Institute of Education

UCL

1970 British Cohort Study -

Age 16 Survey Leisure and Television Diaries User Guide

User guide (Version 1)

April 2022

CENTRE FOR
LONGITUDINAL
STUDIES

UK RI

Economic
and Social
Research Council

Contact

For data queries, please contact help@ukdataservice.ac.uk

Questions and feedback about this user guide should be sent to clsfeedback@ucl.ac.uk.

How to cite this guide

Altintas, E., Campbell, S., Dennison, K., Gershuny, J., Peters, A., Shepherd, P., (2022) *1970 British Cohort Study Age 16 Survey Leisure and Television Diaries User Guide.* London: UCL Centre for Longitudinal Studies.

You should also acknowledge CLS following the guidance from https://cls.ucl.ac.uk/data-access-training/citing-our-data/

This guide was first published in April 2022 by the UCL Centre for Longitudinal Studies.

> Centre for Longitudinal Studies.
>
> Social Research Institute
>
> University College London
>
> 20 Bedford Way, London WC1H 0AL

The UCL Centre for Longitudinal Studies (CLS) is an Economic and Social Research Council (ESRC) Resource Centre based at the UCL Institution of Education (IOE), University College London. It manages four internationally-renowned cohort studies: the 1958 National Child Development Study, the 1970 British Cohort Study, Next Steps, and the Millennium Cohort Study. For more information, visit www.cls.ucl.ac.uk.

This document is available in alternative formats. Please contact the Centre for Longitudinal Studies for questions and feedback about this user guide: tel: +44 (0)20 7612 6875; email: clsfeedback@ucl.ac.uk

Contents

About the 1970 British Cohort Study

BCS70 is a multi-disciplinary longitudinal birth cohort study that follows the lives of more than 17,000 people born in England, Scotland and Wales in a single week of 1970.

Since 1970, cohort members have been surveyed at ages 5, 10, 16, 26, 30, 34, 38, 42 and most recently at age 46-48.

Over the course of cohort members' lives, BCS70 has collected information on health, physical, educational and social development, and economic circumstances among other factors.

BCS70 is run by the Centre for Longitudinal Studies (CLS), a research centre in the UCL Institute of Education, which is part of University College London.

BCS70 is core-funded by the Economic and Social Research Council (ESRC).

1. Introduction

This user guide accompanies the Leisure and Television Diaries data collected during the 1970 British Cohort Study (BCS70) Age 16 Survey, designed by Professor Neville Butler.

1.1 The BCS70 Age 16 Survey

The BCS70 Age 16 Survey, known as 'Youthscan' was carried out by the International Centre for Child Studies in 1986, under the direction of Professor Neville Butler of Bristol University.

The survey aimed to monitor the continued social, physical and educational development of the BCS70 cohort throughout their teenage years. It involved no less than 19 questionnaire instruments.

The Age 16 Sweep was the last time that parents were interviewed, being asked about family circumstances, health, income, housing and education.

A medical assessment took body measurements, blood pressure, pulse, vision, hearing and co-ordination measures.

Cohort members themselves completed a large number of questionnaires covering many aspects of life including education, leisure activities, smoking, drinking, relationships with family and friends, beliefs and aspirations for the future.

Cognitive assessments included Reading Test, Arithmetic, Vocabulary, Spelling and BAS Matrices.

Teachers provided information about the school's characteristics and the pupil's ability, behaviour and attendance.

In particular (for the purposes of this specific documentation), the teenage participants also completed 4-day diet, leisure and television diaries covering diet.

Full details regarding the conduct of fieldwork are provided in the BCS70 Age 16 Survey User Guide:

https://cls.ucl.ac.uk/wpcontent/uploads/2017/07/Guide_to_the_16-year_data.pdf

1.2 Collection of the Leisure and Television Diaries

The leisure and television diary data has not hitherto been available for analysis.

The data were coded into a machine readable format at Bath University under the direction of Sally Jones and Jonathan Gershuny during 1989 and 1990 and further processed by the Centre for Time Use Research in 2015 at the University of Oxford by Evrim Altintas and Giacomo Vagni. This work was supported by the British Academy Postdoctoral Research Fund, "The importance of the everyday" and John Fell Research Fund.

In 2020 and 2021, Stuart Campbell and Jonathan Gershuny at the Centre for Time User Research (now based at University College London Institute of Education) made final adjustments for release.

