
Subject: Re: declaring child survey in Stata

Posted by [Trevor-DHS](#) on Thu, 05 Mar 2015 20:48:50 GMT

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

- 1a) Why are you pooling the data? What benefit do you get from pooling the data over running two separate analyses? Are you trying to compare changes over time?
- 1b) You cannot use the PSU and stratum variables as they are. You need to create new PSU and stratum variables that are specific to each survey, e.g. `egen newpsu = group(survey_year v001)` and `egen newstrata = group(survey_year v023)`
- 1c) You probably have to denormalize the weights for your analysis. There are several posts on the forum about denormalization.
- 1d) your `svyset` command should look more like `svyset newpsu [pweight=wgt], stratum(newstrata)` where `wgt` is the denormalized weight. Even if you weren't denormalizing, you would need to first divide `v005` by 1000000, but as you are pooling data you should be denormalizing the weights.
- 2) If you don't use `svyset` and `svy: regress` (or similar commands) then your tests of significance will be incorrect.
- 3) For the weight, see how it is used above in the `svyset` command. Whenever you use an `svy:` command Stata will refer to the weight in the `svyset` command.
- 4) Using the weight when not using the `svy` commands:
`gen wgt=v005/1000000`
`tab var1 var2 [iw=wgt]`
Using the weight with the `svy` commands:
`gen wgt=v005/1000000`
`svyset v001 [pweight=wgt], stratum(v023)`
`svy: tab var1 var2`
[this assumes the strata are in `v023` - sometimes they are in `v022` and sometimes they need to be created. See other posts concerning the strata variables to use.]
- 5) The variables you mention are standard variables, but I would never copy and paste data as you suggest. You should use the commands for combining datasets:
<http://www.stata.com/manuals13/u22.pdf>
-