Understanding Society User Support - Support #1586

Merging UKHLS waves and optimal weight

09/29/2021 01:08 PM - Gonçalo Gameiro

Status:	Resolved	Start date:	09/29/2021
Priority:	High	% Done:	100%
Assignee:			
Category:			

Description

Hello everyone,

I am doing my master thesis and I am interested in studying the integration/assimilation process of immigrants in the UK namely in two areas: economic integration and cultural assimilation. For this, I thought I would use the UKHLS data set waves 1-10 (or 2-10 in order to use the sample that participated in the BHPS & IEMB), most specifically the indresp.dta of all rounds. I have some questions that i am going to expose here:

1- What is the best way to join all waves into one data set? Option 1: merge all individual level files using PIDP and the xwavedat; Option 2: simply append all individual level files (including xwavedat)?

2- In terms of optimal weight, given that I want to do a longitudinal analysis from round 2-10 and I intend to use the 5min survey, I am doing the following: **svyset i_psu[pw=i_ind5mub_lw], strata(i_strata)** . Is this correct? If I the decide to use waves 1-10 I should do **svyset i_psu[pw=i_ind5mus_lw], strata(i_strata)** right?

3- Now a more theoretical question: Ideally I would like to compare some variables for natives and immigrants (from the 5min survey). How can I be sure a question is asked to both groups? When tabulating some variables (e.g. **bysort a_ukborn: tab a_resjobdeny3** , it seems this question was only asked to non-uk born is this right? Additionally, some variables appear to being asked to the IEMB but when tabulating that variable (e.g. by doing tab f_mabroad) I get 89.79% of observation with the value 'Only available for IEMB'. How can I analyse those observations? Do I need to access the data in a different way? Is it in a different file?

4- How can I interpret proxy variables? Can I take any value from it?

Thanks in advance and I apologise for such 'newbie' questions,
Gonçalo

History

#1 - 09/29/2021 04:09 PM - Gonçalo Gameiro

And if it is merging that I should do, is it 1:1 or m:1?
Thanks!

#2 - 09/30/2021 10:00 AM - Gonçalo Gameiro

Typo: when I say svyset i_psu[pw=i_ind5mus_lw], strata(i_strata) I meant **svyset j_psu[pw=j_ind5mus_lw], strata(j_strata)**, given that it is wave 10

#3 - 09/30/2021 01:10 PM - Understanding Society User Support Team

- Status changed from New to Feedback

- % Done changed from 0 to 80

- Private changed from Yes to No

1. Do you want to set up the indresp data in long format? Then please look at the syntax file which shows you how to do this. In this case of course no merging is needed, you need to append the files. If there are variables you need from xwavedat to this long format file, then do a m:1 matching as the long format file will be at the pidp-wave level while xwavedat is at the pidp level.
Go to this page "https://www.understandingsociety.ac.uk/documentation/mainstage/syntax" and see under "Merging individual files across waves into long format"

2. Yes your choice of weight is correct. Get that from j_indresp and merge it to this long format file created in step 1. The svyset code you suggested is correct but you may get an error message for SE estimation and so this additional code will work.
svyset j_psu [pw=j_ind5mus_lw], strata(j_strata) singleunit(scaled)

3. The extra five minutes questions were asked of the boost sample households and a comparison sample from the GPS and ethnic minorities in GPS who were living in low ethnic minority concentration areas in W1. To know who gets asked a question please take a look at the questionnaire and see the "Universe" field below the question. Please also see this user guide
"https://www.understandingsociety.ac.uk/sites/default/files/downloads/documentation/mainstage/user-guides/user-guide-ethnicity-immigration-researc

h.pdf"

4. When someone is not avaialble to participate in the interview, they are offered the option that someone close to them could fill a shorter version of the questionnaire on their behalf. As other people, no matter how close, are not expected to know subjective, attitudinal information about the sample member, this proxy questionnaire only include factual questions.

You can also do our online Moodle course. https://www.understandingsociety.ac.uk/help/training/online/introduction-course

Best wishes,
Understanding Society User Support Team

#4 - 10/12/2021 02:28 PM - Understanding Society User Support Team

- Assignee set to Understanding Society User Support Team

#5 - 12/14/2021 01:33 PM - Understanding Society User Support Team

- Status changed from Feedback to Resolved

- Assignee deleted (Understanding Society User Support Team)

- % Done changed from 80 to 100