2. Contents of the Leisure and Television Diaries

Cohort members contacted in 1986 and 1987 were asked to complete a four-day Leisure Diary (covering a consecutive Friday, Saturday, Sunday and Monday), to answer some detailed questions about these four days and give details of their television viewing over the same period, in the form of a Television Diary. A copy of the A5 booklet used to collect this information is provided in the BCS70 Age 16 Survey- Document S: Leisure and Activity Diary: https://cls.ucl.ac.uk/wp-content/uploads/2017/07/docs.pdf

The leisure and television diary data are available as four separate datasets. The numbers in brackets indicate the number of cohort members in each of the datasets:

- three leisure diary files (7,243)
  - o the leisure diary episode file
  - o the leisure diary calendar file (derived from the leisure diary episode file)
  - o the leisure diary day aggregate file (derived from the leisure diary episode file)
- one television diary event file (6,946)

An outline of each diary is given below.

2.1 Leisure Diary Episodes

Of the 11,622 cohort members who completed one or more sections of the 16-year Youthscan study, 7,243 (62%) attempted to complete a leisure diary. Some only managed two or three days. Others left parts of the diary blank or provided entries that were illegible.

75 per cent of the diaries were collected between June and September, July being the most frequent month. Users working on time spent in educational activities should take a particular note of this.

The respondents were asked to provide 4-day diaries starting from Friday. Not all the respondents returned 4-day diaries and some returned extra diaries on Tuesday, Wednesday and Thursday. Most of the diaries on Tuesday are 1-2 episodes, giving information on activities done shortly after Monday midnight. Please see Appendix 5 for more information on extra diaries.

Variable EPNUM shows the episode number in the diary sequence. In several cases there were "empty episodes"; that is, a row where all variables other than identifier were missing. These extra empty episodes were deleted and episode number adjusted accordingly.

All the diaries start at 0:00 and end at 24:00 (see variables CLOCKSTH/ CLOCKSTM and CLOCKENDH/CLOCKENDM). In the original data file, diary start times were not identical. An extra episode was added to diaries that do not start at midnight. In some diaries the missing activity between midnight and first reported activity has been imputed and coded as "imputed sleep". Please see Appendix 5 for more information on imputation procedures.

The variables START and END show the cumulative start and end time of the activity. All diaries start at 0 minute and end at 1440 minute at 24:00.

Variable DURATION is the duration of each episode in minutes. The sum of this variable equals to 1440 minutes for each diary.

For WHOWITH, which contains the person cohort members were with during activities, Youthscan coders were given instructions to give priority to the diarist's family, then her or his friends, with others given the lowest priority. This variable

precludes analysis of the small number of cases where both a member of the family and friends/others were present.

Episodes

This dataset contains episode data for up to 4 days' activities for each diarist. The information provided in the completed diaries was broken down into 485,059 single events, or "episodes". Each event records a set of characteristics (activity, location, co-presence with others etc); the period of time during which none of these characteristics changes is defined as an "episode". Adopting the methodology commonly used in time-budget research, this episode data can either be analysed as a collection of single days or manipulated to provide a simulation of the activities undertaken in a given week.

Activity codes

Each activity (main, secondary and tertiary) was coded to 4-digit activity numeric codes adapted from a set used to code time-budget data collected as part of two ESRC-funded national surveys, undertaken in 1983/4 (Gershuny *et al*, 1986) and 1987 and a coding frame developed by the BBC for time-budget research carried out in the early 1980s (BBC Daily Life Survey, 1983-4, see Dyer, 1984 and Marles, 1985). The first two digits in the four-digit system run from 01 to 40 indicating a major activity group. The last two digits provide detail within each of these forty categories. A list of major activity groups is shown in the table below.

Primary and secondary activities were also coded using a now old 69 category list code from the Multinational Time Use Study (MTUS) (variables MAIN and SEC) and primary activity was also coded to the above 40 category list major activity group code (variable AV).

Major Activity Groups

1	Paid work	21	Walks, rambles
2	Paid work at home	22	Religious practices
3	Second job	23	Civic organisation
4	School/classes	24	Cinema, theatre
5	Work/school travel	25	Dance, party
6	Cook, wash up	26	Social club
7	Housework	27	Pubs
8	Odd jobs	28	Eating out
9	Gardening	29	Visiting friends
10	Shopping	30	Listening to radio
11	Child care	31	TV/video
12	Domestic travel	32	Listening to music
13	Dressing/hygiene	33	Study
14	Personal services	34	Read books
15	Eating at home	35	Read periodicals
16	Sleep/naps	36	Relaxing
17	Leisure travel	37	Conversation
18	Excursions	38	Entertain friends
19	Active sport	39	Knitting, sewing
20	Watching sport	40	Pastimes, hobbies

Reporting of activities

The free-format layout of the Youthscan diaries, where diarists could specify their own time-intervals, means that diarists varied in the way they chose to report their activities. Some diarists included very brief activities as main activities whereas others chose to report several different activities within large chunks of time. Because of the diary format, diarists may have failed to report subsidiary or very short activities.

