
Subject: Re: Household weighted prevalence estimates
Posted by [Bridgette-DHS](#) on Wed, 10 May 2023 12:17:07 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from Senior DHS staff member, Tom Pullum:

Your intuition is correct. Rescaling the weights with any multiplier will have the same effect on the numerator and on the denominator, Means, proportions, percentages, rates, and ratios are unaffected by re-scaling the weights. This applies to subpopulations too. Means, etc., defined for geographic subdivisions or for wealth quintiles, or any other covariates are unaffected by rescaling. You can just use the weights as they are.

If you were to pool surveys and use the weights in the separate surveys, the overall means, etc., would be affected by the relative sizes of the samples. You would then want to re-scale in proportion to the population sizes to get unbiased estimates of the pooled population. But we recommend against such pooling because the surveys are conducted at different times and aggregations are not samples from a well-defined population. Also, the results would be dominated by the largest country. Glad you are not planning to produce pooled estimates.

The observation that "weighted counts are not population estimates" is based on the DHS scaling of the weights, such that the mean of hv005 in the HR file is 1 (multiplied by 1,000,000 to get an integer). Similarly, the mean of v005 in the IR file is 1 and the mean of mv005 in the MR file is 1. If, say, you used the HR file, with hv005 as a frequency weight, to get a count of the number of households that have electricity, the total would be (approximately) the number of households in the sample that have electricity, times 1,000,000. That would NOT be interpretable as an estimate of the number of households with electricity in the country.
