reports that activity at lastsweep incorrectly recorded (ECONLAST/N9ECLAST=3) OR CM not interviewed in last 2 sweeps (FFSWEEP2=2) and reports not still doing activity from lastsweep/January 2004 or not continously (N9CURACT<>1 or N9CURACS<>1)	173	4.1	No unfinished previous activity	19	The data (including start date) for corrected activity from lastsweep/Jan2004 (jslot=80) is taken from sweep9 variables N9ACTIV,N9STMO, N9STYR, N9JEMPST to N9JSC. The end of the corrected activity from lastsweep/Jan2004 (jslot=80) is set as reported in variables N9AENDY N9AENDM in first activity history record (or set to that sweeps interview date if N9AENDY in first activity history record is missing, if just N9AENDM missing, endmonth set to June) . The end of each activity in the activity histories is set as reported in variables N9AENDY N9AENDM in the following activity (or set to the start of that activity if variable N9AENDY in the following activity is missing). Activities added as reported in activity history records with start dates set to the end of the previous activity as reported in variables N9AENDY N9AENDM (or sweep 9 interview date if variable N9AENDY mssing and CM reports still doing activty at sweep9) . Sweep9 current activity is the last activity history record which CM reports as still doing and continuously (N9ACURAC=1 and N9ACURAS=1). Data for this activity is taken from sweep9 current activity data (variables ND9ECACT, N9CEMPST to N9CSIC)
			4.2	Start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) > start date for most recent previous activity.	94	As for sweep9 subgroup 4.1 above plus.... End dates for most recent previous activity set to the earliest of that sweeps interview date/ start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth).
			4.3	Start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) <= start date for most recent previous activity and CM last interviewed in sweeps 7 or 8.	26	As for sweep9 subgroup 4.1 above plus.... End date of most recent previous activity is set to its start date. Start date of corrected activity from lastsweep/Jan2004 (jslot=80) set to start of most recent previous activity.
			4.4	Start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) <= start date for most recent previous activity and CM not interviewed in sweeps 7 or 8.	27	As for sweep9 subgroup 4.1 above plus.... End date of most recent previous activity is set to that sweeps interview date. Start date of corrected activity from lastsweep/Jan2004 (jslot=80) is set to most recent previous activity sweeps interview date.
			4.5	Start year for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) missing and CM last interviewed in sweeps 7 or 8.	4	As for sweep9 subgroup 4.1 above plus.... End date of most recent previous activity is set to that sweeps interview date. Start date of corrected activity from lastsweep/Jan2004 (jslot=80) is set to sweep7/sweep8 interview date.
			4.6	Start year for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) missing and CM not interviewed in sweeps 7 or 8.	3	As for sweep9 subgroup 4.1 above plus.... End date of most recent previous activity is set to that sweeps interview date. Start date of corrected activity from lastsweep/Jan2004 (jslot=80 and sweep9 current activity) is set to January 2004.