Researchers planning analyses of activities such as listening to the radio, that often take place in the background of other activities, are advised to make use of secondary and tertiary activity code variables (SECONDACT, THIRDACT), since relying exclusively on the main activity variable (MAINACT) may under-estimate their occurrence.

This file will be of limited use to researchers analysing leisure behaviour over a period of time or aggregated over a day or week. Although the file includes all of the raw information a considerable amount of data manipulation is required to investigate questions such as "how many minutes per day were spent reading at home?".

The episode file enables analysis of activity variables constructed by combining various of the diary fields. Multitasking may be examined by combining the primary and secondary activities in a single episode; differentials in co-presence among different sorts of activities may be explored by combining activity with fields from the WHOWITH variable.

Please see section 3.4 for a description and further notes on the variables in this file, and the Appendix for detailed notes on corrections and imputations applied to these data.

2.2 Leisure Diary Calendar

This file gives a five-minute by five-minute account of all the diary fields, in a sequence of five (primary, secondary, tertiary, location, co-presence) 288 element vectors, where each case/row of data corresponds to a day. It also contains two

further 288 element vectors giving primary and secondary activities coded to the modern Multinational Time Use Study activity classification. The calendar file is the most easily approachable format for most serious analytic needs. Information lost, as compared to the episode file, is episodes of less than 2.5 minutes duration, of which there are very few.

Please see section 3.5 for a description and further notes on the variables in this file.

2.3 The Diary Day Aggregate file

The diary day aggregate file contains totals of the number of minutes per day spent in each of forty main activity groups and is for researchers who need summary time budget data, without having very specific questions. One use for this file is to investigate how long the diarists in the Youthscan cohort spent on each of the forty major activity groups. For most research purposes, the two activity categories "dress/personal care" and "sleep" should be amalgamated as so many teenagers reported their nightly activities as "went to bed; got up".

Aggregate time-use data may also be used to show how the amount of time spent in one activity affects the time spent in other ways. For example, the data may be used to construct curves comparing the elasticity of a flexible type of activity – such as participating in sport – with more fixed or regulated types of activities such as paid work.

Please see section 3.6 for a description and further notes on the variables in this file.

2.4 The Television Diary Event

This dataset contains 59,786 records – a record for each separate "episode" recorded in the television diaries over the relevant four-day period.

The television event file can be used to extract information about a particular programme, answering questions such as, "How many of the Youthscan diarists watched Coronation Street on Monday evenings?" and to provide answers to questions about which programmes the diarists had watched at a particular time on particular days, e.g. "What was being watched at 11.30 pm on Saturday evenings?". The Audience Research codes also allow analysis about the general type of

programme being watched. For each of these analyses, other items of data may be added – including the diarist's rating of the programme.

Duration of event

The television diary events can also be aggregated to show the total time per day spent watching different types of programmes. It is also possible to derive a daily or simulated-week-average for the total time spent watching television. However, many of the television diary events lack an estimate for duration. Also, it is not possible to differentiate between those diarists who didn't fill in the television diaries and those who never watched television. However, the total times can be compared with the equivalent figures derived from the "television and video" events in the leisure diaries.

Type of programme

The variable MAINACT summarises the type of TV programme or video watched. Where the programme's name was clear, the coders entered the code as "5000" and went on to record the programme name. Where the diarists had used general headings such as "children's TV", "cricket" or "film" the coders used a sequence of activity codes from 5001 to 5023, to preserve the information about the programme type. Code "5888" was used where a television diary event was probably a continuation of the last recorded event. Code "5999" was used to indicate where no TV programmes or videos had been watched/recorded by the diarists on a particular day.

Name of TV programme

The variable TVSTR2 provides information on the name of the programme watched (TVSTR is the original text provided by the coders – TVSTR2 is the recoded version). At the time of data entry, a decision was taken to restrict the length of original programme names (TVSTR), since they are usually identifiable from twelve characters. However, coders used a number of abbreviations and contractions that have made the data unnecessarily cumbersome. For example, variable TVSTR includes numerous different values for the weekly music programme "Top of the Pops" including "Top of Pops", "Top Pops", "Topofthepops", "Top of the p" and

"Top/pops".For this reason, TVSTR has been omitted from the deposited data in favour of the more useful recoded version.

The variable TVLIKE indicates how much the diarists enjoyed the programme or video they had watched.

Video or cable

The VIDEO variable shows those cases where the diarists had watched a film on video or cable. Diarists recorded watching less than six per cent of the programmes on video or cable. The extremely small number of cable viewings (50 or 0.08 %) reflects the low subscription rates for cable TV in the mid-1980s when the industry was in its infancy.

The variable IBACODE classifies the TV programmes according to the National Audience Research Codes (NARC). These codes are based upon information supplied by the Independent Broadcasting Authority (IBA). Approximately 80 per cent of TV diary events recorded by the diarists have a valid code for this variable.

