Understanding Society User Support - Support #1666

Clarification of weights when joining datasets and comparing COVID and main survey waves

03/07/2022 02:11 PM - William Shufflebottom

Status:	Resolved	**Start date:**	03/07/2022
Priority:	Normal	**% Done:**	100%
Assignee:	Understanding Society User Support Team		
Category:	Data management		

Description

Hi,

My name is William and I'm working on a social capital publication. I'm a little unsure of which weights we should be using in the "svydesign" in RStudio.
We are interested in the impacts of COVID-19 lockdowns on measures of social capital and are looking at variables in the COVID-19 wave 8 datasets primarily. However, we are also doing some comparisons between COVID-19 wave 8 and the Main Survey waves 6, and 10 where variables are the same. We have also left-joined domains by PIDP to our COVID-19 wave 8 from the xsample and Main Survey wave 11 to include missing domains.
I've looked at the guidance on which weights to use and I'm still a little unsure. We are comparing COVID-19 wave 8 (with joined wave 11 domains) to Main Survey Wave 6 and 10. So, we have the weights:

- Wave 6 - indinub_lw, and indinui_xw (no ui version for the longitudinal lw)
- Wave10 - indinui_lw, and indinui_xw
- COVID-19 Wave 8 - betaindin_xw, and betaindin_lw

We are only producing the mean, standard errors, confidence intervals, coefficient of variation. I think it's just the xw weights we need but I'd like to clarify which weights we should be using for our analysis if that's alright?

Kind regards

Will

History

#1 - 03/07/2022 10:06 PM - Understanding Society User Support Team

- Category set to Data management

- Status changed from New to Feedback

- Assignee set to Understanding Society User Support Team

- % Done changed from 0 to 50

- Private changed from Yes to No

Hi William,

Are you analysing the data cross-sectionally, that is, for example, a mean from Wave 8 Covid-19 survey with a mean from the main survey Wave 6, or longitudinally (in other words, analysing within person change), for instance, analysing whether an employed person became unemployed or a married person divorced. As a general rule, for a cross-sectional analysis you need to use _xw weights, for a longitudinal analysis _lw weights. Please let us know if this answers your question.

Best wishes,
Understanding Society User Support Team

#2 - 03/08/2022 10:27 AM - William Shufflebottom

Hi,

Thanks for your quick response. We are comparing means so that is the *xw then. Thanks. For our variables where we are making comparisons between the means for Wave 8 of the COVID study and with waves 6, and 10 of the Main survey, am I right in saying then that we should use the 'betaindin_xw' weight for running our estimates for the COVID-19 Wave 8 variable list and indinui_xw for waves 6 and 10 of the main survey (as the guidance states that we should use _ui* post wave 6)? Thanks again for your clarifying this.

Best

Will

#3 - 03/08/2022 10:35 AM - William Shufflebottom

update: or, do we use the xw for indicators with no comparison and then have to make our own weights for comparisons between the main survey and a covid survey?

#4 - 03/14/2022 01:40 PM - Understanding Society User Support Team

Hi William,

If you are analysing the main survey data separately from the Covid-19 survey data, so first generating the means and them comparing them, then yes, you should use the indinui_xw for the main survey and the betaindin_xw for the Covid-19 survey. Using ui_xw weights is recommended because this way you will use the whole sample available, so including the IEMBS subsample, other _xw weights available exclude this subsample, in other words, their value for the respondents from the IEMBS subsample equals 0.

Best wishes,
Understanding Society User Support Team

#5 - 06/06/2022 08:57 PM - Understanding Society User Support Team

- Status changed from Feedback to Resolved

- % Done changed from 50 to 100