Swp9 group		Number of cases	sweep9 subgroup		Number of cases	Work history amendment
5	CM reports that activity at lastsweep incorrectly recorded (ECONLAST/N9ECLAST=3) OR CM not interviewed in last 2 sweeps (FFSWEEP2=2) and script error (-2) at variables N9ACTVT2 to N9CURACT	93	5	No subgroups for sweep9 group5 cases	93	No sweep9 activities added to activity history dataset. Variable WSWEEP09 set to 2 (in sweep 9 but no activity data).
6	CM reports not doing same activity as sweep 7 or 8 or no longer doing corrected activity from lastsweep/Jan2004 but did not enter activity history loop (n9ahistf=1). Asked about sweep9 current activity in variables N9CURAC2 to N9CURSTY.	213	6.0x	Cm reports no longer doing sweep7/8 current activity (includes subgroups 6.00 N=13, 6.02 N=152, 6.03 N=28)	193	End date of most recent previous activity is set to that activity sweep's interview date. Plus treament for sweep9 current activity (slot96) dependent on value of x (second decimal value) as described below.
			6.2x	Start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) > start date for most recent previous activity (includes subgroups 6.24 N=1, 6.25 N=6)	7	The data (including start date) for corrected activity from lastsweep/Jan2004 (jslot=80) is taken from sweep9 variables N9ACTIV,N9STMO, N9STYR, N9JEMPST to N9JSC. End dates for most recent previous activity set to the earliest of that sweeps interview date/ start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth). Plus treament for sweep9 current activity (slot96) dependent on value of x (second decimal value) as described below.
			6.3x	Start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) <= start date for most recent previous activity and CM last interviewed in sweeps 7 or 8. (includes subgroups 6.34 N=3).	3	As for sweep9 subgroup 6.2x above but... End date of most recent previous activity is set to its start date. Start date of corrected activity from lastsweep/Jan2004 (jslot=80) set to start of most recent previous activity.
			6.4x	Start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) <= start date for most recent previous activity and CM not interviewed in sweeps 7 or 8. (includes subgroups 6.45 N=3, 6.46 N=1)	4	As for sweep9 subgroup 6.2x above but... End date of most recent previous activity is set to that sweeps interview date. Start date of corrected activity from lastsweep/Jan2004 (jslot=80) is set to most recent previous activity sweeps interview date.
			6.5x	Start year for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) missing and CM last interviewed in sweeps 7 or 8. (includes subgroups 6.54 N=1)	1	As for sweep9 subgroup 6.2x above but... End date of most recent previous activity is set to that sweeps interview date. Start date of corrected activity from lastsweep/Jan2004 (jslot=80) is set to sweep7/sweep8 interview date.

Swp9 group		Number of cases	sweep9 subgroup		Number of cases	Work history amendment
			6.6x	Start year for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) missing and CM not interviewed in sweeps 7 or 8. (includes subgroups 6.60 N=1, 6.65 N=4)	5	As for sweep9 subgroup 6.2x above plus.... End date of most recent previous activity is set to that sweeps interview date. Start date of corrected activity from lastsweep/Jan2004 (jslot=80 amd sweep9 current activity) is set to January 2004.
			6.x0	Refused at n9curac2 (includes subgroups 6.00 N=13, 6.60 N=1)	14	Treament for end of most recent previous activity(before sweep9) dependent on value of x, first decimal value) as described above plus.. End of corrected activity from lastsweep/Jan2004 (jslot=80), if exists, set to own start. Sweep 9 current activity (slot96) is created with JACTIV=-8 (dk) and start date of sweep9 interview date set = n9cursty/m or sweep9 interview date if n9cursty missing.
			6.x2	Start of sweep 9 current activity (variables n9cursty/m) >=start of most recent previous activity (includes subgroups 6.02 N=152)	152	Treament for end of most recent previous activity(before sweep9) dependent on value of x, first decimal value) as described above plus.. Sweep 9 current activity data (slot96) is taken from sweep9 current activity data (variables ND9ECACT, N9CEMPST to N9CSIC) with start date set to n9cursty/m or sweep9 interview date if n9cursty missing.
			6.x3	Start of sweep 9 current activity (variables n9cursty/m) < start of most recent previous activity (includes subgroups 6.03 N=28)	28	As for sweep9 subgroup 6.x2 above but... Sweep 9 current activity data (slot96) start date set to most recent previous activity sweeps interview date.
			6.x4	Start of sweep 9 current activity (variables n9cursty/m) >=start of (includes subgroups 6.03 N=28)(jslot=80) and interviewed in sweeps7 or 8 (includes subgroups 6.24 N=1, 6.34 N=3, 6.54 N=1))	5	As for sweep9 subgroup 6.x2 above but... End of corrected activity from lastsweep/Jan2004 (jslot=80) set to the earliest of [sweep7 / sweep8 interview date , n9cursty/m] Sweep 9 current activity data (slot96) start date set to n9cursty/m or sweep9 interview date is n9cursty missing.
			6.x5	Start of sweep 9 current activity (variables n9cursty/m) >=start of corrected activity from lastsweep/Jan2004 (jslot=80) and not interviewed in sweeps7 or 8 (includes subgroups 6.25 N=6, 6.45 N=3, 6.65 N=4))	13	As for sweep9 subgroup 6.x2 above but... End of corrected activity from lastsweep/Jan2004 (jslot=80) set to the earliest of [January 2004 , n9cursty/m] Sweep 9 current activity data (slot96) start date set to n9cursty/m or sweep9 interview date if n9cursty missing.

