



CHAPTER 1: OVERVIEW OF THE MTUS DATASET

1.1 Introduction

The origins of the Multinational Time Use Study (MTUS) go back to the 1980s following an initiative of Professor Jonathan Gershuny. The idea was to create a cross-nationally harmonised set of time use surveys composed of identically recoded variables. The first version of the MTUS dataset comprised some 20 countries, and has since been regularly expanded. The fifth version, named *World5*, comprised 35 surveys and was at the basis of the landmark study '*Changing Times*' by J. Gershuny (2000). The *World5* version of the dataset was restricted to the population aged 20 to 59 years old, even though the original versions of many of the surveys covered a wider range of respondents. The MTUS coordinating committee (<http://www.timeuse.org/mtus/coordinators/>) expanded to the whole age range of respondents from Release 2 of Version 5.52. This User's Guide describes the harmonised aggregate files for the Release 3 version of *World5.5* of nearly 40 surveys. Future releases of this User's Guide will gradually cover all the surveys in the MTUS archive.

The MTUS archive, located at the Centre for Time Use Research in the Department of Sociology at the University of Oxford includes:

- **Original files:** Episode and aggregate files, not generally available for distribution except by arrangement with the data provider. In some cases, CTUR only holds an interim version of the original data or only the MTUS version of the data.
- **World 5.53:** A harmonised aggregate file with all surveys from all countries except those surveys collected by the Australian Bureau of Statistics, Statistics Finland and Statistics Sweden, with diarists of all ages and with additional variables, but only with the old 41-digit activity codes. In this file, each row represents a 24-hour observation (diary).
- **World 5.8:** Two harmonised aggregate files (one for adult diarists aged 18+ and a separate file for diarists aged less than 18) with all surveys from all countries except those surveys collected by the Australian Bureau of Statistics, Statistics Finland and Statistics Sweden, with the same survey, household and diarist variables in *World 5.53*, but with summary time in both the old 41-category list and in a new and



updated 69 activity code list. Again, each row represents a 24-hour diary.

- **World 6.0:** Two harmonised episode files (one for adult diarists aged 18+ and a separate file for diarists aged less than 18) with all surveys from all countries except those surveys collected by the Australian Bureau of Statistics, Statistics Finland and Statistics Sweden, with the identifier variables, sex, age, and diary details: main activity, coded in both the 41 and 69 category frames; secondary activity in the 69 category frame, location, mode of transport whether ICTS were used, and who else was present. In these files, episodes represent a change in main activity, secondary activity, or location.
- **Restricted files** in versions 5.53, 5.8 and 6.0 for Australia, Finland and Sweden, are released separately and **require additional permission to access.**
- **Supplementary files** with additional variables for the UK (<http://www.timeuse.org/mtus/supplementary/>) and the USA (American Heritage Time Use Study – AHTUS <http://www.timeuse.org/ahtus/>).

Three variables, EDUCA, INCORIG and EMPINCLM, contain the original categories from each survey, and hence the labels are not included in the mega files. All the other variables are harmonised across surveys and not affected by this merge. We report the labels in the readme files for each individual survey (<http://www.timeuse.org/mtus/documentation/surveys/>).

We split the megafiles for the most recent versions into child and adult files for two reasons. First, the age of the youngest diarists vary highly across the surveys. Also, as children's time use differs markedly from the activities of adults, excluding children from the main file increases the harmonisation of the adult files (users can always either make use of the Version 5.53 or use the identifier variables to match the adult and child files together into the full surveys). Additionally, there are a number of surveys that only sampled young people. We hope in the future to include some of these surveys in the child files.

1.2 Format and structure of the datasets

The harmonised surveys are saved in **SPSS**, **STATA**, and **Flat text** formats. In future we also plan to make files available in SAS and R. People who download the flat text files can access the variable and value labels from the User Contributions page (<http://www.timeuse.org/mtus/contributions/>). Users are advised to either download the SPSS or STATA labels, change the extension form .sps or .do to .txt, then you can open the labels in a text editor.



In most descriptive analyses, MTUS Users are encouraged to use all cases and disregard the fact that the total number of cases (i.e. diaries) correspond to a smaller number of respondents for some surveys. However, when carrying out analyses based on inferential statistics, MTUS Users should be aware of the non-independence of cases and should use appropriate statistical techniques for those surveys which collected multiple diaries from the same respondent.