Please see section 3.7 for a description and further notes on the variables in this file

2.5 Additional diary records for 15 cases

There are an extra set of different leisure diary records for the following 15 cases: *B11008Z, B11040Z, B11135F, B11146J, B11835C, B13863Q, B14973Z, B18711V, B21264U, B23389Q, B23519G, B24606J, B24976H, B25414G, B25574X*

It is possible that for these cases the diarist completed two diaries (in each of the two sets with the same BCSIDs the diarist's sex is the same). For these cases only the first recorded set is retained.

2.6 Television diary programme codes

In the television diary both MAINACT and IBACODE, due to being individually coded (mentioned above), can differ sometimes in the interpretation of the same television event. For example, there are 164 instances where an event coded as a film (5004) for MAINACT is not considered a film according to the IBACODE. Users are encouraged to utilise TVSTR2 when considering such discrepancies.

2.7 Leisure diary cases with limited information

The variable RISKDIARY is a flag in all three leisure diary datasets which can be used to identify cases with limited information, specifically 7 or fewer timeslots recorded per day or time-missing data for one or more days

3. The Leisure and Television Diaries research datasets

3.1 Licencing and Data access

All datasets are available from the UK Data Service (UKDS).

All users of the data need to be registered with the UKDS. Details of how to do this are available at https://www.ukdataservice.ac.uk/get-data/how-toaccess/registration.

The Leisure and Television diary data from the surveys have been supplied to the UKDS under End User Licence (EUL). These datasets can be downloaded once the End User Licence access conditions have been accepted by the user.

3.2 Available datasets

- bcs4_leisure_diary_episode: 483,999 rows

- bcs4_leisure_diary_calendar: 28,431 rows

- bcs4_leisure_diary_aggregate: 7,243 rows

- bcs4_television_diary_event: 59,786 rows

3.3 Identifiers

BCSID: *This variable uniquely identifies each diarist and it links the survey to the other waves of BCS70.*

EPISODENUM (Leisure diary episode file): *Identifier for each chronological recorded activity episode for a cohort member*

DAYNUM (Leisure diary Calendar file): *Identifier for each recorded day in a cohort member's diary*

EVENTNUM (Television diary event file): *Identifier for each chronological recorded event for a cohort member*

3.4 Leisure Diary Episode file: variable description and notes

Variable name	Variable description
EPISODENUM	Secondary identifier aggregate episode number across the four diary days
SEX86	Sex of diarist
DAY	Day diary kept
MONTH	Month diary kept
YEAR	Year diary kept
DIARYDAY	Day of week diary kept
EPNUM	Episode number within day
CLOCKSTH/ CLOCKSTM	Activity start time on 24-hour clock (hour/minute)
CLOCKENDH/ CLOCKENDM	Activity end time on 24-hour clock (hour/minute)
START/END	Start/ end minute of the activity (of 1440 min)
DURATION	Duration of activity in minutes
MAINACT	Original primary activity
SECONDACT	Original secondary activity
THIRDACT	Original tertiary activity
MAIN	Primary activity, 69 category list
SEC	Secondary activity, 69 category list
AV	Primary activity, 40 category list
WHERE	Location of the activity
WHOWITH	Who else was present during the activity
TOTEP	Total number of episodes in a diary
MISSINGTIME	Total number of minutes missing in this diary
RISKDIARY	Flag for diaries with limited information
STRT	Episode limit check - start
ENDD	Episode limit check - end

3.5 Leisure Diary Calendar file: variable description

Variable name	Variable description
DAYNUM	Secondary identifier aggregate diary day
SEX86	Sex of diarist
DAY	Day diary kept
MONTH	Month diary kept
YEAR	Year diary kept
DIARYDAY	Day of week diary kept
EPNUM	Episode number
TOTEP	Total number of episodes in a diary
MISSINGTIME	Total number of minutes missing in this diary
RISKDIARY	Flag for diaries with limited information
MAIN1-MAIN288	Primary activity
SECN1-SECN288	Secondary activity
THRD1-THRD288	Tertiary activity
PRI1-PRI288	Primary activity 2021 MTUS
SEC1-SEC288	Secondary activity 2021 MTUS
WHERE1-WHERE288	Location
COPR1-COPR288	Co-presence

3.6. Diary Day Aggregate file: variable description

Variable name	Variable description
SEX86	Sex of diarist
DAY	Day diary commenced
MONTH	Month diary commenced
YEAR	Year diary commenced
RISKDIARY	Flag for diaries with limited information
TOTAL_DAYS	Total number of days recorded
TOTAL_DURATION	Total recorded duration
TOTAL_EPISODES	Total recorded episodes
TOTAL_ACTIVITY_1	Total time spent: paid work

