

Understanding Society User Support - Support #412

Weights for BHPS and Understanding Society

09/09/2015 02:04 PM - Andreas Wiedemann

Status:	Closed	Start date:	09/09/2015
Priority:	High	% Done:	100%
Assignee:	Olena Kaminska		
Category:	Weights		
Description Hello, I've merged the BHPS with the BHPS-subset of Understanding Society to create a longitudinal panel of BHPS respondents up until 2012 (i.e. I use the BHPS portion of Understanding Society). I am not entirely sure which weights I should use for the analysis. I've read the documentation of both dataset, but it is still not clear which weights are the best for my purpose. My goal is to re-create the same underlying population in both datasets, either for the UK or GB. Most importantly, however, I want to be consistent across these two dataset in order to analyze trends in, e.g., income over a time span covering both datasets. Most of my variables of interest are at the household level, but some are at the individual level. Should I use the longitudinal BHPS weights (indin91_lw for individuals or the cross-sectional hhdenbh_xw for households)? And do I have to use weights only in the Understanding Society-part or also in the BHPS part of my panel. Many thanks for your help, Andreas			

History

#1 - 09/11/2015 04:08 PM - Olena Kaminska

If you are doing longitudinal analysis since 1991, you should use indin91_lw weight. You use the weight once in your model - presumably your model will have some information from different years.

hhdenbh_xw are suitable only for cross-sectional household analysis representing only one year.

Hope this helps,

Olena

#2 - 09/11/2015 05:40 PM - Andreas Wiedemann

Great, thanks. This makes sense in a regression framework, but what if I'd like to plot means of a variable over time, say income? Do I have to multiply all values for each individual per year by indin91_lw? And similarly if I'd like to look at household level variables, should I then multiply by hhdenbh_xw?

#3 - 09/22/2015 11:57 AM - Olena Kaminska

You can use in-built stata programs to use weights (which are much easier and save you from making mistakes), such as weight [pw=weight] or svy commands.

For means you can do for example

mean var [pw=weight_name]

If you just want mean of a variable for each wave (and you are not interested in changes, i.e. who changed etc.) then you should use each wave separately and use cross-sectional weights. Depending on the variable and which instrument it comes from you can pick the weight. You will use BHPS weights until 2009 and the 'ub' weights (combined for UKHLS and BHPS weights).

Hope this helps,

Olena

#4 - 10/07/2015 09:10 AM - Redmine Admin

- Status changed from New to Closed

- % Done changed from 0 to 100

#5 - 11/10/2015 10:42 AM - Gundi Knies

- Category set to Weights

- Assignee set to Olena Kaminska

- Target version set to BHPS