Subject: Re: When weights are not supported
Posted by Bridgette-DHS on Tue, 11 Nov 2014 17:26:54 GMT
View Forum Message <> Reply to Message

Another response from Tom Pullum:

I agree with what Professor Cameron said. The fixed and random effects should partially remove the influence of the weights, as well as adjust for unmeasured sources of variation.

I suggest that you first apply, to the same data, a simpler statistical model--single-level rather than multi-level-- that allows you to use weights. Apply that model with weights and then without weights to see how robust the estimates are--that is, to see whether there is a dramatic difference between the weighted and unweighted estimates when the fixed/random effects are included. (The fixed/random effects would be included in both the weighted and unweighted versions.) If there IS a dramatic difference, then maybe you should wait until the multi-level model has been revised to include weights.

There is another way to check robustness. I call this the construction of "fake" weights. First adjust v005 in the way you would like, for example so that the weighted number of cases in each survey is the same or is proportional to the population of the country, etc. Then think of v005 as an fweight rather than a pweight. You could remove a factor of 10,000 so that the arbitrary inflation factor is 100 rather than 1,000,000.

Say that the adjusted v005, still with the factor of 1,000,000, is called v005r. Then try these lines:

First run the model without any weights.

Then:

replace v005r=round(v005r/10000)
expand v005r

and then re-run the model on these data (without any weights).

Neither model will actually use weights, but the second data set has been artificially weighted through the expand command.

There is a substantial issue here--the expanded data file will be about 100 times bigger than the original file and probably too big. To make this strategy feasible you could do some sub sampling. OR you could work from a sample of the original data. The main objective would be to get a sense of how much distortion there will be in the estimates if you use fixed/random effects for countries and strata and do not weight.