TOTAL_ACTIVITY_2	Total time spent: paid work at home
TOTAL_ACTIVITY_3	Total time spent: second job
TOTAL_ACTIVITY_4	Total time spent: school/classes
TOTAL_ACTIVITY_5	Total time spent: travel to/from work
TOTAL_ACTIVITY_6	Total time spent: cook, wash up
TOTAL_ACTIVITY_7	Total time spent: housework
TOTAL_ACTIVITY_8	Total time spent: odd jobs
TOTAL_ACTIVITY_9	Total time spent: gardening
TOTAL_ACTIVITY_10	Total time spent: shopping
TOTAL_ACTIVITY_11	Total time spent: childcare
TOTAL_ACTIVITY_12	Total time spent: domestic travel
TOTAL_ACTIVITY_13	Total time spent: dress/personal care
TOTAL_ACTIVITY_14	Total time spent: consume services
TOTAL_ACTIVITY_15	Total time spent: meals, snacks
TOTAL_ACTIVITY_16	Total time spent: sleep
TOTAL_ACTIVITY_17	Total time spent: free time
TOTAL_ACTIVITY_18	Total time spent: travel excursions
TOTAL_ACTIVITY_19	Total time spent: active sport
TOTAL_ACTIVITY_20	Total time spent: passive sport
TOTAL_ACTIVITY_21	Total time spent: walks
TOTAL_ACTIVITY_22	Total time spent: religious activity
TOTAL_ACTIVITY_23	Total time spent: civic activity
TOTAL_ACTIVITY_24	Total time spent: cinema, theatre
TOTAL_ACTIVITY_25	Total time spent: dances, parties
TOTAL_ACTIVITY_26	Total time spent: social club
TOTAL_ACTIVITY_27	Total time spent: pub
TOTAL_ACTIVITY_28	Total time spent: restaurant
TOTAL_ACTIVITY_29	Total time spent: visit friends
TOTAL_ACTIVITY_30	Total time spent: listen to radio
TOTAL_ACTIVITY_31	Total time spent: TV, video

TOTAL_ACTIVITY_32	Total time spent: listen to CDs records
TOTAL_ACTIVITY_33	Total time spent: study
TOTAL_ACTIVITY_34	Total time spent: read books
TOTAL_ACTIVITY_35	Total time spent: read paper/magazine
TOTAL_ACTIVITY_36	Total time spent: relax
TOTAL_ACTIVITY_37	Total time spent: conversation
TOTAL_ACTIVITY_38	Total time spent: entertain friends
TOTAL_ACTIVITY_39	Total time spent: knit, sew
TOTAL_ACTIVITY_40	Total time spent: other leisure
TOTAL_ACTIVITY_41	Total time spent: missing

3.7 Television Diary Event file: variable description

Variable name	Variable description
EVENTNUM	Secondary identifier television event number
SEX86	Sex of diarist
DIARYDAY	Episode day of the week
EORDER	Order of events within day
DD	Date diary commenced – day
MM	Date diary commenced – month
YY	Date diary commenced – year
START	Start time sequence
DURATION	Duration supplied by diarist – minutes
MAINACT	Main activity code
TVSTR2	Amended TV text
TVLIKE	How much programme was enjoyed
VIDEO	Video or cable (according to diarist)
IBACODE	Programme type: IBA codes

Appendix: Changes and corrections made to the Leisure Diary Episode file

This section describes the changes, corrections and imputation procedures made to the original episode data file. An analysis of the data with no attention to the information provided in this section can lead to biased estimates and inaccuracy in the results.

I. The cases where total minutes spent in the diary were less than 1440 fixed

There are 1440 minutes in each day, hence total amount of time in a diary (including the missing activities) should sum up to 1440. In some of the cases total amount of time in the diary was less than 1440 minutes whether due to the missing CLOCKEND variable in the last episode or START variable in the first episode. In a few other cases those variables had missing information in the middle of the day. Each of these cases were checked and corrected with reference to the previous or next episode.