Swp9 group		Number of cases	sweep9 subgroup		Number of cases	Work history amendment
			6.x6	Start of sweep 9 current activity (variables n9cursty/m) <start of corrected activity from lastsweep/Jan2004 (jslot=80) and not interviewed in sweeps7 or 8 (includes subgroups 6.46 N=1)	1	As for sweep9 subgroup 6.x2 above but... End of corrected activity from lastsweep/Jan2004 (jslot=80) set to January 2004. Sweep 9 current activity data (slot96) start date set to n9cursty/m or sweep9 interview date if cursty missing.
7	CM refuses at n9eclast(whether doing same activity as sweep 7 or 8 N=12) or n9curact (whether still doing corrected activity from lastsweep/Jan2004 N=17). Asked about sweep9 current activity in variables N9CURAC2 to N9CURSTY.	29	7.0x	CM refused at n9eclast (includes subgroups 7.00 N=15, 7.01 N=1, 7.02 N=6, 7.03 N=2).	24	End date of most recent previous activity before sweep9 is set to that activity sweep's interview date. Plus treament for sweep9 current activity (slot96) dependent on value of x (second decimal value) as described below.
			7.2x	Start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) > start date for most recent previous activity (includes subgroups 7.20 N=1)	1	The data (including start date) for corrected activity from lastsweep/Jan2004 (jslot=80) is taken from sweep9 variables N9ACTIV,N9STMO, N9STYR, N9JEMPST to N9JSC. If missing set to -8 don't know. End dates for most recent previous activity set to the earliest of that sweeps interview date/ start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth). Plus treament for sweep9 current activity (slot96) dependent on value of x (second decimal value) as described below.
			7.5x	Start year for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) missing and CM last interviewed in sweeps 7 or 8. (includes subgroups 7.50 N=1)	1	As for sweep9 subgroup 7.2x above but... End date of most recent previous activity is set to that sweeps interview date. Start date of corrected activity from lastsweep/Jan2004 (jslot=80) is set to sweep7/sweep8 interview date.
			7.6x	Start year for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) missing and CM not interviewed in sweeps 7 or 8. (includes subgroups 7.60 N=2, 7.65 N=1)	3	As for sweep9 subgroup 7.2x above but... End date of most recent previous activity is set to that sweeps interview date. Start date of corrected activity from lastsweep/Jan2004 (jslot=80 amd sweep9 current activity) is set to January 2004.
			7.x0	Refused at n9curac2 (sweep9 current activity asked, includes subgroups 7.00 N=15, 7.20 N=1, 7.50 N=1, 7.60 N=2)	19	Treament for end of most recent previous activity(before sweep9) dependent on value of x first decimal value) as described above plus... End of corrected activity from lastsweep/Jan2004 (jslot=80), if exists, set to own start. Sweep 9 current activity (slot96) is created with JACTIV=-8 (dk) and start date of sweep9 interview date set = n9cursty/m or sweep9 interview date if n9cursty missing.
			7.x1	No unfinished previous activity (includes subgroups 7.01 N=1).	1	Sweep 9 current activity data (slot96) is taken from sweep9 current activity data (variables ND9ECACT, N9CEMPST to N9CSIC) with start date set to n9cursty/m or sweep9 interview date is n9cursty missing.

