
Subject: Re: Strata with single PSU

Posted by [user-rhs](#) on Mon, 17 Mar 2014 14:25:47 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hi Nina,

A couple of things. First, are you sure that the ones that were dropped because of missing outcome truly had missing outcome? That is, it is missing because the question was not asked for that person or in that survey round for that country. Have you looked in the country-specific variables to make sure that the variables were not stored in there?

Second, if people are being dropped out of your model because of missing explanatory variables, you have bigger issues than the svyset problem you are describing. Perhaps they are missing values on just one, but not all of your explanatory variables. If you have substantial amounts of people being dropped from the model because of missingness of values, I suggest you find out what variables are causing people to drop out of your model. If they are dropping out in large amounts because of missing values on a small amount of variables, you can create a flag for missing for each variable where the value is 1 if the value is missing and 0 otherwise, and recode the missing values in your explanatory variables to 0. That way, you get to keep all of your observations that are not missing on the outcome in the model. The coefficient on the missing flags may not have meaningful interpretation, but at least you are not selectively losing people over one or two variables with missing values out of the 10 or however many you have in your model.

Either way, seems like you might be dealing with selection issues here. I would investigate the explanatory variables first and see if creating those flags helps. Then, I would look at the outcome variable. Are they missing at random or is there endogeneity/self-selection into answering/not answering the question used for your outcome?

HTH,
RHS