

Understanding Society User Support - Support #1125

Weighting problem

01/07/2019 04:39 PM - Theodora Kokosi

Status:	Resolved	Start date:	01/07/2019
Priority:	High	% Done:	100%
Assignee:	Theodora Kokosi		
Category:	Weights		
Description Dear all, Happy new year! I am facing a problem with the weights from the youth dataset. I am running an analysis using the Understanding Society's complex sampling design that takes into account the psu, the strata and the weights. The weight I am using is the combined cross-sectional youth weight from wave 3 (c_ythscub_xw), however, when I am asking Stata to first run the survey command (svy) and then impute the dataset it doesn't allow me to proceed. I have discovered that if I remove the weight from the analysis it runs fine so I was wondering if there is a problem with it since I am dealing with data from different sources (adults, household, youth and other linked data to the household or the adults). What is the best way to deal with the weighting when it comes to using data from different levels although my main outcomes are from the youth dataset, hence, I am using the aforementioned weight? Thank you in advance.			

History

#1 - 01/08/2019 10:44 AM - Olena Kaminska

Dear Theodora,

Can you just clarify - what are you imputing: is it one specific variable (item nonresponse) or is it persons (unit nonresponse)? Have you checked whether you still have missingness that you are imputing if weight is not 0? It is possible possible that the weight corrects already for the nonresponse that you are imputing, so the imputation is not needed.

Also, can you tell me which command you are using for imputation?

Thanks,
Olena

#2 - 01/08/2019 12:55 PM - Stephanie Auty

- Status changed from New to In Progress

- % Done changed from 0 to 10

- Private changed from Yes to No

#3 - 01/08/2019 03:31 PM - Theodora Kokosi

Dear Olena,

Thank you for your reply. I am imputing 5 variables. The 3 of them are linked data about air pollution, greenspace and deprivation and the other 2 are from the individual respondents' dataset about the highest educational qualification and the neighbourhood cohesion.

I haven't checked whether I still have missingness if weight is not 0 but I will check now. Thanks!

My decision for imputing was that in the fully adjusted models where all covariates are included the sample drops a lot.

My command is:

- 1) I start by defining the complex samples design: `mi svyset psu [pweight=c_ythscub_xw], strata(strata) singleunit(scaled)`
- 2) Then I define the variables I want to impute: `mi register imputed (variables)`
- 3) I then define the regular variables: `mi register regular (variables)`
- 4) Finally: `mi impute monotone (regress) (imputed variables)=(regular variables), add(20) nomonotonechk`

After the dataset is set I run simple regressions by typing:

mi estimate: svy: regress

However!

- When I run this the imputation syntax I get this message: missing imputed values produced

This may occur when imputation variables are used as independent variables or when independent variables contain missing values. You can specify option force if you wish to proceed anyway.

- Consequently, I forced the imputation and it worked fine but in the fully adjusted models the sample is almost by 1000 cases smaller as if the imputation was not done.

What do you think?

Many thanks,
Dora

#4 - 01/09/2019 05:45 PM - Stephanie Auty

- Status changed from In Progress to Feedback

- Assignee changed from Olena Kaminska to Theodora Kokosi

- % Done changed from 10 to 80

Dear Dora,

Our remit at the User Forum is to answer queries related to Understanding Society data and provide general advice about how to manage the data. Given the number of users we have I'm afraid we cannot advise on individual users' analysis specifically.

We do not provide training in statistical methods but there are a wide range of courses available - NCRM provide a use list of course held across the country <http://www.ncrm.ac.uk/training/>.

If you are using Stata, you may be able to find some answers if you ask on the Stata forum, statalist.

Best wishes,
Stephanie

#5 - 01/10/2019 03:56 PM - Theodora Kokosi

Dear Stephanie,

Thank you for your reply. My initial query was to get general advice about the weights and not statistical advice.

Since I was asked what commands I am using in my analysis, I decided to be as explanatory as possible to help your colleague understand the situation.

I presume that the weights are not the problem in my analysis so I will refer to another forum.

Thank you for your help.

Best wishes,
Dora

#6 - 01/11/2019 10:50 PM - Olena Kaminska

Dora,

It is hard to judge what goes wrong from your description - the issue could be in many places.

I have just one comment for you: conceptually you don't need svy for imputing linked data. Maybe try to impute it in the original data source? This may or may solve the problem - but this is another point to think about.

Best wishes,
Olena

#7 - 08/10/2022 11:30 AM - Understanding Society User Support Team

- Status changed from Feedback to Resolved

- % Done changed from 80 to 100