Swp9 group		Number of cases	sweep9 subgroup		Number of cases	Work history amendment
			7.x2	Start of sweep 9 current activity (variables n9cursty/m) >=start of most recent previous activity (includes subgroups 7.02 N=6)	6	Treament for end of most recent previous activity(before sweep9) dependent on value of x first decimal value) as described above plus.. Sweep 9 current activity data (slot96) is taken from sweep9 current activity data (variables ND9ECACT, N9CEMPST to N9CSIC) with start date set to n9cursty/m or sweep9 interview date if n9cursty missing.
			7.x3	Start of sweep 9 current activity (variables n9cursty/m) < start of most recent previous activity (includes subgroups 7.03 N=2)	2	As for sweep9 subgroup 7.x2 above but... Sweep 9 current activity data (slot96) start date set to most recent previous activity sweeps interview date.
			7.x5	Start of sweep 9 current activity (variables n9cursty/m) >=start of corrected activity from lastsweep/Jan2004 (jslot=80) and not interviewed in sweeps7 or 8 (includes subgroups 7.65 N=1)	1	As for sweep9 subgroup 7.x2 above but... End of corrected activity from lastsweep/Jan2004 (jslot=80) set to the earliest of [January 2004 , n9cursty/m] Sweep 9 current activity data (slot96) start date set to n9cursty/m or sweep9 interview date if n9cursty missing.
	CM entered activity history records but did not report last activity history activity as still current or not continously	88	8.0x	Cm reported not doing same activity as previous sweep (7 or 8).	78	End date of most recent previous activity before sweep9 is set as reported in variables N9AENDY N9AENDM in first activity history record (or set to that sweeps interview date if N9AENDY in first activity history record is missing, if just N9AENDM missing, endmonth set to June) . Activities added as reported in activity history records with start dates set to the end of the previous activity as reported in variables N9AENDY N9AENDM (or sweep 9 interview date if variable N9AENDY mssing and CM reports still doing activty at sweep9) . The end of each activity in the activity histories, except the last, is set as reported in variables N9AENDY N9AENDM in the following activity (or set to the start of that ac tivity if variable N9AENDY in the following activity is missing). Plus treament for last activity history activity and sweep9 current activity (slot96) dependent on value of x (second decimal value) as described below.

Swp9 group		Number of cases	sweep9 subgroup		Number of cases	Work history amendment
8			8.2x	Start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) > start date for most recent previous activity (includes subgroups 8.22 N=3, 8.23, N=2, 8.24 N=3)	8	End dates for most recent previous activity (before sweep9) set to the earliest of that sweeps interview date/ start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) The data (including start date) for corrected activity from lastsweep/Jan2004 (jslot=80) is taken from sweep9 variables N9ACTIV,N9STMO, N9STYR, N9JEMPST to N9JSC. The end of the corrected activity from lastsweep/Jan2004 (jslot=80) is set as reported in variables N9AENDY N9AENDM in first activity history record (or set to that sweeps interview date if N9AENDY in first activity history record is missing, if just N9AENDM missing, endmonth set to June) . Activities added as reported in activity history records with start dates set to the end of the previous activity as reported in variables N9AENDY N9AENDM (or sweep 9 interview date if variable N9AENDY mssing and CM reports still doing activty at sweep9) . The end of each activity in the activity histories, except the last, is set as reported in variables N9AENDY N9AENDM in the following activity (or set to the start of that ac tivity if variable N9AENDY in the following activity is missing). Plus treament for last activity history activity and sweep9 current activity (slot96) dependent on value of x (second decimal value) as described below.
			8.4x	Start date for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) <= start date for most recent previous activity and CM not interviewed in sweeps 7 or 8. (includes subgroups 8.44 N=1)	1	As for sweep9 subgroup 8.2x above but.... End date of most recent previous activity is set to that sweeps interview date. Start date of corrected activity from lastsweep/Jan2004 (jslot=80) is set to most recent previous activity sweeps interview date.
			8.6x	Start year for corrected activity from lastsweep/Jan2004 (n9styr n9stmth) missing and CM not interviewed in sweeps 7 or 8. (includes subgroups 8.62 N=1)	1	As for sweep9 subgroup 8.2x above but.... End date of most recent previous activity is set to that sweeps interview date. Start date of corrected activity from lastsweep/Jan2004 (jslot=80 and sweep9 current activity) is set to January 2004.
			8.x1	CM reports doing same activity as the last activity history activity now but not continously (N9ACURAC=1 and N9ACURAS=2, -8, -9 in last acthist rec). Start dates entered for asked sweep9 activity the same as start dates for last activity history activity (includes subgroups 8.01 N=5)	5	Treament for end of most recent previous activity(before sweep9), corrected activity from lastsweep/Jan2004 (slot80) and activity history records except the last, dependent on value of x first decimal value) as described above plus.. Last activity history activity is dropped. The data (including start date) for asked sweep9 current activity(jslot=96) is taken from sweep9 variables ND9ECACT,N9CURSTM, N9CURSTY, N9CEMPST to N9CSIC. If N9CURSTY is mmissing, start is set to sweep9 interview date.

