Understanding Society User Support - Support #1212

UKHLS/BHPS Harmonized data set

07/18/2019 09:54 AM - Martin Binder

Status:	Resolved		**Start date:**	07/18/2019
Priority:	Normal		**% Done:**	100%
Assignee:	Martin Binder			
Category:	Data inconsistency			

Description

Dear Madam or Sir,

thank you for putting together such a great harmonized version of both data sets, I think this is really helpful.

I have noticed a potential inconsistency in the harmonization of the marriage status variable, though, which I think might prove problematic. As per the harmonised user guide, the variable „mlstat" (indresp, section household) should be harmonized over BHPS and UKHLS. I don't think this is done based on the correct UKHLS variable. While the BHPS variable mlstat contains yearly information on the different marriage statuses for an individual, from the UKHLS wave 2 onwards, the variable mlstat contains extreme amounts of -8 inapplicable values. In my understanding, this is due to people without marriage status change not being carried in the yearly value according to their status but according to this -8 status. This is a huge change in the coding frame between the two data sets and I think it would be better to harmonize the UKHLS timeframe using the marstat variable, which contains the yearly values for marriage status and would thus lead to an overall marriage status variable that uses the same coding between BHPS and UKHLS. Or am I missing something here?

Secondly, I was wondering whether there is particular reason why lfsato and sclfsato (life satisfaction) is not harmonized across the two data sets. Coding frames only seem minimally different to me and I was wondering whether there are substantive concerns about merging this that I am not aware of.

Thanks for your time and great effort!

Best,

Martin

History

#1 - 07/19/2019 05:25 PM - Gundi Knies

Dear Martin,
according to the question text and notes in the UKHLS questionnaire, the variable _mlstat should be routinely combined with _marstat from the hhgrid (see: the indall data file) when the respondent confirmed their marital status (_mlstatchk==1). This does not appear to have happened and we will inspect & correct the data processing routine. The next data release is planned for Nov 2019. In the meantime, please update the -8 with the value of _marstat on _indall for those where _mlstatchk==1. Sorry for any inconvenience caused.

As regards the life satisfaction variable - yes, it is fine to treat the BHPS and UKHLS variables as identical. The reason the variable names have not been aligned is this:
The UKHLS variable was first asked in the self-completion paper-and-pencil instrument in Waves 1 and 2 and it had been decided to mark out these variables by the 'sc' prefix. However, since Wave 7 some interviews are conducted online and for these respondents all questions are done as a self-completion. This raised the question of whether or not we should continue marking out self-completion variables in the UKHLS by the 'sc' prefix and whether to impose the UKHLS variable names to the harmonised BHPS. We decided not to rename the hBHPS variables just yet but to wait for user feedback.

Thanks for spotting the data inconsistency and for taking the time to let us know!
Best wishes,
Gundi

#2 - 07/19/2019 05:27 PM - Gundi Knies

- Category set to Data inconsistency

- Assignee set to Martin Binder

- Target version set to X M

- % Done changed from 0 to 80

- Private changed from Yes to No

#3 - 07/20/2019 11:05 AM - Martin Binder

Glad I could help pointing this out and thanks for detailed response!
Best,
Martin

Gundi Knies wrote:

> Dear Martin,
> according to the question text and notes in the UKHLS questionnaire, the variable _mlstat should be routinely combined with _marstat from the hhgrid (see: the indall data file) when the respondent confirmed their marital status (_mlstatchk==1). This does not appear to have happened and we will inspect & correct the data processing routine. The next data release is planned for Nov 2019. In the meantime, please update the -8 with the value of _marstat on _indall for those where _mlstatchk==1. Sorry for any inconvenience caused.
>
> As regards the life satisfaction variable - yes, it is fine to treat the BHPS and UKHLS variables as identical. The reason the variable names have not been aligned is this:
> The UKHLS variable was first asked in the self-completion paper-and-pencil instrument in Waves 1 and 2 and it had been decided to mark out these variables by the 'sc' prefix. However, since Wave 7 some interviews are conducted online and for these respondents all questions are done as a self-completion. This raised the question of whether or not we should continue marking out self-completion variables in the UKHLS by the 'sc' prefix and whether to impose the UKHLS variable names to the harmonised BHPS. We decided not to rename the hBHPS variables just yet but to wait for user feedback.
>
> Thanks for spotting the data inconsistency and for taking the time to let us know!
> Best wishes,
> Gundi

#4 - 08/07/2019 11:32 AM - Stephanie Auty

- Status changed from New to Resolved

- % Done changed from 80 to 100