
Subject: Re: A little question about pooling data

Posted by [Reduced-For\(u\)m](#) on Sun, 15 Apr 2018 20:14:40 GMT

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

There are a number of ways to do that. One would be what you describe - take a weighted average of the (weighted) country specific averages. Your second weighting (the average of the country-level estimates) could be weighted in any way you want. One difficulty with that is getting good confidence-interval estimates... which is why people sometimes pool the data together to do the estimation in one step (and because that simple method won't work for more complicated estimates).

The second option is to simply append all the data together and re-scale the DHS weights in such a way that you effectively get back out the weighting that you do want. One way would be to normalize each individual survey/country to have their weights sum to 1* and then apply those weights in a regression context (or using the svy: prefix in front of a Stata command). This would be effectively weighting each country equally.

Or, you could take those "sum-to-1" weights from the previous step, and multiply them by some population of interest to get a "representative population weighted average", which is more like the thing you describe (the thing in my first paragraph). The difficulty here is getting the appropriate population size for the appropriate population... that has to come from outside the DHS.

If you need more details on these, many of these problems have discussed on the boards under the name of "de-normalizing" weights... if you don't find what you need there, feel free to ask for more specifics about one of the methods.

*To get within-country weights to sum up to 1, you just get the total sum of weights for each survey round, and divide the DHS-given weights with the sum-of-weights (the Stata "egen, by" command is good for this).