Subject: Re: Using weights for merged PR&IR file NFHS-4(2015-16), India
Posted by Bridgette-DHS on Tue, 12 Feb 2019 16:01:30 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

I just merged the IR and PR files for the survey. There are 699,686 women in the IR file and all of them are in the PR file. The correlation between v005 and hv005 is 0.9982. Switching from v005 to hv005 for your weights cannot affect the problem you are having. I doubt that the problem has anything to do with the weights. Have you tried the decomposition without using weights at all? That's something to check.

In the PR file, the RSBY variable is sh55d. It is binary (0/1). I have not used the Fairlie decomposition but I see that it is specifically for binary variables. The caste variables, sh35 and sh36, respectively, have 4 and 5 categories, respectively. There are 4x5=20 combinations. There are 20x2 combinations with sh55d, and I see that none of them are empty, so your problem is not due to empty cells.

Is the Fairlie method designed for categorical predictors? Could that be the problem?

My recommendation, whenever you have an analytical problem, is that you simplify the data setup as much as possible. But a bigger question is whether you need to apply the Fairlie method, or any other decomposition method, in this setup. You can easily identify which categories or combinations of sh35 and sh36 have high enrollment in RSBY and which ones have low enrollment, just using cross-tabulations. If you want to account for other variables, you can use logit regression. I wonder what other forum users would suggest.