
Subject: Weighting Data for Pooled BDHS Household Members Dataset
Posted by [suzzon](#) on Mon, 03 Apr 2017 17:35:24 GMT

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

Dear DHS experts,

I am using the Bangladesh DHS 1993/94, 1996/97, 1999/2000, 2004, 2007, 2011 and 2014 Household members files. Using DHS forum guidance, I have de-normalized the weights (hv005). Then I have appended all the BDHS Household members files. I am using comparative wealth index (CWI) proposed by Dr. Shea Rutstein and Sarah Staveteig as socioeconomic status variable. I have the following specific questions:

1. Should I use the de-normalized hv005 (denormw) as sample weight in further statistical analysis?
2. Or Should I compute a new household members weight variable (hhmemwt) from the de-normalized hv005 and use it as sample weight in further statistical analysis? The procedure of this Household members weighting (hhmemwt) for CWI has been mentioned in CWI computation procedures (SPSS syntax).

I am providing the Stata code for calculation of these two weight variables (i.e. denormw and hhmemwt) here:

**** Denormalization of household members weight provied in household member files

```
gen denormw=(hv005/1000000)*(total number of residential households in the country/ total number of households interviewed in the survey)
```

*** HH members weighting for Index

```
gen hhmemwt=denormw*hv012
```

Thank you for your time.

Moin

Subject: Re: Weighting Data for Pooled BDHS Household Members Dataset
Posted by [Bridgette-DHS](#) on Thu, 06 Apr 2017 11:53:54 GMT

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

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

There have been many forum posts on the topic of the weights. My recommendations continue to be the following. First, when using the pooled file you should only look at trends or differences between surveys. The pooling of several surveys into a single large file is only for data processing convenience. You should not do analyses that combine the surveys in the sense of calculating a single mean for two surveys, because such a number would not refer to a well-defined population and is not interpretable. Second, for looking at trends or differences between surveys, I recommend adjusting the weights so that the total weight for each survey is the same. I have described how to do this in Stata. Your alternative, to weight in proportion to population size (in terms of number of households) is also possible, especially when working with multiple surveys from a single country. I would not recommend using your hmemwt. Within DHS, we would never apply a multiplier of hv012.

Subject: Re: Weighting Data for Pooled BDHS Household Members Dataset
Posted by [habt_lancs](#) on Wed, 13 Dec 2017 17:25:20 GMT

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

hello,

I am using the household members recode by pooling the surveys conducted in different periods. And in this case i have to de-normalize the weight by using the following procedure:

$$HV005^* = HV005 \times (\text{total number of residential households in the country at the time of the survey}) / (\text{total number of households interviewed in the survey})$$

Unlike, the population of female 15 -49 data that can be obtained from UN, i do not know how to get data on "total number of residential households in the country at the time of the survey".

any help please...

Best,

Subject: Re: Weighting Data for Pooled BDHS Household Members Dataset
Posted by [Bridgette-DHS](#) on Thu, 14 Dec 2017 14:44:48 GMT

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

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

The household members file, or PR file, has one record per person, You could multiply the

weights in that file (hv005) by the total population of the country (preferably the total population residing in household), divided by the total sample size (which is the same whether weighted or unweighted). The problem you describe would only be relevant if you were using the HR file, which has one record per household.

There are many components of uncertainty when you do this, regarding the restriction to the population in households, the choice of reference date, the fact that some population estimates will be inter-censal estimates and others will be post-censal projections, the inaccuracies in the census counts, the possibility of using information about age/sex and place of residence within the country, etc. I personally would not alter the weights in this way. There have been many postings on this topic.

Subject: Re: Weighting Data for Pooled BDHS Household Members Dataset
Posted by [habt_lancs](#) on Thu, 14 Dec 2017 14:55:43 GMT

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

Thanks so much. Very helpful.

One point though

You said : "I personally would not alter the weights in this way."

I also read your suggestions in a different thread; you seem to be unconvinced of this way of de-normalizing (re-weighting).

Since i am using many surveys of a country (Ghana) pooling together, what do you think is the best way to solve this problem.

De-normalizing this way or just simply using the pooled original weight (either hv005 or v005) by dividing it by 1000000?

To elaborate on what i am doing, if it helps, i am using a diff-in-diff estimation strategy where my interest variable is measured at region level and i cluster SE at region level.

best,