Subject: Stratification for 1996, 1999, 2004-2005 dataset
Posted by mariachiara on Mon, 18 Sep 2017 08:58:28 GMT
View Forum Message <> Reply to Message

Dear DHS Forum Users,

My research project will try to estimate the impact of foreign land acquistions on children malnutrition in Tanzania. As such, I am using only children datasets and I pooled together the children datasets for 1996, 1999, 2004-2005, 2010 and 2015-2016.

Reading several threads, I understood that I have to:

1) de-normalize the sample weights as I am using several rounds of survey. This can be done by making sure that the weights for each round sum up to one and then multiplying those weight by the eligible population, i.e. kids under five (which I found here https://esa.un.org/unpd/wpp/Download/Standard/Population/). Is there someone who can confirm this procedure?

2) setting my dataset as survey data using the "svyset" command in Stata. Here is the point where I am stuck.
I have read other threads where they indicate the code for correctly using "svyset" for different surveys round, but my problem is that I can't identify the strata variables for 1996, 1999 and 2004-2005 datasets.
For the 2010 and 2015-2016 datasets, the strata variable is v022 and the stratification is well explained in the final report. In those datasets stratification is done at the regional and urban/rural levels, so for each region there are two strata (urban and rural). So, in 2010 there are 26 regions in Tanzania and V022 lists 52 strata, while in 2015 there are 30 regions and v022 lists 59 strata (one region is considered totally urban here).

In 2004-2005, problems start. Even though the sampling frame is the same as the 2010 survey (the 2002 census), the variable v022 lists 221 strata so I don't think that this variable is correctly identifying strata. However, I read in detail the final report and I found no information on the stratification. The same is true for the 1996 and 1999 surveys, where variable v022 lists 177 and 84 strata, respectively. However, the final report for the 1996 rounds says "The list of PSUs for the 1996 TDHS survey was stratified by each of the 20 regions (for the mainland) and within each region by urban and rural areas".

Therefore, I am not sure how to identify strata for those three surveys and I would greatly appreciate any help that you can give me! I apologize if the questions are silly but I have never worked with survey data and I am trying to understand how DHS works.

Thanks a lot.

Best,

Mariachiara

Subject: Re: Stratification for 1996, 1999, 2004-2005 dataset
Posted by Liz-DHS on Mon, 18 Sep 2017 18:29:29 GMT
View Forum Message <> Reply to Message

Dear User, A response from Senior sampling expert, Dr. Ruilin Ren,
Quote:
1. Demoralize the household weight HV005 or the individual weight V005 in DHS does not need the population size of the indicator to be calculated. For de-normalizing HV005, it is always multiplying the weight by the total number of households of the country at the time of the survey and dividing by the total number of households interviewed in the survey. For de-normalizing V005, it is always multiplying the weight by the total number of women 15-49 of the country at the time of the survey and dividing by the total number women 15-49 interviewed in the survey. The population size of the indicators to be calculated using household weight or individual weight have nothing to do with the de-normalization.
2. For surveys older than 2010, the stratification variable HV022 or V022 might not be coded correctly in the DHS data, it is safe to use HV024 crossing HV025 or V024 crossing V025 to get the sampling stratum, because in most of the DHS surveys, stratification is the crossing of region by urban rural.

Subject: Re: Stratification for 1996, 1999, 2004-2005 dataset
Posted by Reduced-For(u)m on Mon, 18 Sep 2017 18:48:44 GMT
View Forum Message <> Reply to Message

I like this - we should probably just stick with "Demoralizing weights" as our phrase of choice. It seems to capture most of our feelings about trying to get weighting right.