Subject: Re: Weighting after de-normalization
Posted by Reduced-For(u)m on Wed, 18 Mar 2015 01:14:19 GMT
View Forum Message <> Reply to Message

"if I understand well because I pool surveys of different phases from different countries, I will have two PSU. First, at country level and then at household level, right?" - actually you just want 1 PSU number for each sample PSU, it just needs to be unique across years and countries. So for example, you could generate your PSU numbers by giving each country a number between 10 and 99 and then generating a PSU by "PSU*1000000 + Year*100 + CountryNumber" ... something like that (you could also probably concatenate variable in some way, you just need to create a unique number of each country-X-survey-X-PSU

"Once I de-normalize the weights, fix the PSU and append the datasets, the database is ready for analysis, right" - yep. You just need to use svyset and the svy: prefix before your regressions (so as to actually use the weights and PSUs).

"Do, I need to take into account some special treatment for the weight and PSU for the above two analysis, apart from what I already mentioned?" - nope, just set the svyset*.

This is just a technical note: you will be implicitly weighting regressions here by not just probability weight, but also by the sum of the total weights for the country (that is, by the number of survey rounds and the size of the country, which determine the values of the de-normalized weights). If you have big countries or countries with many more survey rounds than other countries, they will get higher weight in your regression. Maybe they should, maybe they shouldn't - that is just an issue regarding interpretation of the regression coefficients.