People with older versions of the MTUS data will note that in the past, we converted full-week surveys into a day representing average time in each activity across the week. From the release of Version 5.53 through all future releases, each individual diary day appears independently.

Also in contrast with previous versions of the MTUS, we no longer exclude diaries missing large amounts of time. Instead, we retain all diaries, but include a marker variable (badcase) which distinguishes quality from low-quality diaries. In addition, our recommended weight, propwt, assigns a zero-value to all low-quality diaries. Thus, when the data are weighted in analysis, low-quality diaries are excluded.

Some original surveys additionally include row cases for non-respondents who do not complete a diary. Nevertheless, most of the surveys do not include specific information on non-respondents in the data files. The MTUS format provides a suitable platform to analyse good quality diaries as well as low-quality diaries, but users would need to take greater account of original survey information to investigate people who do not respond at all. When original surveys include case rows for non-diarists (people with 24 hours of no reported activity), we delete the non-diary cases.

Further, from the release of Version 5.53, we have undertaken additional data preparation work before identifying low-quality diaries, as well as expanding the definition of a low-quality diary to consider factors in addition to missing time.

The time use diary is a narrative account and not a series of quantitative questionnaire answers. People can give a full account of their day without necessarily completing all columns of the diary for all potential time slots. For example, when offered a main activity and a mode of transport column, some diarists opt to make a time-saving minimal entry of writing “drove car” or “train” in the mode of transport column while not recording anything in the main activity column when their activity is transport. We consider entries across the whole diary to define the main activity, not just the information recorded in the main activity column.



We also impute activities in three specific instances. In cases where the main activity is missing at the beginning or end of the diary day, where the diarist records being at home or at another person's home, where the subsequent activities after the initial gap at the beginning or the preceding activities before the final gap at the end of the diary day are the sorts of activities which typically follow or precede sleep (have a drink, low-activity leisure like watch TV or read, dress/undress, personal care) we presume the missing activity to be sleep. (We make this assumption both on account of the pattern of activity and as time-diary surveys tend to start the observation window as a point of the day when most people in the population are asleep). Second, in circumstances of a missing period of 25 or fewer minutes that precedes leaving home to travel somewhere else or following returning home after travel from activities away from home, we assume the missing activity to include a combination of personal and household care. The documentation specific to each survey (conversion files) includes information on the number of cases amended by these procedures. Finally, where the diarist gives location information indicating that they are not at home but records no activity, we mark the case with the category unknown activity away from home on the newer 69 category activity code frame (though these cases are treated as missing time on the older 41 category activity list).

After this data preparation work, we define any diary to be low quality, when:

- continues to have 91 or more minutes of missing time,
- it has fewer than 7 episodes,
- it is missing two or more of four basic activities, or
- it was filled in by a diarist whose age or sex is not known.

We define the four basic activities as those necessary for basic functioning on a day-to-day basis:

- 1) eating or drinking (measured by time in these activities, or time recorded working with food (set or clear table, food preparation, cooking and the like), or the diarist being in a location where they are likely to be around food and drink, that is attending a feast or being at a pub or in a restaurant);
- 2) sleep or rest (including do nothing, think, time out, or take a work break)
- 3) personal care (including assumed self care preceding or following travel and receiving personal services, such as at hair dresser or doctor)
- 4) exercise and/or travel (including leisure excursions, gardening, walk dogs, imputed travel where no activity is recorded but the diarist records a change of location or records a mode of transport).



Diaries with large volumes of missing time do not account for enough of the day to allow imputation of what is likely to have taken place in the missing periods.

Very low episode diaries and diaries missing basic activities do not give complete accounts of the day, and both lead to over-estimates of the activities the diarist did record and under-estimates of the activities the diarist did not record. Age and sex are two variables highly associated with specific time use patterns, and these variables are required for the creation of the recommended weight.

