
Subject: Re: Interpretation of Rescaled household level weights for India-NFHS4
Posted by [Bridgette-DHS](#) on Fri, 24 Jan 2020 15:00:16 GMT

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

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

The huge variation in national weights for the NFHS4 is unfortunate, but it arises mainly because of the huge variation in the populations of the states and districts. Below I will paste a list of the mean weight (hv005/1000000) in each state, ordered from smallest to largest. Within each state, there is considerably less variation in the weights.

I agree with your specification of svyset and the use of subpop() with svy. The advice from Stata Technical Support is good but I really don't think re-scaling the weights will eliminate the problem, because Stata automatically re-scales pweights so they add to 1. The location of the decimal point in the weight is irrelevant. You can define wt=hv005 or wt=hv005/1000000 or wt=hv005/100000000 or use any other multiplier or divisor for hv005 and the results (with pweight) will not change. I suggest that you try that, but as I said I doubt that it will eliminate your problem.

I suggest that you add this line right after you load the data: "recast double hv005". That will change hv005 to double precision, and I believe all calculations involving hv005 will then include many more decimal places, in effect averting a matrix calculation which involves division by zero, which is what causes singularity in the variance/covariance matrix. The only alternative I can suggest is to drop the cases with the very lowest values of hv005, because their impact on any national results will be negligible, and that's another way to avoid dividing by zero. Please try this and let us know whether it works.

File Attachments

1) [HV024.JPG](#), downloaded 1312 times
