Subject: Creating panel - using weights
Posted by Nora on Sat, 02 Jul 2016 17:06:08 GMT

View Forum Message <> Reply to Message

Hello,

1-I'd like to collapse the data to create a panel of observations across several years. So in this case, I should collapse across which variable to ensure that my units are constant over the years? Do I need to use the GPS data-set?

2-How do I use the weights in stata on the individual and household level data? Is it only run along with the regression command? If I create summary statistics or graphical figures do I need to specify the weight option? Is there a way to weight the whole data-set like in SPSS?

3-If I am collapsing my data-set, do I use the weight along with the collapse option? If so, would I still need to put it again as an option in the regression command or summary statistics tabulation?

Thanks.

Subject: Re: Creating panel - using weights
Posted by Reduced-For(u)m on Mon, 04 Jul 2016 18:17:24 GMT

View Forum Message <> Reply to Message

Hi Nora,

Responding to your points:

1. a "panel" of what? Countries (with the time dimension being survey year)? If this is the case, you want to use the original weights given in the dataset to "collapse" each survey round down into one observation - that is, use the weights to generate the averages that become the values of your panel. After that, in the regression, it is up to you whether you want to weight those by, say, relative country population or something like that. But you would use the DHS weights in the creation of the aggregates for your panel, and then some other weight (or no weight) to adjust for population sizes in the regression (which is run on the weighted means...if I'm understanding you correctly).

2. See above.

3. Yes, use the weight when you collapse.

Happy to follow-up if I'm not understanding your data setup correctly (as in, don't understand what "panel" you are trying to create).

Subject: Re: Creating panel - using weights

Posted by Nora on Mon, 04 Jul 2016 19:23:35 GMT

Thanks for your reply.

I want to create a panel of clusters (or any constant geographical measure other than governorates due to their low number). But as far as I know the household within each clusters are not the same each year (is this information correct?) so collapsing by clusters and appending the data-set for the years of interest would not be 100% correct because households change, wouldn't it?

Subject: Re: Creating panel - using weights
Posted by Nora on Mon, 04 Jul 2016 20:13:18 GMT

I meant merging* the data-set rather than appending it.

Subject: Re: Creating panel - using weights
Posted by Reduced-For(u)m on Mon, 04 Jul 2016 21:41:24 GMT

The problem isn't that household's change (that is true in any "repeated cross-section" that gets turned into a "panel" of aggregates), the problem is that the clusters change from year to year. That is, they don't represent the same areas at all. Even the "regions" sometimes change from survey to survey. On top of that, clusters are not large enough to be representative samples, and so even if they were constant across time, you'd have problems.

So basically, at a unit smaller than the "region", it is impossible to build the panel you want. At the "region" level (which as you say has too few units), it MAY be possible, but would depend on the country (whether regions stay constant across time).

I know some people have done some GPS stuff about imputing/estimating narrow-area values in repeated surveys to generate a panel like you describe, but it is not really advised by DHS*, and I think often times they are stretching the data far too thin for their own purposes (at least in a panel context, in a cross-section there are times when this might be OK).

Subject: Re: Creating panel - using weights
Posted by Nora on Tue, 05 Jul 2016 13:31:40 GMT

That's very helpful. Thank you very much for your replies!

Subject: Re: Creating panel - using weights
Posted by kingx025 on Sun, 22 Apr 2018 18:30:19 GMT
View Forum Message <> Reply to Message

The other poster explained very well the issue of clusters changing over time and having small numbers, but there is one bit of good news:

IPUMS-DHS has created integrated REGION variables for countries in the DHS, so you can be sure that you are looking at the same geographic footprint across surveys. Here are the two integrated geography variables for Egyptian regions in IPUMS-DHS:
 https://www.idhsdata.org/idhs-action/variables/GEO_EG1988_20 14#codes_section
 https://www.idhsdata.org/idhs-action/variables/GEOALT_EG1988 _2014#codes_section
There are also GIS files for the integrated geographic regions, here:
https://www.idhsdata.org/idhs/gis.shtml

If you are looking across multiple Egyptian DHS surveys and want consistent geographic footprints for regions across years, use one of the above variables.

Miriam King