We make **five exceptions in relation to these quality rules.**

- Diarists may not record any travel when their travel episodes are very short, but may record a pattern in the diary that lets us know that they did travel and where in the day the travel took place. In such diaries, you will find patterns of continuous reports of activity, and a change of location (such as eating breakfast at home then doing paid work at the office) with no report of travel in between the change of location. We handle these cases by making a flag variable for unreported travel present (0=no such missed travel, 1=missed travel). If the diary includes 2 of the 4 basic activities, one of the 2 missing activities is exercise or travel and the diary is flagged as including missing travel, then we count this diary as a good diary. We do not alter the diary record in such cases, and users of the MTUS would have to make their own adjustments to the entries made by the diarists if they wish to account for such travel.
- Diarists may not record any personal care when their episodes of personal care are short. In such diaries, you will find patterns of continuous reports of activity, and transitions where personal care is highly likely to have occurred (sleep for 2+ hours at home followed by other activities with no care, eating meals at home where the meal consumption lasts at least 10 minutes followed by other activities with no record of personal care). If such patterns are present, we make a flag variable for unreported personal care (0=no, 1=yes). If the diary is missing 2 basic activities, and one of these missing activities is personal care and we can flag this diary as having unreported personal care patterns, then we count this diary as a good diary. We do not alter the diary record in such cases, and users of the MTUS would have to make their own adjustments to the entries made by the diarists if they wish to account for such personal care.
- Diaries of carers (either the variable “carer” flagging cases of people who look after an adult needing assistance =1 for yes, or the diary includes any time in any form of adult or childcare) who otherwise meet the other 4 good diary criteria count as good diaries.



- Diaries including only 2 of the basic activities but that have at least 12 episodes where the diarist reports being at home all day (defined as no travel but eloc=1 – own home, or eloc=2 – other’s home for at least 1000 minutes), but otherwise meet the other 4 good diary criteria count as good diaries.
- Other diaries including only 2 of the basic activities and 15 or more episodes count as good diaries.

Some original surveys additionally include row cases for non-respondents who do not complete a diary. Nevertheless, most of the surveys do not include specific information on non-respondents in the data files. The MTUS format provides a suitable platform to analyse good-quality diaries as well as low-quality diaries, but users would need to take greater account of original survey information to investigate people who do not respond at all. When original surveys include case rows for non-diarists (people with 24 hours of no reported activity), we delete the non-diary cases.

1.3 Surveys included

The following table lists all the surveys included in MTUS, as well as the surveys that we hope to include in coming years.

Table 1.1 – List of surveys included and versions available

Country	Surveys and years	Versions available		
Austria	1992	5.53	5.8	6.0
	2008-09	Hope to include in future		
Australia	1974	5.53	5.8	6.0
	1987	5.52		
	1992	5.52		
	1997	5.52		
	2006	5.53		
Belgium	1965	5.0		
	1999	Hope to include in future		
Bulgaria	1965	Hope to include in future		
	1988	5.0		
	2001-02	Hope to include in future		
Canada	1971	5.52		
	1981	5.52		
	1986	5.52		



	1992	5.52		
	1998	5.52		
	2005	Hope to include in future		
Denmark	1964	5.53		
	1975	Hope to include in future		
	1987	5.53		
	2001	5.53		
Estonia	1999-2000	Hope to include in future		
Finland	1979	5.52		
	1987	5.52		
	1999	5.52		
France	1966	5.0		
	1974-75	5.0		
	1986	Hope to include in future		
	1998	5.53	5.8	6.0
Germany	1965	W5.0 only for now		
	1991-92	5.53	5.8	
	2001-02	5.53		
Hungary	1965	5.0		
	1976-77	5.0		
	1986-87	Hope to include in future		
	1999-2000	Hope to include in future		
India	1998-99	Hope to include in future		
Ireland	2007-08	Hope to include in future		
Israel	1990	5.0		
Italy	1979-80	5.0		
	1989	5.53		
	2002-03	5.53		
Japan	1976	Hope to include in future		
	1981	Hope to include in future		
	1986	Hope to include in future		
	1991	Hope to include in future		
	1996	Hope to include in future		
	2001	Hope to include in future		
	2006	Hope to include in future		
Latvia	2003	Hope to include in future		
Lithuania	2003	Hope to include in future		
Netherlands	1975	5.52		
	1980	5.52		
	1985	5.52		
	1990	5.52		