Swp9 group		Number of cases	sweep9 subgroup		Number of cases	Work history amendment
			8.x2	CM reports doing same activity as the last activity history activity now but not continously (N9ACURAC=1 and N9ACURAS=2, -8, -9 in last acthist rec). Start dates entered for asked sweep9 activity not the same as start dates for last activity history activity (includes subgroups 8.02 N=26, 8.22 N=3, 8.62 N=1)	30	As for sweep9 subgroup 8.x1 above but... End of last activity history activity set to = start of same last activity history activity The data (including start date) for asked sweep9 current activity(jslot=96) is taken from sweep9 variables ND9ECACT,N9CURSTM, N9CURSTY, N9CEMPST to N9CSIC. If N9CURSTY is mmissing, start is set to sweep9 interview date.
			8.x3	CM reports not doing same activity as the last activity history activity now. Start dates entered for asked sweep9 activity the same as start dates for last activity history activity (includes subgroups 8.03 N=17, 8.23 N=2)	19	As for sweep9 subgroup 8.x1 above but... Last activity history activity is dropped. The data (including start date) for asked sweep9 current activity(jslot=96) is taken from sweep9 variables ND9ECACT,N9CURSTM, N9CURSTY, N9CEMPST to N9CSIC. If N9CURSTY is mmissing, start is set to sweep9 interview date.
			8.x4	CM reports not doing same activity as the last activity history activity now. Start dates entered for asked sweep9 activity later than start dates for last activity history activity (includes subgroups 8.04 N=27, 8.24 N=3. 8.44 N=1)	31	As for sweep9 subgroup 8.x1 above but... End of last activity history activity set to = start of same last activity history activity The data (including start date) for asked sweep9 current activity(jslot=96) is taken from sweep9 variables ND9ECACT,N9CURSTM, N9CURSTY, N9CEMPST to N9CSIC. If N9CURSTY is mmissing, start is set to sweep9 interview date.
			8.x5	CM reports not doing same activity as the last activity history activity now. Start dates entered for asked sweep9 activity earlier than start dates for last activity history activity (includes subgroups 8.05 N=2)	2	As for sweep9 subgroup 8.x1 above but... End of last activity history activity set to = start of same last activity history activity The data (except start date) for asked sweep9 current activity(jslot=96) is taken from sweep9 variables ND9ECACT, N9CEMPST to N9CSIC. Start date for asked sweep9 activity (jslot=96) is set to sweep9 interview date.
			8.x6	CM reports not doing same activity as the last activity history activity now. Start dates entered for both asked sweep9 current activity and last activity history activity are missing (includes subgroups 8.06 N=1)	1	As for sweep9 subgroup 8.x1 above but... Last activity history activity dropped as no start or end dates The data (except start date) for asked sweep9 current activity(jslot=96) is taken from sweep9 variables ND9ECACT, N9CEMPST to N9CSIC. Start date for asked sweep9 activity (jslot=96) is set to sweep9 interview date.

32

Swp9 group		Number of cases	sweep9 subgroup		Number of cases	Work history amendment
9	CM refused when asked if still doing same activity as last sweep or corrected activity from lastsweep/Jan2004 continously (n9cjcont/ N9curacs)	5	9.1	CM refused when asked if still doing same activity as last sweep continously (n9cjcont)	4	End of most recent previous activity(before sweep9) set to that activity sweeps interview date. A dummy activity history activity (slot81) is set up with start=sweep9 interview date, all activity details to be copied from most recent previous activity. c lastslot or slot80 (for corrected activity at lastswp or Jan2004)
			9.8	CM refused when asked if still doing corrected activity from lastsweep/Jan2004 continously (N9curacs)	1	End of corrected activity at lastswp or Jan2004 set to start of corrected activity at lastswp or Jan2004 . A dummy activity history activity (slot81) is set up with start=sweep9 interview date, all activity details to be copied from corrected activity at lastswp or Jan2004

Appendix 4: NCDS Activity Histories SPSS Dataset Filename

Dataset	SPSS dataset file name
NCDS Activity histories dataset	NCDS_activity_histories.sav