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

Subject: Re: Replicating Step 7 and Level-2 Weight from Report 27
Posted by Bridgette-DHS on Wed, 12 Jul 2023 11:02:30 GMT
View Forum Message <> Reply to Message

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

You appear to have divided the weight by 1000000 twice. The first gen command did follow the instructions for denormalizing the survey weight and already restored the survey weight's scale. Your second command, "gen DHSwt_2=newwt/1000000" is not needed.

Subject: Re: Replicating Step 7 and Level-2 Weight from Report 27
Posted by Yawo on Wed, 12 Jul 2023 21:20:55 GMT
View Forum Message <> Reply to Message

Thanks for your response. I understand the procedure, so just to be clear, I have two related questions:

Q1: Given that I am combining the data for both men and women, are the following codes ok now?

**************** STAGE A ********************
* 7: Prepare the adjusted final/individual adjusted weight for men and women: dividing population of women or men by the appropriate number interviewed; de-normalize the final weight */

gen newwt =.
replace newwt =(v005/1000000)*(4764000/10060) if sex==1
replace newwt =(v005/1000000)*(5538000/5135) if sex==2
label variable newwt "Population adjusted sample weight"

************ STAGE B: APPROXIMATE LEVEL-1 AND LEVEL-2 WEIGHTS ************
/*Step 1: De-normalize the final weight, using approximated normalization factor. */
gen d_HH = newwt * (M/m_c)

/* Step 2: Approximate the level-2 weight (get f variation factor; set alpha based equal split

```
gen f = d_HH / ((A_h/a_c_h) * (M_h/S_h))
scalar alpha = 0.5
gen wt2 = (A_h/a_c_h)*(f^alpha)
```

/* Step 2B - OPTIONAL: Generate a revised level-2 weight by replacing it with its geometric mean [if needed to get resolve the "weights in variable wt2 not constant within groups defined by: v001" error - based on advice by Tom here in [message #21803]:https://userforum.dhsprogram.com/index.php?t=msg&amp ;amp ;amp ;amp;th=9098&goto=21803&#msg_21803

```
gen ln_wt2=log(wt2)
egen mean_wt2=mean(ln_wt2), by(v001)
gen rev_wt2=exp(mean_wt2)
drop ln_wt2
```

/* Step 3: Approximate the level-1 weight */ gen wt1 = d_HH/wt2

Q2: If rev_wt2 is used, then should we adjust the formula for wt1 as well: gen rev_wt1 = d_HH/rev_wt? By the same token, should we also change to the rev_wt2 in the loop (i.e, gen wt1_`i' = d_HH/wt2_`i')?

With much appreciation for your assistance.

Cy

Subject: Re: Replicating Step 7 and Level-2 Weight from Report 27
Posted by Bridgette-DHS on Mon, 17 Jul 2023 11:36:14 GMT
View Forum Message <> Reply to Message

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

In the MR file the weight is mv005, so I think you want mv005 rather than v005 in the line for sex==1. But maybe you have already renamed mv005....

The steps look ok. As I believe I said earlier, I would be cautious about pooling the MR and IR files. Note that the age range for men is usually not the same as the age range for women.

Subject: Re: Replicating Step 7 and Level-2 Weight from Report 27
Posted by Yawo on Tue, 18 Jul 2023 12:07:56 GMT

Thanks very much - yes, I have am aware of your cautions about pooling male and female data together and also noted that women have different ages (15-49) vs men (15-54). I also renamed the male variables from mv* to v* before the merge.

But there is the second part of my post and question - relating to the use of the revised wt2 - which was not addressed. if we used a revised/ geometric mean version of wt2 (rev_wt2), as alluded to earlier, does it mean that we have to:

(1) use the revised wt2 in generating wt1 (i.e, d_HH/rev-wt2 instead of d_HH/wt2)?
(2) similarly, should we also change to the rev_wt2 in the loop (i.e., gen wt1_`i' = d_HH/wt2_`i')?

Thanks, again Tom and Bridgette - for all your assistance.

best, Yawo

Subject: Re: Replicating Step 7 and Level-2 Weight from Report 27
Posted by Bridgette-DHS on Wed, 19 Jul 2023 16:21:12 GMT

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

My answer is "yes" to (1) and (2). Good luck with this!

Subject: Re: Replicating Step 7 and Level-2 Weight from Report 27
Posted by Yawo on Wed, 19 Jul 2023 18:08:22 GMT

thanks so much. This is very helpful.

best, Yawo