	1994	5.52		
	2000	5.53	5.8	6.0
	2005	5.53	5.8	6.0
New Zealand	1998-99	Hope to include in future		
	2008-09	Hope to include in future		
Norway	1971	5.52		
	1981	5.52		
	1990	5.52		
	2000	5.52		
Peru	1966	Hope to include in future		
Poland	1965	5.0		
	1975-76	Hope to include in future		
	1984	Hope to include in future		
	2001	Hope to include in future		
Portugal	1999	Hope to include in future		
Romania	2001	Hope to include in future		
Slovak Republic/ Czechoslovakia	1965	5.0		
	2006	Hope to include in future		
Slovenia / Yugoslavia	1965	5.0		
	2000	5.52		
South Africa	2000	5.52		
Spain	2002-03	5.53	5.8	6.0
Sweden	1991	5.52		
	2000	5.52		
Turkey	2006	Hope to include in future		
United Kingdom	1961	5.52		
	1974-75	5.53	5.8	6.0
	1983-84	5.52		
	1987	5.52		
	1995	5.53	5.8	6.0
	2000-01	5.53	5.8	6.0
	2005	5.53	5.8	6.0
USA	1965-66	5.52		
	1975-76	5.52		
	1985	5.53	5.8	6.0
	1992-94	5.53	5.8	6.0
	1994-95	5.53	5.8	6.0
	1998-2001	5.52		
	2003-04	5.53	5.8	6.0
Totals: 22 countries	60 surveys	5.53: 20	5.8: 15	6.0: 13



1.4 Technical information on the surveys

The table below contains key information on the sample size, age of respondents, response rate, etc. for each of the surveys included in the MTUS. These tables summarise the information in the top table of the readme files.

Table 1.2 Technical information on the time use surveys

Country ¹	Year	Age	Sample Size ²	Survey Period (# months) ³	Response rate (%)	Diary (# days)	Type of diary	Time interval	Household members ⁴
AUS	1974	18+	1491	7	67% A/W 58% Melb	1	On day	Free	No
	1987	15+	1011	2	74.2%	2	On day	15min	Yes
	1992	15+	7045	11	82.9%	2	On day	5 min	Yes
	1997	15+	7246	8	72.0%	2	On day	5 min	Yes
	2006	15+	13617	8	82.5%	2	On day	5 min	Yes
OST	1992	10+	25233	2	47.0%	1	On day	30 min	Yes
CAN	1971	18-64	2141	8	72.0%	1	On day	Free	Yes
	1981	15+	2686	3	46.0%	1	On day	Free	No
	1986	15+	9946	3	78.9%	1	On day	Free	No
	1992	15+	9815	12	77.0%	1	Recall	Free	No
	1998	15+	10749	12	77.6%	1	Recall	Free	No
DEN	1964	15+	4069	2	80.4%	1	Recall	30/15 min	In limited cases
	1987	16-74	3584	3	72.7%	1	Recall	15 min	No
	2001	16-74	4000	4	70.0%	2	On day	10 min	In limited cases
FIN	1979	10-64	12038	4	81.0%	2	On day	10 min	No
	1987	15+	7758	12	74.0%	2	On day	10 min	No
	1999	10+	10561	12	52.0%	2	On day	10 min	Yes
FRA	1985/6	15+	16047	12	66.9%	1	On day	5 min	Yes
	1998/9	15+	15441	12	88.3%	1	On day	10 min	Yes
GER	1991/2	12+	7200	4	Quota	2	On day	5 min	Yes
	2001/2	10+	11919	12	95.5%	3	On day	10 min	Yes
ITA	1988/9	3+	38110	12	70.0%	1	On day	Free	Yes
	2002/3	3+	55773	12	91.8%	1	On day	10 min	Yes
NET	1975	12+	1309	1	76.0%	7	On day	15 min	No
	1980	12+	2730	1	54.0%	7	On day	15 min	No
	1985	12+	3263	1	54.0%	7	On day	15 min	No
	1990	12+	3158	1	49.0%	7	On day	15 min	No
	1995	12+	3227	1	20.0%	7	On day	15 min	No
	2000	11+	1813	1	25.0%	7	On day	15 min	No
	2005	12+	2204	1	37.0%	7	On day	15 min	No
NOR	1971/2	16-74	3040	12	58.0%	2 & 3	On day	15 min	No
	1980/1	16-74	3307	12	65.0%	2	On day	15 min	No
	1990/1	16-79	3097	12	64.0%	2	On day	15 min	No
	2000/1	9+	3211	12	50.0%	2	On day	10 min	Yes
SPA	2002/3	10+	46774	12	86.0%	1	On day	10 min	Yes