II. All the cases given a start time of 0:00

Most of the diarists started their diaries by coding their very first activity after they got up. This resulted with inconsistent start times at the beginning of the diaries. We therefore added an extra episode to each diary which did not start at 0:00 and imputed the missing time between the last reported activity at 00:00 in the previous day and the first reported activity. When there was no information for imputation, we left this period as missing. When imputed we always used the "imputed X" activity code to distinguish these episodes from the actual reports of the diarists.
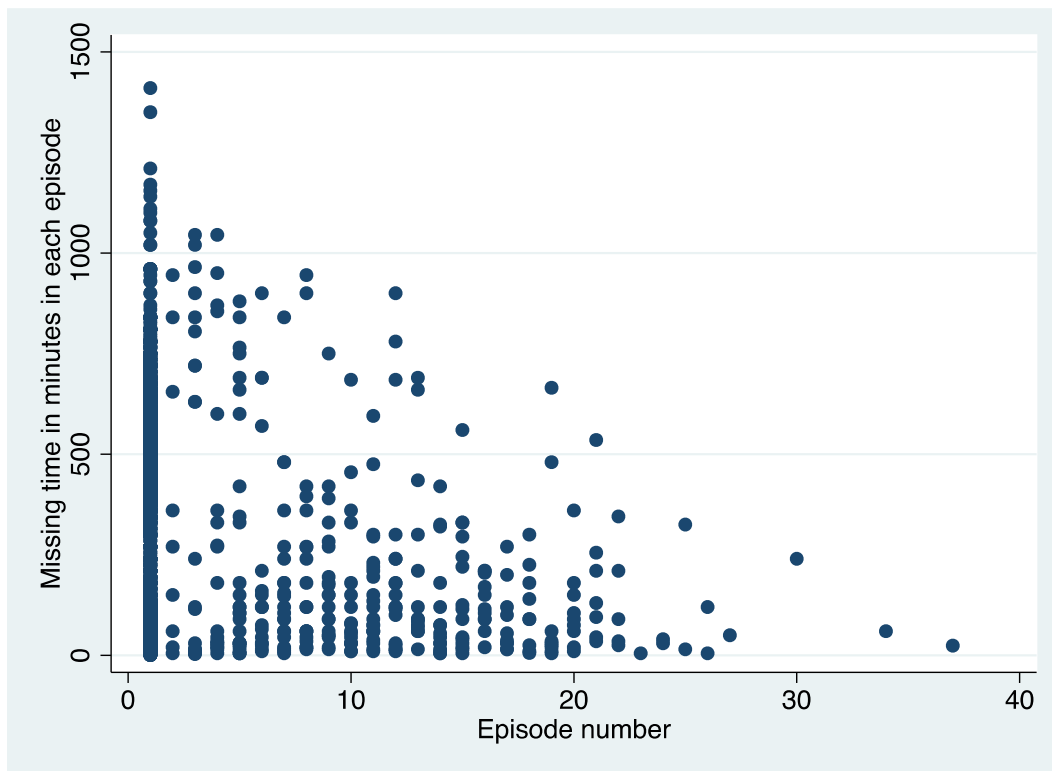
III. Activity imputation procedures

Figure 1: The distribution of missing time in each episode by episode number (positive missing time only)

Figure 1 shows the distribution of the missing time in each episode, by episode number. Episodes with 0 missing time are excluded. There are 9057 episodes where main activity is missing (out of 485,059) episodes, a very small proportion. Yet, the average duration of this missing time is quite considerable (300 minutes). This is because many respondents started to complete the diary at inconsistent times each day, leaving chunks of unreported time at the beginning of the day.

The longer the imputed period is, the more questionable it should be considered and users should be very careful while making inference using imputed activity categories. Depending on the research questions, some users might be better advised to follow more restricted procedures for the imputation, or consider the imputed periods as *missing*.

Imputed activity categories

1300 Imputed personal/ household care

1707 Imputed travel

1606 Imputed sleep

1809 Imputed time away home

3600 Imputed time at home (not necessarily personal/ household care)

Imputation procedures

If the diary has fewer than 5 episodes, no imputation is made and the missing period is coded as -7. This is because diaries with less than 5 episodes yield very little information for a reasonable imputation procedure.

There are around 850 4+ diaries (846 Tuesday diaries, 5 Wednesday and 6 Thursday diaries). No imputation was made for 4+ diaries in this version of the data. Please see item VI for more information on extra diaries on this point.

Note that in all these cases, the last reported activity of the previous day ends at midnight.

For Saturday, Sunday and Monday diaries, the following rules were used while imputing the first episode of the diaries.

i. If the last reported activity of the previous day is get up/ go to bed and the first reported episode of the following day is get up/ go to bed, the missing period in between is coded as *imputed sleep (1606);* <u>irrespective of the duration or timing of the period</u>. 8732 episodes are coded as imputed sleep according to this procedure. The durations of these episodes vary between one minute to 975 minutes, while 95 per cent is 11 hours or less. In 85 per cent of the cases (7476 episodes) the first reported activity (get up/ go to bed) following the imputation started between 6AM and 11:30AM.

ii. If the last reported activity of the previous day is main sleep and the first reported episode of the following day is get up/ go to bed the missing period in between is coded as *imputed sleep (1606);* <u>irrespective of the duration or timing of the episode.</u> 1643 episodes are coded as imputed sleep according to this procedure.

