
Subject: weighting urban rural divide , multiple countries years

Posted by [pedrorosado](#) on Mon, 09 Oct 2023 13:15:54 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hello all,

I'm currently working with DHS dataset, creating a pooled dataset of all the surveys conducted in Africa. The objective is to create a comparable wealth index, similar to the International Wealth Index. Over the course of this work several questions arouse.

Firstly:

- Imagining the case of a single country survey analysis, if I want to create statistics comparing the Urban to rural population, how could that be achieved, in terms of weighting? From what I understand the sampling method takes into account this divide, and samples accordingly, so, by which weight should populations be weighted?

Secondly:

- In the Index construction, some questions for which I'm interested in are either not present for specific surveys/years, or some households do not have a recorded answer (9 or 99 or 999). From what I read the observations are not self-weighted, so my questions is: Is there a common/accepted way of weighting for this type of surveys? And if so, which? Additionally, does this affect the cluster/stratum weights (as in, if a household did not answer a question, in a certain cluster, and it gets a smaller weight for that cluster, does that also reduce the cluster weight in the sample?).

Thank you for your time,
Pedro

Subject: Re: weighting urban rural divide , multiple countries years

Posted by [Bridgette-DHS](#) on Mon, 09 Oct 2023 14:32:57 GMT

[View Forum Message](#) <> [Reply to Message](#)

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

There have been many posts on weights. They are also discussed in the Guide to DHS Statistics. When analyzing a single survey, you should use the weight option in any estimation commands to get unbiased estimates. The weight is hv005 in the HR and PR files, v005 in the IR, KR, and BR files, and mv005 in the MR and CR files. The weights take into account the difference between the urban/rural composition in the sample and in the population. The only purpose of the weights is to produce unbiased estimates.

We mostly use Stata, and you will find several examples of weighting with Stata on the forum. We have GitHub code in SPSS and R, as well as Stata, with many examples of weighting and "svy" adjustments for clusters and strata. The construction of recodes and indices does not include weights.

<https://github.com/DHSProgram>
