Subject: Re: Weighting district-level data
Posted by Reduced-For(u)m on Fri, 18 Dec 2015 17:47:39 GMT
View Forum Message <> Reply to Message

1) It should be very easy to re-calculate the district-by-survey-round means again, this time using the weights:

collapse Y [pweight=weight], by(district)

2) You can not do the probability weighting after you collapse - you can only adjust for population differences across districts after you collapse.

3) the clustering is easy and you don't have to do anything until you run the regressions: reg Y X, cluster(district)

4) if you have 200 "districts" does that mean that PSU is your "district"? If so, now you have an entirely new problem - namely, that "districts" are probably changing from round to round. How are you matching your "districts" from survey to survey, and what (administratively or in DHS-terms) is the "district" you are using?

So to sum up: you have to weight when you collapse (construct district-level means), you cluster in the regression, and I am not sure that you really have the information you think you do if you are using such small administrative regions (because usually these change from survey round to round, but maybe Egypt is different).