

Essential notes for using the 1975-76 data

This survey collected information from the same people over four waves. The background variables are longitudinal, and can be treated as such appropriately. Whether four diary days truly represent longitudinal change at the individual level is a different question. Time-diary studies collect at two levels of sample: personal data (and in some cases household or other aggregated unit data), and, separately, activity patterns over 24-hour days. Time-use surveys reveal what patterns of activity occur in the general population and which patterns are associated with which groups of people. Time use surveys do not reveal the range of patterns in which any particular individual engages – and to answer questions about changes in an individual’s general behaviour, a researcher would need substantially more detailed information from that person, collected through such qualitative methodologies as life history interviewing or long-term diary collection. Study of person-level behaviour change entails a considerably greater researcher input and respondent commitment than can generally be feasibly expected with national-sample surveys.

This means two things, First, it does not matter if an individual is asked to complete the diary on a usual or unusual day as the random element of the sample should capture something close to the true proportion of ordinary and non-typical days from the population engaged in those behaviour patterns on the sampled days. Second, in cases of longitudinal time use surveys, such as this 1975-76 data set, the data can be used to see what patterns of behaviour are associated with people who do and who do not experience different changes over time, but the data are not suited to say whether individual “x” has changed usual behaviour. Some features of the way the diarists completed the diaries may show longitudinal (or previous participation in a diary study) effects, but a case can be made for treating each diary as a separate observation.

There are two additional good reasons for using four waves of the 1975-76 survey. Small numbers of diaries can produce peculiar results, as an unusual pattern may cause undo effect. Using all four waves increases the sample size of diary days. Second, the four waves of data collection span a whole year. The 1965-66 survey collected data over six months, and the more recent surveys collected data over a whole year, which allows for inclusion of seasonal activity variation in models. If you restrict your analysis to only the first wave, in addition to having a small sample size, you also have only autumn activity patterns – whereas with all four waves you have four seasons of activity patterns. In addition, roughly half the diary days were collected on weekend days, and the other half on week days. The proportion of days collected during different waves varies. Researchers may wish to take this into account when analyzing the data, though the weights will balance the distribution of the days of the week.

In some limited circumstances where a researcher has a main aim to calculate a participation rate, the fact that up to four diaries were completed by a single person may be an issue for concern. Indeed some previous research has revealed that comparability issues can arise in relation to participation rates when researchers compare surveys which collected only one diary per respondent with surveys that collected one week of diaries or two to three diaries per participant in the same week. The reason that such comparability issues arise is that people engage in weekly and monthly as well as daily cycles of behaviour. If the activity of interest is an activity in which people engage virtually every day (sleep, eat, personal care, etc.), then the number of diaries has little impact. For activities where many of the people who engage in the activity do so on a

regular but not a daily basis (for instance voluntary activity or exercise), participation rates can vary significantly when the survey collects more than one diary over a short period. In the case of this 1975-76 survey, though, most diaries are collected at intervals of two or more months, and thus are not likely to produce the comparability issues that arise when comparing one-diary-per-participant surveys with surveys collecting two to seven diaries in the same week. In most circumstances, using these diaries as separate observations will produce reasonable results. If the researcher has a concern about this issue, we recommend that they consider one of two strategies:

- Select for the good diary sample ($recwght$ or $xtimewt = 0$). The attrition weights account for the return of a good diary in each wave. Then randomly sample one diary per person. This will yield a sample with diaries from the full year rather than simply from the first two months of data collection.
- Use clustered errors in models, clustering around respondent's identification number (pid). This option will address the problem of errors differing with up to four diaries coming from the same person rather than from four different people.

Users also should note the differences between the AHTUS versions of this file and the versions of this dataset used up until this point in most publications. Most previous research has used files of summed time in activities for only the first wave of data collection and for spouses as well as main respondents. The AHTUS version of this data file includes significant data cleaning, all the context variables, and all four waves of data with attrition weights for main respondents only. The numbers from this dataset will differ from previously published numbers. The closest equivalent to the previous files (though with errors corrected) is in the supplementary file for this survey which covers both the spouse and main respondent diaries (if the user selects for the first wave only).