iii. If the last reported activity of the previous day is personal hygiene/ got ready and or other household activities at home and the first reported episode of the

following day is get up/ go to bed, then the following imputation procedures are followed:

      a. If the missing period is between midnight and 1AM, the missing hour is coded as *imputed personal/ household care (1300)*.
      b. If the first reported activity is reported between 1:01AM and 4:59AM, the period is coded as *imputed time at home (3600)*.
      c. If the first reported activity was reported at or after 5AM, the missing period is coded as *imputed sleep (1606)*.

iv. If the last reported activity of the previous day is work related, personal care, domestic work or leisure activity that took place at home and the first reported episode of the following day is get up/ go to bed which is also in the same home; then the following imputation procedures are followed:

      a. If the missing period is between midnight and 1AM, the missing hour is coded as *imputed personal/ household care (1300)*.
      b. If the first reported activity is reported between 1:01AM and 4:59AM, the period is coded as *imputed time at home (3600)*.
      c. If the first reported activity was reported at or after 5AM, the missing period is coded as *imputed sleep (1606)*.

v. If the last reported activity is arrive home/ go out, the location is at home and the previous location is different from home then we assume that the respondent arrived home (rather than going out). If the next reported activity of the following day is get up/ go to bed, the following imputation procedures are followed:

      a. If the missing period is between midnight and 1AM, the missing hour is coded as *imputed personal/ household care (1300)*.
      b. If the first reported activity is reported between 1:01AM and 4:59AM, the period is coded as *imputed time at home (3600)*.
      c. If the first reported activity was reported at or after 5AM, the missing period is coded as *imputed sleep (1606)*.

vi. If the last reported activity of the previous day is get up/ go to bed and the location is home, and the first reported activity of the previous day also takes place at home the time in between is *imputed sleep (1606)*.

vii. If the last reported activity of the previous day and the first reported activity of the following day is travel (indicated by activity codes or location codes) and the duration of the missing period is less than an hour, then the missing period is coded as *imputed travel (1707)*.

viii. If the last reported activity of the previous day is leisure outside home and the first reported activity of the following day (before 3AM) is travel, arrive home/ get out or leisure activity out of home; the three hour period in between is coded as *imputed time away home (1809)*.

ix. If the last reported activity of the previous day is activity reported at home and the first reported activity is also activity reported at home and the missing period ends at or before 10AM, it is coded as *imputed time at home (3600)*.

x. If the last reported activity is arrive home (confirmed by location codes) and the first reported activity is an activity at home and the location is no another home, the time in between is *imputed time at home (3600);* <u>irrespective of the duration or timing of the episode.</u>.

xi. If the last reported activity is an household activity and the first reported activity is activity on the following day is get up/ go to bed, the following imputation procedures are followed:

     a. If the missing period is between midnight and 1AM, the missing hour is coded as *imputed personal/ household care (1300)*.

     b. If the first reported activity is reported between 1:01AM and 4:59AM, the period is coded as *imputed time at home (3600)*.

     c. If the first reported activity took place between 5:00 AM & 12:00 PM, the missing period is coded as *imputed sleep (1606)*.

xii. If the last reported activity is baby sitting for other people's children and the location is not own home, and the next reported activity is arrive home/ get out or travel an hour or less of missing time is *imputed away from home (1809)*.

xiii. If the last reported activity is visiting friends and relatives and the first reported activity is either a form of travel or hanging out at someone else's house the missing period in between is coded as *time away from home (1809); irrespective of the duration or timing of the episode.*

For Friday diaries the following rules were used while imputing the first episode of the diaries.

i. If the first reported activity is get up/ go to bed and the activity is reported between 5:00 AM and 11:00AM the missing period in between is *imputed sleep (1606)*.

ii. If the first reported activity is get up/ go to bed and the activity is reported between 1:01 AM and 4:59 AM the missing period in between is *imputed time at home (3600)*.

iii. If the first reported activity is get up/ go to bed, personal care, eat/ drink and the missing period is one hour or less activity the missing period is *imputed personal/ household care (1300)*.

iv. If the first reported activity is leisure household activities such as watching TV and the missing period is an hour or less the missing period is *imputed time at home (3600)*.

v. If the first reported activity is between 5:00 AM and 9:30 AM and it's a "morning activity" such as personal care or eating/drinking, time in between is imputed sleep