Country ¹	Year	Age	Sample Size ²	Survey Period (# months) ³	Response rate (%)	Diary (# days)	Type of diary	Time interval	Household members ⁴
SLO	2000/1	10+	4500	12	52.5%	2	On day	10 min	Yes
RSA	2000	10+	14553	3	94.0%	1	Recall	10 min	Yes
SWE	1990/1	20-64	3943	9	75.0%	2	On day	10 min	No
	2000/01	20-99	3976	12	50%	2	On day	10 min	No
UK	1961	15+	2363	1	69.8%	7	On day	30 min	Yes
	1974/5	5+	3583	4	60.0%	7	On day	30 min	Yes
	1983/4	14+	1525	2	51.0%	7	On day	15 min	Yes
	1987	16+	3035	1	70.0%	7	On day	15 min	Yes
	1995	16+	1875	1	93.0%	1	Recall	15 min	No
	2000/1	8+	11667	12	45.0%	2	On day	10 min	Yes
	2005	16+	4941	10	59%	1	Recall	10 min	No
USA	1965	18-64	1243	7	74.0%	1	On day	Free	Yes
	1975/76	18+	2406	3	72.0%	1	On day	Free	No
	1985	12+	5358	12	55.2%	1	On day + recall	Free	Yes
	1992/4	0+	9386	12	63.0%	1	Recall	Free	No
	1994/5	18+	1199	13	64.6%	1	Recall	Free	No
	1998/1	18+	1700	12	56.0%	1	Recall	Free	No
	2003	15+	20720	12	57.0%	1	Recall	Free	No
	2004	15+	13973	12	57.3%	1	Recall	Free	No

Notes:

- 1- More countries have carried out time use surveys. A more comprehensive list is available at the MTUS web site: <http://www.timeuse.org/information/studies/>
- 2- Unless otherwise indicated, the sample size refers to the number of individuals. The actual number of cases is larger in surveys where 2 or 3 diaries were collected.
- 3- 'Period' refers to different collection periods throughout the year.
- 4- Indicates whether or not more than 1 household member was included in the survey.

1.5 File naming conventions

We have standardised the way files are named in MTUS. The name of each file distinguishes:

- The country (2 or 3-letter code) (see table below);
- The year of the survey (4-digit);
- The version of the archive (World5.3, 5.8, 6.0);
- The type of file (extensions 'sav' or 'dta' for data files, and extensions 'sps' or 'do' for programme files).

For example, Release 3 of the World 5.5 version of Canada 1992 is called 'Can1992W553.sav', which should be read as:

Country: Can

Year: 1992



Version: W553
Type: sav (the SPSS version)

Note that in surveys for which the data collection spread over more than 1 year, SURVEY takes the value of the year when the data collection began. For example, Finland 1987/88 is referred to as FIN1987. For exceptions to this rule, see the survey-specific README documents.

Country	Code	Country	Code
Argentina	ARG	Latvia	LAT
Australia	AUS	Lithuania	LTH
Austria	OST	Netherlands	NET
Belgium	BEL	New Zealand	NWZ
Brazil	BRA	Norway	NOR
Bulgaria	BUL	Peru	PER
Canada	CAN	Poland	POL
China	CHI	Portugal	POR
Denmark	DEN	Republic of Korea	RKR
Estonia	EST	Romania	ROM
Finland	FIN	Russia	RUS
France	FRA	Slovak Republic/Czechoslovakia	SPA
Germany	GER	Slovenia/Yugoslavia	SLO
Hungary	HUN	South Africa	RSA
India	IND	Spain	SPA
Ireland	IRE	Sweden	SWE
Israel	ISR	Turkey	TUR
Italy	ITA	United Kingdom	UK
Japan	JPN	United States of America	USA

1.6 Missing value conventions

We use three codes to mark missing values, and a separate fourth convention for weights and identifier variables that are not present, as follows:

- “-7” refers to situations for we can create a variable for this survey, but we cannot create the variable for this diarist (or diary) as the respondent was not asked for the information on this diary or because the information is not relevant to that respondent (such as the employment status of a spouse for a person who is single and not living with a co-habiting partner). Although this missing value option potentially applies to all variables, it is mainly used for



AGEKIDX, AGEKID2, WORKHRS, EMPSP, PARNTID1, PARNTID2, PARTID and EMPINCLM.

- “-8” refers to situations where we can create the harmonised variable for the study, but no information is recorded for this case (item non-response or insufficient information to create the variable for that case).
- “-9” refers to situations for which the harmonised variable could not be computed for the survey (with exceptions for weights and case identifier variables – although the identifier of spouse or of parents can have a -8 value if this could not be created for a case). Note that we use -9 with the time use activity variables to distinguish true 0s (the diarist did not record any time in this activity, though in theory they could have done so) from cases where no time is recorded in the activity because we could not create this time use category for this survey.

