
Subject: Replicating Step 7 and Level-2 Weight from Report 27

Posted by [Yawo](#) on Sat, 08 Jul 2023 02:48:43 GMT

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

Hello,

I am replicating the calculation of level weights, so I can clone them for other countries. I realized that the codes provided are almost identical for calculating the weights for HH and for women in Zimbabwe 2015. For example, step 7 advises us to use the following formula to generate the sample weights:

```
gen DHSwt = v005 / 1000000
```

However, on pg 10 of the report, there are directions (bullet 6 & 7) to get projected population figures for women for the survey year and divide those by number of women interviewed. But the code provided for Zimbabwe did not follow this approach. Out of curiosity, I went ahead and retrieved the pertinent population figures and calculated a new weight as follows:

```
gen newwt =  
replace newwt =(v005/1000000)*(4098000/9955)  
gen DHSwt_2 = newwt/1000000
```

Of course, the figures are different, as I expected.

Question 1: Why is it the case that the procedures for approximating household and individual weights are identical for the Zimbabwe illustration and that the population figures were not used? More specifically, under what conditions would you use the projected population figures or not?

Question 2: Secondly, I fitted the test run (below) using the new weights:

```
svyset v001, strata(v022) weight(wt2) singleunit(centered) || _n, weight(wt1)  
svy: melogit v474 i.sex ||v001:, or
```

I received the much-discussed error: "weights in variable wt2 not constant within groups defined by: v001" Following the advice of Tom reported in this post:

<https://userforum.dhsprogram.com/index.php?t=msg&th=9098 &start=0&>, I generated the geometric mean version of wt2, as Tom advised below, and avoided the dreaded error.

Given that wt1 is revised from wt2? (ie: `gen wt1 = d_HH/wt2`), should we adjust wt1 as well with the revised wt2 (i.e., `gen rev_wt1 = d_HH/rev_wt2`)?

Question 3 : But I have a follow-up question about the code below:

```
bysort v001: gen dif = 0  
replace dif = 1 if v001 == v001[_n-1] & wt2 != wt2[_n-1]  
browse if dif == 1  
count if dif==1
```

I can see the code is trying to check for an error or consistency, and replacing missing values. But could you please clarify what it is exactly doing, and what answer one should be looking for? I get very different results of the count variable depending on whether I use the original wt2 or the revised version:

wt2 count if dif==1 reports 7,652 cases
rev_wt2 count if dif2 == 1 reports 0 cases

Thanks in advance... Yy