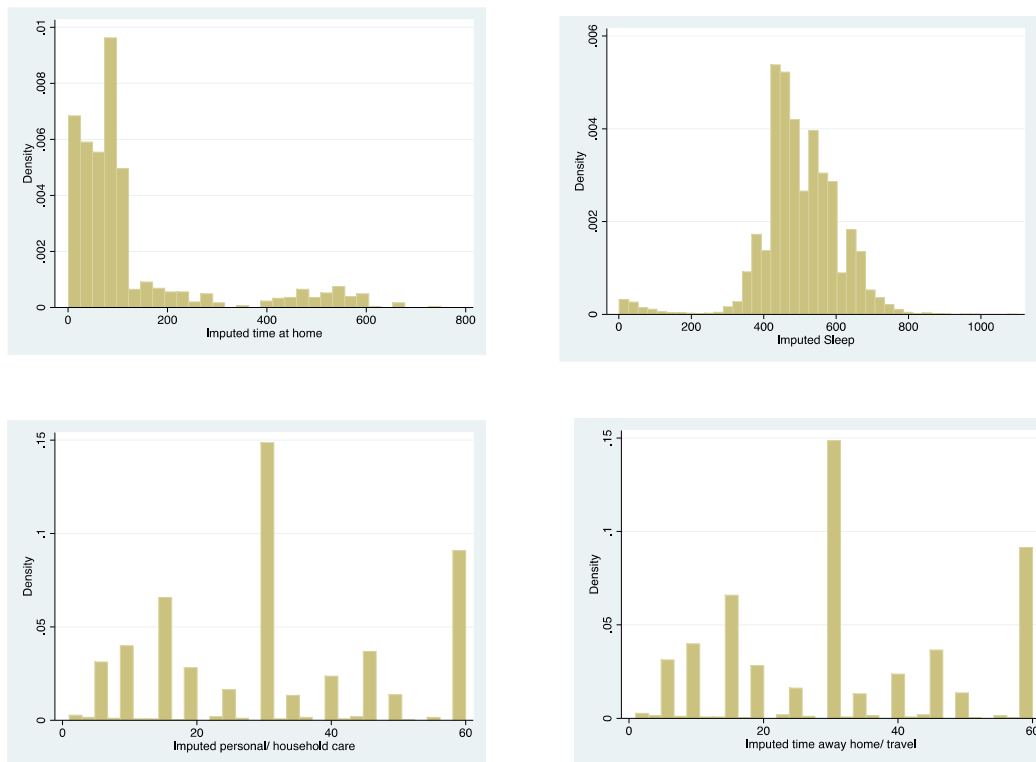Figure 2 shows the distribution of imputed activity durations.

Figure 2. Distribution of imputed episodes

At the beginning of the imputation there were 25 439 initial episodes with missing activity code. At the end of these procedures 2865 episodes still left missing and coded as 9999 (Missing, undecipherable). In total, 1% of the episodes are coded as missing. 38 per cent of episodes with missing codes are from diaries with four or less diaries.

Note that missing location and who-with variables are not imputed and left as missing.

IV. 1605 Sleep unspecified

This category is added to distinguish getting up/ going to bed from main sleep activities. In the original data there are four categories for time spent in bed and sleep related activities, namely:

1302 get up/ go to bed

1601 Main sleep

1602 Short naps and snoozes

1603 Being sick, ill in bed

The first of these categories is supposed to indicate a relatively short period of 'falling into sleep' or 'waking up' immediately before or after main sleep. It is therefore supposed to be an activity with relatively short duration. In 57 per cent of the episodes with 1302 get up/ go to bed code, the duration is half an hour or less, in 71 per cent it is an hour or less; and in 7 per cent of the cases it is 8 hours or longer. Clearly, such cases indicate a coding error on behalf of the coder. 1302 episodes that were an hour or more were coded as 1605 "Unspecified sleep" in order to distinguish a relatively long napping/ sleeping time from a getting up/ going to bed activity. Note that this activity category is different from 1302 "Imputed sleep". We do not use 1605 to fill a missing period. We this code only when the original code is 1302 but the period is longer than an hour.

V. Diary completion day

There are three separate variables – DAY, MONTH, YEAR - showing the diary commencement day, month and year respectively. These variables are left in this version of the data with no corrections or changes.

We created a new DIARYDAY variable to identify the day of the week the diaries are kept for. Note that in most cases DAY shows the diary start day, hence it is same for all diaries; whereas DIARYDAY shows the day of the week the diary is completed.

15 per cent of diary days could not be identified due to missing information on DAY, MONTH or YEAR.

VI. Extra diaries:

224 of Tuesday diaries have only one episode, while 461 have two and 106 have three episodes. 94 % of Tuesday diaries having 3 or less episodes indicate that most of these diaries are in fact continuation of Monday night, rather than full Tuesday diaries. These diaries can still be informative but in this version of the data we did not work on +4 diaries very much.

References

Dyer, N. (1984) *Daily life in the 1980s*, BBC Broadcasting Research Findings, No. 10

Gershuny, J.I., Miles, I., Jones, S.E., Mullings, C., Thomas, G. and Wyatt, S. (1986) "Time Budgets: Preliminary Analysis of a National Survey", **Quarterly Journal of Social Affairs**, 2(1), pp 13-39

Marles, V. (1985) *Listening and viewing amongst the unemployed*, BBC Broadcasting Research Findings, No. 11