Subject: Weights for analysis of HIV+/HIV- urban and rural populations
Posted by Bouboulina on Wed, 17 May 2017 19:14:00 GMT
View Forum Message <> Reply to Message

Hello,

I read the DHS guide to statistics and need help interpreting this note: "The sum of the sample weights only equals the number of cases for the entire sample and not for subgroups such as urban and rural areas."

I want to compare proportions of various variables between urban and rural HIV-positive populations (stratified by sex) in the 2010 Burundi and 2015 Rwanda surveys. Since my analysis will exclude HIV-negative persons, is it still correct to use the HIV05/1000000 as the weight? Or do I need to do something additional to this weight?

I'm also interested in comparing urban HIV-positive to urban HIV-negative populations and rural HIV-positive to rural HIV-negative populations (also stratified by sex). Same question here: is it still correct to use HIV05/1000000 as the weight?

Thanks so much!

Subject: Re: Weights for analysis of HIV+/HIV- urban and rural populations
Posted by amage on Sat, 16 Sep 2017 19:17:32 GMT
View Forum Message <> Reply to Message

Hello,

Following up on this question, which I posted under a different user name in February (I had to re-register).

I'd be so grateful for help with this matter. To give a little more context, I'm trying to compare a cohort of HIV-positive patients from clinics in Rwanda, Burundi, and the DRC to the population-representative samples captured by the DHS to determine whether the cohort is representative of the HIV-positive populations in these countries in terms of several variables, including sex, age, BMI, marital status, pregnancy status, etc.

Could you let me know if the HIV05 variable is the correct weight variable to use for this purpose? Please see my previous message in this thread as well.

Thanks so much,
Anna

Subject: Re: Weights for analysis of HIV+/HIV- urban and rural populations
Posted by Bridgette-DHS on Wed, 20 Sep 2017 12:09:14 GMT
View Forum Message <> Reply to Message

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

I assume you are using Stata with pweights and svyset. The weight can be hiv05 or hiv05/1000000. You will get exactly the same results with either one.

If you want to restrict an analysis to a sub population, then you can do it by including, at the end of the estimation command, something like "if hiv03==0". However, it would be better to construct a variable that is 1 for this subpopulation, e.g. "gen hivneg=1 if hiv03==0" and then put "svy, subpop(hivneg): " in front of (on the same line as) the estimation command. There are good reasons for using the subpop option with svy, even though (in some checks I have done) it produces almost exactly the same results as the "if" command.

Perhaps another alternative would be to include hiv03 as a covariate on the right hand side of the model. If its coefficient is not significant, then you don't need hiv03 in the model. You have to be sure to take account of any values of hiv03 other than 0 and 1. Note that if hiv03 is NA (that is, ".") then hiv05 will be too.