Understanding Society User Support - Support #762

inconsistency between xwavedat and indresp on parental occupation?

04/08/2017 05:13 PM - Carolina Zuccotti

Status:	Closed	Start date:	04/08/2017
Priority:	Normal	% Done:	100%
Assignee:	Carolina Zuccotti		
Category:			

Description

Hello,

I am a bit puzzled with the parental information contained in the indresp file and in the xwavedat file.
The xwavedat variables are supposed to be based on indresp data, but I am not sure why I find the following inconsistencies.
(Perhaps this is my mistake too)

Let me give an example, base don data I am using. I have added parental information in xwavedat to b_indresp.
I have the following variables:

. sum b_masoc90_cc b_masoc00_cc masoc00_cc masoc90_cc

```
Variable |        Obs       Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
b_masoc90_cc |      1,734   54.04268    25.38314        10         99
b_masoc00_cc |      1,734   551.8899    260.5772       111        925
  masoc00_cc |     19,497   580.8958    269.3696       111        925
  masoc90_cc |     23,434   57.34847    26.26591        10         99
```

NOTE THAT I HAVE RECODED VALUES BELOW 0 TO MISSING.

I next do the following tests, with these results:

- tab masoc00_cc b_hhorig = this only has values for the UKHLS sample
- tab masoc90_cc b_hhorig = this has values for both UK and BHPS sample, but BHPS NI is missing

- tab b_masoc00_cc b_hhorig = this has values for both UKHLS and BHPS sample
- tab b_masoc90_cc b_hhorig = this has values for both UKHLS and BHPS sample

I would have expected that the variables from the xwavedat include more cases, and not less.

But there is even a more strange thing in this data:

tab b_masoc90_cc if masoc90_cc==.

I would have expected a "no observations" sign, assuming that masoc90_cc is based on the indresp files (as it is stated here: https://www.understandingsociety.ac.uk/documentation/mainstage/dataset-documentation/wave/xwave/datafile/xwavedat/variable/masoc00_cc): however, I get 1731 cases with a valid value. This means that there are only 3 cases that have a value in both indresp and xwavedat.

Can you explain why this is the case? It seems as if in order to get full parental data I need to merge the values in indresp with the values in xwavedat.

I then tested another variable, to double check this: b_racel and racel_dv

. sum b_racel racel_dv

```
Variable |        Obs       Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
     b_racel |      3,892   4.763361    11.60202         1         97
    racel_dv |     52,065    3.03609     7.31321         1         97
```

. tab b_racel if racel_dv==.
no observations

- This is what I expected.

I would be very grateful if you could clarify what is included in the xwavedat file in terms of parental information, or if there is a document where I could find more information.

Many thanks in advance!!!
And please let me know if you need further clarification; or if I am doing the wrong tests :)

Carolina

History

#1 - 04/13/2017 02:14 PM - Victoria Nolan

- Status changed from New to In Progress

- Assignee set to Alita Nandi

- % Done changed from 0 to 10

- Private changed from Yes to No

Dear Carolina,

Many thanks for your enquiry - the team is looking into it and will get back to you as soon as we can.

Best wishes, Victoria.

On behalf of the Understanding Society data user support team.

#2 - 06/07/2017 04:21 PM - Victoria Nolan

- Status changed from In Progress to Closed

- Assignee changed from Alita Nandi to Carolina Zuccotti

- % Done changed from 10 to 100

Dear Carolina,

Further to your separate correspondence with us, we are now closing this issue, as it has been resolved outside of the forum.

Best wishes, Victoria.