

Imputation for missing values and disaggregation of information

We did not over-write information recorded in the original surveys. In the case of the background and demographic information, we have included a number of flag variables that mark cases of inconsistent information. We did not mark information that appears inconsistent in the diaries as such identification is less straightforward. We discuss the reasons for this decision in the inconsistencies section which follows this.

We did impute diary information only for cases where we could make logical inferences to fill in gaps. We also made some imputations to disaggregate information using material provided in the diaries. The full detail of the imputation and disaggregation is documented in the syntax files used to construct the harmonized datasets.

Imputed time use activities have separate codes, and thus are easily distinguished from the originally coded activities. We have included six imputed codes for:

- **sleep and rest** (main activity only). We define imputed sleep as early hours time at home or another home where the activities before or after the block include personal care activities associated with getting up or preparing for bed and where little other sleep is recorded in the diary.
- **imputed eating** (secondary activity only). Where diaries included at least 15 episodes yet include no episodes of eating or drinking, but the diarist does report food preparation and or setting then clearing the table, we assume that the person is likely to have consumed some food or drink during these activities, and if no secondary activity was reported, then we coded secondary eating time.
- **imputed personal or household care** (main activity only). We defined imputed personal care in gaps following main night sleep where the next activities lead up to leaving the house, and also activities where a person has been away from home, then has the gap immediately upon arriving home, before undertaking other activities. Typically, these gaps are short duration activities. In 1975-76, those imputed personal care gaps preceded by travel tend to be followed either by sleep or travel.
- **imputed social activity** (main activity only for 1965-66, 1992-94, and 2003; some cases of secondary activity for 1975-76 and 1985). We defined imputed social time either from original codes where the diarist reported to activity of the person with whom they interacted but not their own activity and time where no activity was recorded but other people were recorded as present. In the case of the 1975-76 and 1985 data, there is a code for recording the activity of someone else who was present rather than the diarist's own activity. These cases are recoded as imputed social activity when they appear as either a main or a secondary activity.
- **imputed time away from home** (main activity only). We define imputed time away from home as time when the person has been traveling, following time at an unspecified other (not missing) location away from home or another home
- **imputed travel** (main and secondary activity). We defined imputed travel when the person changed location from their own home or another home to another specified location (not to an unknown location or to a general "other" location), or changed from one type of specified (not general "other") location to another specified type of location but did not indicate an episode of travel. We did not impute travel when the

change of location was from an away from home location to a restaurant, bar or café, as a number of diaries use such a change of location to mark walking into the restaurant or food court next to shops in a shopping mall or the canteen at a workplace.

We also disaggregated some activities. The older heritage datasets distinguish physical and general child care provided only to a child aged under 5 from physical and general care provided to a child aged 5 to 17 or to children of mixed age groups. The 2003 ATUS only includes a physical and general child care code, but the who else is present file allowed us to determine if the child or children present during the vast majority of such activity were aged less than 5 or older.

In the 1975-76 study, spouses completed diaries which only collected main activity and location, but not secondary activity or who else is present. For all datasets, we imputed whether domestic animals and whether shop or professional service staff are present from the activities. We also imputed the presence of children from some of the child care activities. For the spouse diaries in 1975-76, we also imputed the presence of the spouse by using the main respondent's diary – if the main respondent reported being with her or his spouse, then we imputed the spouse being present during the corresponding activity on the spouse diary as well.

The 1965-66 data combined animal care with gardening. We examined gardening and animal care in the later surveys, and noticed a number of patterns. Very few gardening episodes last for less than 15 minutes. A large number of animal care episodes are a short duration and fall into particular series of circumstances:

- taking place in the kitchen or dining room in the house around the time the diarist eats or is engaged in food preparation or washing up dishes;
- taking place in the house when the person first wakes up in the morning or returns from a long spell away from the house;
- taking place in the middle of the night, when the person wakes up, then goes back to bed shortly thereafter;
- travel episodes involving walking starting and ending at home (these episodes tended to be longer – a median of 20 minutes, compared to the other animal care episodes, which are generally less than 15 minutes, and sometimes often 5 or fewer minutes).

These cases are coded as animal care in the 1965-66, while the remaining cases in the combined gardening/animal care category recorded as gardening.

For all years, we used the location information to distinguish paid work at home from paid work elsewhere. For 1992-94, we had to code cycling and walking as secondary activities where the main activity is traveling and the mode of transport if cycling or walking. For 2003, we identified eating at the workplace and eating in restaurants by using the code for eating and drinking and the location code. We also used location (at home versus at other people's home or other places) to distinguish receiving or visiting friends from social communication at home.

There is one instance where we did not disaggregate. This case relates to eating at work for the 1992-94 NHAPS data. In that survey, only a general eating category is coded. While this dataset does include a variable with many of the verbatim responses for some activities, 98.1% of the

entries for the “eating/drinking” code are listed as “eating meals or snacks”. This survey includes the most detailed list of locations, but these location codes do not identify whether the location is a workplace, with the exception of one code – “at work, no specific location”, which is applied to a mere 15 episodes. Instead, this survey allows people to determine if an episode of paid work took place in a laundry mat, library, office, factory, restaurant, etc. The problem this presents is that it is not possible to say if an episode of paid work took place in a library or restaurant as this is *the location where the person normally works* or because this is *the location of a meeting* on that particular day. We did investigate the level of eating at work episodes we could impute if we only counted cases where the activity before and after the eating was main paid work, but this gives a result very much below the level of eating at work in the earlier surveys. Very little of this activity is recorded in 2003. We decided not to disaggregate this activity, though users are free to make their own estimates if this identification is important to their research.

John Robinson has noted that future potential work using the original verbatim responses for the 1985, 1992-94 (as well as 1995 and 1998-2001) datasets may allow future disaggregation, such as distinguishing watching one’s child act in a school play as opposed to watching other amateur theatre or professional theatre, or to distinguish watching one’s child play sports at school as opposed to being a spectator at other sporting events.