Subject: Re: declaring child survey in Stata
Posted by Trevor-DHS on Tue, 17 Mar 2015 01:27:26 GMT
View Forum Message <> Reply to Message

I'll respond to the 3 posts in order. First, though, are you using Stata or SPSS? The first code examples were in Stata, but the data files you sent were in SPSS.

First message

1a)
I don't understand how the 2003 data are not affected if the policy change took place in 1997.

You don't need to use year variables - you just need a variable that differentiates between the two surveys, so you can use your created variable v0 (phase). From this you can easily create your dummy variables.
gen d2003 = (v0==3)
gen d2008 = (v0==4)

1b) I don't understand your comment about VOO1 and V001. I only see v001 in your dataset. You can use the code I gave above but with v0 instead of year to create your new psu and new stratum variables.
egen newpsu = group(v0 v001)
egen newstr = group(v0 v023)
these can be done before or after pooling the data, but you need v023 in the dataset.

1d) I don't know what you mean by pseudo panel level, but I think your weight variable is ok after dividing by 1000000.

4) v022 provides a pairing or grouping of PSUs known as implicit strata that used to be for the calculation of sampling errors. We no longer recommend that approach, but rather to use the explicit strata that were defined for the survey and are found in v023.

5) You can do the pooling of datasets just as easily in Stata, using the append command.

Second message:

1) After looking at thew report, there are in fact 40 strata in the 2003 survey (see appendix B of the 2003 report), and I believe 36 in 2008 (I can't access the report due to a slow connection where I am currently). For the 2003 survey you can recode v0, sprovin and v025 to produce a variable with 40 categories that matches the strata given on page 169 of the 2003 report. Do something similar to produce the strata used for the 2008 survey.

Alternatively, you can use a more approximate definition of strata and just use v023. For the 2008 data you can create v023 as follows:
egen newv023 = group(v024 v025)
check that the coding of the resulting variable matches the codes used for 2003. You would then create a variable that separates these by survey using v0 as described earlier.
(I don't recommend this, but it probably won't make much difference in your significance test

results).

2) You can include whichever region variables you wish to as independent variables. The variables used as strata and the variables used as independent variables do not have to match. See 1) just above about strata - it is not v024, but the 40 strata shown on page 189 (for 2003).

Third message:

DHS data are not panel data - the respondents, households, and clusters are not the same from one survey to the next - so I would not be declaring the data as panel data. You should be using the svyset and svy commands in your analysis.