Subject: Re: accounting clustering effects of women's data when using baby-based analysis
Posted by Reduced-For(u)m on Sat, 28 Sep 2013 21:40:46 GMT
View Forum Message <> Reply to Message

Hey R,

Glad I could be some help, and can answer a few of your questions. The results below I'll respond to in another post because there is something kinda awesome (awesome from an applied stats perspective, not for your research) about that. Anyway, on these questions:

1 - I think you figured it out, but basically you can't use both "svy" and specify a V/C matrix - the svy command is telling stata how you want to weight and compute standard errors, so you can't then tell it to use some other standard error calculation. So when you want to weight and specify an SE computation too, then drop the "svy".

2 - This is tricky, but here is the basic idea: both are specifications of the V/C matrix for error terms - errors within a "group" are allowed to be correlated in some way and have some heteroskedatasticity. In the RE model, you are parametrically modelling this V/C matrix - the within-group off-diagonals are a parameter you estimate that is constant (in some way) within a group. You are giving structure to how the matrix should look, because of your assumptions about the data.

The "clustering" V/C matrix (called an "arbitrary" V/C by some) sets up "groups" similar to an RE estimator, but now you make no parametric assumptions "within group" about the structure of the V/C matrix. Each off diagonal is just E'E (residuals squared or person A's residual times person B's for A,B in same cluster). It allows the model more freedom within group. This tends to lead to higher standard error estimates than the RE model because you are imposing less structure (tradeoff between SE size and believability).

A really good, and fairly readable paper on this is Cameron and Miller's "Robust Inference with clustered data" which goes over REs and clustering and bootstraps.

3 - hmmmm... I'm not sure about these variables, so I'll just say that my main point was that you want a weight that is designed to be used at the level of your analysis, in this case the child. So I would use the weight that comes with the child recode, and not the one from the woman recode. Maybe a DHS staffer can help with this better than I can.

OK. More below on your regression results.