Subject: Re: Pooled Cross sections
Posted by Reduced-For(u)m on Wed, 15 Jun 2016 19:48:30 GMT
View Forum Message <> Reply to Message

I can "weigh in" (get it!?) on a couple of these....

3. I don't think this makes sense. Weighting is about adjusting for the probability of showing up in the sample. But what might make sense in this description relates to causal effects. If the effect of some variable of interest is common across all people, then you don't need to weight because one observation is as good as any other one. But, if there is unobserved heterogeneity in treatment effects across people, then the average treatment effect you estimate would be biased, because it would over-weight the treatment effect some people got (those with a low sampling weight) and under-weight others (those with a large sampling weight). So if your model is wrong in certain ways, weighting can maybe make it less wrong.

4. Clustering standard errors and using weights address two different problems. Weights will affect both your p-values and (to a lesser extent) your standard errors. Clustering will ONLY affect your standard errors (and CI/pval). My simulations suggest that clustering at too small of a level when using pooled DHS rounds can lead to SE that are way, way too small. The appropriate level to cluster at depends on exactly what you are doing. One good paper on thinking about that, which is relevant to the DHS context, is the famous "How much should we trust difference-in-difference" paper:

http://economics.mit.edu/files/750

If you tell me what you are trying to estimate (in general) I can maybe give some guidance on how to cluster. But the standard econometric thinking on the matter seems to apply fairly well to the DHS (if you think of appending multiple DHS like appending multiple rounds of the CPS or whatever other household survey is popular in your field).