

Understanding Society User Support - Support #1708

Usual and last pay - question on imputation and missing values

05/30/2022 03:33 PM - Rebeka Balogh

Status:	Feedback	Start date:	05/30/2022
Priority:	Normal	% Done:	80%
Assignee:			
Category:	Income		
Description			
Hi there,			
I'm constructing indicators relating to an individual's current employment in wave 8. My sub-sample consists of employees aged 25 to 60 who reported good self-rated health and no psychological distress in the previous wave.			
The two indicators in question are			
1. the relative pay level (using net labour income fimnlabnet_dv), and			
2. pay volatility (whether there's a discrepancy between usual net pay in current (main) job (paynu_dv) and last net pay in current (main) job (payn_dv)).			
While the ratio of missing observations for relative pay level are low, that's not the case for pay volatility. The reason for that being that while fimnlabnet_dv (and one of its components, paynu_dv) is imputed, payn_dv isn't. Beyond the number of missing cases, I'm more concerned that pay volatility won't be missing at random as certain respondents might be more likely to refuse to answer questions on their pay (usual net pay and overall net labour income will be imputed for them but lay pay won't be). Could you confirm this? With your experience on the income variables, do you think it's safer to drop the indicator on pay volatility altogether? For the analysis I'm doing, it'd ideal if the assumption of missing at random could hold. I've read the UKHLS income data guidance but couldn't find information on non-response to income questions.			
Thanks very much			

History

#1 - 05/30/2022 03:40 PM - Rebeka Balogh

- last pay

#2 - 05/31/2022 09:24 AM - Understanding Society User Support Team

- Category set to Income
- Status changed from New to In Progress
- Private changed from Yes to No

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

We aim to respond to simple queries within 48 hours and more complex issues within 7 working days.

Best wishes,
Understanding Society User Support Team

#3 - 06/15/2022 01:07 PM - Understanding Society User Support Team

- Status changed from In Progress to Feedback
- % Done changed from 0 to 80

Hi Rebeka,

Thanks for your question. Missing data is unlikely to be missing at random, and so dropping missing cases will likely mean your sample is selected. As a robustness check, you could try imputing missing values or creating weights to adjust for this selection.

I note that the way you calculate volatility is non-standard. Usually people define volatility by comparing the current report to the lagged report. If you used this definition, it might help with your missing data issue. So you can measure volatility as a function of fimnlabnet_dv minus its lag (you would then want to use one of the longitudinal weights but as your analysis is longitudinal, I guess you are already doing that).

Hope this helps.

Best wishes,
Understanding Society User Support Team