There are cases where an original weight is not present. In these cases, we use “0” rather than a missing value to indicate that this weight is not present in the study (and anyone attempting to use this weight would find they have no cases remaining for analysis from the survey). The conventions for the identifiers are set out in detail below.

Users also should note that we do not use missing values for the aggregated or summary time use variables, unless the category is not available for the whole survey. A value of 0 means that the diarist did not record any minutes in the activity (it is impossible to say for certain if this is because the diarist did not do any of the activity or if the diarist actually did undertake the activity but did not report doing the activity in the diary). If a category is not coded in the survey, then the summary value is set to -9 for the whole survey. Users should take notice of -9 values. If one sums time across a variable that cannot be created for a survey without first addressing the missing categories, 9 minutes will be subtracted in error for each category that is not present.

There are no system missing cases in MTUS data files. All cases for all variables have either a valid value or a standardised missing value. The MTUS data files contain no declared missing values. MTUS users need to declare missing values if they choose to do so before running their analysis.

1.7 Data preparation

The World 5.53 and World 5.8 Versions of the harmonised MTUS datasets are restricted to the **aggregate (summary)** version of the time-use surveys. The MTUS includes survey, demographic and socioeconomic information about



respondents (hereafter called diarists) and their households alongside the aggregated time-use variables. Each row case in the W5.53 and W5.8 files represents one 24-hour time diary. For those studies where respondents completed more than one diary, individual diarists appear on a separate row for each diary they completed.

The World 6.0 files include the sequence information, though this version will only be created for a limited subset of datasets as the process of conversion is more cumbersome, especially in the case of older time use surveys where information is reported for uneven intervals. In the World 6.0 episode files, each row case represents one change of main activity, secondary activity or location in a diary. Where the diarist completed more than one diary, the episodes of the subsequent diary or diaries follow the episodes of the first diary. As this file is large, only the identifiers, age and sex are included in the World 6.0 file alongside the diary information. Users will need to match the World 6.0 files with the corresponding World 5.8 file for the additional background variables.

Before beginning the actual conversion, users should undertake three steps to ensure maximum data quality.

DATA PREPARATION STEP 1 – Check alternative options for the MTUS background variables to ensure that you are using the option with the cleanest profile compared with other reported results and the least missing data. If there are options and one is better than others, the choice should be documented in the conversion syntax. In some cases, combinations of original variables are needed to create the MTUS variables.

DATA PREPARATION STEP 2 – Apparently missing main activity time in diaries is not necessarily missing. The point of the diary is to collect information about what people are doing at any point in time. Diarists sometimes do not fill in the main activity column – creating the appearance of missing data, but fill in other information elsewhere in the diary that nonetheless indicates their activity and allows us to properly code the time period. We should recognise that elements of the diary are not always separate. At points of overlap, diarists can record a comprehensible and clear response in the diary in one place but not in others. For example, an entry “too train to work” is both a location/mode of transport and an activity, and this dual meaning is clear even if the entry is written only once in either the main activity column or the location column. The redundancy of writing the same entry in two places is not necessary for the diary to have a full account of a participant’s activities. We recommend the following steps be undertaken where a main activity is missing before converting the data:



- a) Completing a time diary can be an onerous task, and some diarists do not appreciate making redundant entries. Where diaries have a location or mode of transport column and the diarist is travelling, some diarists may write “drive my car to work” or “on the bus” in the mode of transport column and see no point in writing the same entry in the activity column. When main activity is missing but the diarist has recorded a mode of transport during this time period, we recode the missing main activity slot as unspecified travel.
- b) Some diarists get confused while they complete the diary in a hurry, and may record the main activity in the secondary activity column. Another possibility is that a diarist may be undertaking a series of main activities while also doing an extended secondary activity – for instance alternating between care of pets, care of children and housework as main activities while listening to the radio. An item on the radio may be particularly interesting and attract the diarist’s full attention for 10 minutes, but the 10 minutes of main activity radio listening is more efficiently recorded by simply extending the radio listening recorded in the secondary activity column. Where main activity is missing but a valid secondary activity is recorded, we recode the main activity as the reported secondary activity, and recode the secondary activity as no reported second activity. These are cases where the diarist has reported one valid activity.
- c) For short gaps in the early hours at the beginning or end of the diary where the diarist is at home or in the same location where they report sleeping on the diary day and asleep before the gap or asleep following the gap, we recode the gap as imputed sleep.
- d) If a short gap (<25 minutes) occurs at home just before travel or at home just after travel home, set the missing time to imputed personal and household care.
- e) If there is other diary information that illuminates the activity in an episode where there is no recorded main activity, this should be used to identify the activity. As some examples, individual surveys in the past have recorded information as the number of cigarettes smoked during the episode, which television station the diarist watched if they watched TV during the episode, which type of material the diarist read if they read during the episode, and the like. Similar to the instance of the diarist recording a mode of transport but not recording a main activity, a diarist might record that they smoked 10 cigarettes in 15 minutes or watched a specific television station for 45 minutes but not record a main activity.



