
Subject: Weighting the KR for multiple countries
Posted by [DHSStudent](#) on Tue, 23 Apr 2013 20:49:22 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi,

I'm interested in looking at the DHS children's recode by pooling multiple countries (one survey from each, all Phase 6), but I'm sort of confused whether or not to weight the data, and how to go about it if so. I know v005 can be used for a single country, but I've mostly been finding answers related to single countries or the Household Recode so far so I'm unsure.

I'm planning on simply concatenating the 5 datasets.

I'm using SAS, if that matters. Suggestions on coding would also be great, but I'm mostly just confused as to how to go about weighting.

Any guidance would be greatly appreciated. Thanks so much.

Subject: Re: Weighting the KR for multiple countries
Posted by [Reduced-For\(u\)m](#) on Tue, 23 Apr 2013 21:15:22 GMT
[View Forum Message](#) <> [Reply to Message](#)

I am literally working on this right now, but it is a pretty tough question.

So you append 5 surveys together. What do you want the thing you are estimating to represent? An average of all the people that live in those five countries? An average that thinks of a country as an individual? An average that represents the people sampled in the DHS?

My guess is that intuitively you want the first one. So here is how I've been thinking of it, but I would be very happy if someone pointed out why it was wrong. It's based off the attached note (and modified from it) which was originally posted here:

http://userforum.measuredhs.com/index.php?t=tree&th=54&goto=82&#msg_82

gen preweight = (v005/1000000)

Now go get population estimates of each country, merge them as "countrypop"

egen countryn = count(Y), by(country)
*for whatever Y is, your outcome

*your weights, I think, should sum to countryn ... that is something mentioned on this board somewhere, but you can check how close that is once you lose some observations due to missing covariates or whatever. You may need another level where you scale the weights to sum to the number of observations.

gen newweight = preweight*(countrypop/countryn)

and now you can set "newweight" as your weight. This should correspond to a "population average of your five countries".

Note, though, the relative size of your countries may basically dominate any across-country differences. I'd do all the regressions/tabulations separately by country too, just so you know.

File Attachments

1) [Note+on+de-normalization+of+DHS+standard+weight.pdf](#),
downloaded 1399 times

Subject: Re: Weighting the KR for multiple countries
Posted by [bsayer](#) on Wed, 03 Jul 2013 22:31:29 GMT
[View Forum Message](#) <> [Reply to Message](#)

I think you are right on the money regarding "What does the pooled data represent". That is a key question. And are the countries sufficiently similar in the characteristic that you are studying to warrant pooling? If the answer to the second question is yes, then I don't think you need to do anything to the weights, provided you are using survey design capable software. In the context of SAS, this is somewhat iffy, as I don't trust the results.

BTW, in SAS you just use one set statement and list the five data sets. Just make sure the variable names and lengths are the same in all five data sets.
