

CHAPTER 5: WEIGHTS IN MTUS

Time diary analysis requires two levels of weighting. First, as in all surveys, weights are needed to bring the sample in line with the population from which it was drawn. Second, weights also balance seasonal variations and variations by day of the week.

The MTUS aggregate file contains the following two weights:

- OCOMBWT: Original weight (population & day preferred, or whatever original weight is available if not combined);
- PROPWT: Proposed weight (population & day combined weight rescaled if needed) (see below about the rescaling procedure).

The core file contains only propwt.

The episode file does not contain weights for two reasons. First, the weights are calculated at the diary level. Non-response and sample distribution adjustments can be addressed at the level of the diary and the diarist. The meaning of an episode level weight is less clear. There is no way of knowing how many diaries are less detailed than the diarists' actual days or the degree to which activities are under-reported at the episode level. While the sample distribution can be reasonably accurately estimated, it is not possible to estimate the true distributions of episodes in a meaningful way. Second, the meaning of an episode varies by the context of the research purpose (discussed more in the next chapter). As there is not a set unit for all analysis at the episode level, the episode file is best treated as a database from which variables are extracted to match back into the aggregate version for analysis.

Some surveys inflate the sample size by a factor to mirror the size of the whole population of the country. If the original weight is inflated, OCOMBWT should be left inflated.

Nevertheless, to promote consistency among the datasets and to prevent surveys from countries with larger populations from apparently swamping surveys from countries with smaller populations, we deflate the original weight in the computation of PROPWT. The mean of the original weight will sum to the inflation factor. Where survey designs collect diaries on a weekday and a weekend day, it is advisable to use the mean of the weekday diaries to deflate weekday diaries and the mean of the original weight for the weekend diaries to deflate the weekend diaries.

If the survey does not include a weight, OCOMBWT should be set to 0. Researchers have to find official statistics describing the population by age and sex. We recommend the United Nations publication “*World Population Prospects*” which contains time-series (since 1950) of the population by age and sex for each country. Alternative internationally recognised sources, such as the ILO Yearbook also may be used. If the survey has enough cases for you to split age and sex groups by employment status (using emp, employed or not employed) – this means if you have at least 50 cases of working and not-working for each sex and age group – then also include employment status. Note that you may not include employment status for the youngest and oldest diarists if few are working, but include emp for the working age population.

The main MTUS weight, PROPWT, is calculated by:

- 1- Begin from the original survey weight (deflated if the original weight was inflated). If there is no original weight, construct a population weight by dividing the percentage of the population you would expect to be in each age/sex (and in some cases employment status) group by the percentage of cases you actually have in that age/sex (/employment status) category in the sample data (expected/achieved).
- 2- Create a good-diary inflation factor by dividing the total number of diaries collected in the survey by the number of good diaries (total diaries/good diaries).
- 3- Create an interim weight with the same value of the weight in step 1. Set the value of this interim weight to 0 where NOWGHT=0. Multiply your interim weight by the good diary inflation factor you created in step 2 for all cases (remaining 0 where nowght=0).
- 4- Compute ASEWT - a sum of the weights for each age/sex(/employment status) group (across all days of the week).
- 5- Compute two further sums: ASEDayWT - a sum of weights for each age/sex (/employment status) group separately for each day of the week, and also: ASEDayCASE - the sum of the number of cases in each age/sex(/employment status) group for each day of the week. (This step differs from step 4 as now you calculate the sums separately for each day of the week).
- 6- Compute an expected sum of weights (ESW) for each day of the week: $ESW = ASEWT / 7$ (divide by 7 as there are 7 days of the week).
- 7- $PROPWT = (ESW / ASEDayWT) / (ASEDayCASE / ASEDayWT)$
- 8- Check that the weight has calculated correctly. PROPWT should have a mean of 1, and PROPWT should sum to the total number of diary cases. When the weight has been applied, the weighted frequency of the diaries by the days of the week should be evenly distributed (14.3% for each day). Similarly, each age/sex group should have an even distribution of diaries across each day of the week.