Nonetheless, such records do reveal what the diarist was doing, so can be used to complete apparently missing time episodes.

All these changes are making use of information the diarist supplied about their activity, and this procedure eliminates some unnecessary wastage of diaries. All such data cleaning should be fully documented in the conversion programme.

DATA PREPARATION STEP 3 – Check to see if other diary information facilitates the coding of time use activities. Different studies code activities in different ways. Sometimes researchers need to use multiple columns from the diary to code a single activity. For instance, some surveys simply code “eating/drinking” in the main activity, and the location variable is needed to distinguish breaks at work (AV1) from eating meals at home (AV15) from drinking in a pub (AV27) from eating out in a restaurant (AV28) from eating a meal at a friend’s house (AV29). Likewise, location can distinguish socialising with friends at home (AV38) from socialising with friends away from home (AV29) or paid work (main job) at home (AV2) from paid work (main job) away from home (AV1). Who else is present information sometimes is needed to distinguish childcare from adult care. Some cases arise peculiar to only one dataset. For instance, Denmark 1964 includes an original code for all media use, but also has a separate column where diarists indicated what media they were reading, watching, or listening to, and this second column enables the separate coding of listening to the radio (AV30) from watching TV (AV31) to listening to music (AV32). All combinations of information used to code a category should be included in the documentation.

We engage in a series of quality checks in the post-harmonisation process. These quality checks are detailed in Section 11 (Quality checks) of the MTUS coding procedures document available elsewhere on the MTUS web pages (<http://www.timeuse.org/mtus/convert/>).



1.8 Identifying good-quality diaries

Diaries with large volumes of missing time do not account for enough of the day to allow imputation of what is likely to have taken place in the missing periods. Very low episode diaries and diaries missing basic activities do not give complete accounts of the day, and both lead to over-estimates of the activities the diarist does record and under-estimates of the activities the diarist did not record. We make one exception in relation to these quality rules. Some people who care for a young child or an adult on their diary day appear not to record their own rest, consumption or personal care. We count diaries of people who performed any care (of a child, pet or adult) on the diary day who otherwise reported 7 or more episodes and have accounted for at least 22 hours and 30 minutes as quality even if two or more basic activities are missing. Age and sex are two variables highly associated with specific time use patterns, and these variables are required for the creation of the recommended weight, as is the day of the week on which the diary was completed. We also classify diaries missing age or sex of the diarist or the day of the week on which the diary was completed as low quality diaries.

For this reason, the MTUS includes the variable BADCASE to distinguish quality diaries from diaries lacking sufficient standards for analysis. The MTUS team defines any diary to be low quality when:

- continues to have 91 or more minutes of missing time;
- it has fewer than 7 episodes;
- it is missing two or more of four basic activities, or
- it was filled in by a diarist whose age or sex is not known;
- the day of the week on which the diary was completed is not know.

We define the four basic activities as those necessary for basic functioning on a day-to-day basis:

1. eating or drinking (measured by time in these activities, or time recorded working with food (set or clear table, food preparation, cooking and the like), or the diarist being in a location where they are likely to be around food and drink, that is attending a feast or being at a pub or in a restaurant);
2. sleep or rest (including do nothing, think, time out, or take a work break);
3. personal care (including assumed self care preceding or following travel and receiving personal services, such as at hair dresser or doctor);
4. exercise and/or travel (include leisure excursions, gardening, walk dogs).



Note that only quality diaries have positive values in the MTUS weight PROPWT. Low quality diaries should have "0" values for this weight. Also, as previously noted in this document, we keep only diary cases. Cases of non-diarists (effectively row cases for non-respondents) are removed.