Subject: Re: How do I account for clustering within families?
Posted by Reduced-For(u)m on Sat, 12 Apr 2014 21:01:09 GMT
View Forum Message <> Reply to Message

I think there is a confusion here that stems from different disciplines speaking differently, so let me clarify something.

By "clustering" I mean choosing a specification for the variance/covariance matrix of error terms that accounts for within-cluster heteroskedasticity and serial-correlation. In STATA terms, I mean something like "reg Y X, cluster(clustervar)". This is something that relates to getting your standard errors right, but will not in any way affect point estimates.

I think you're question about "controlling" for within-household effects has to do with point estimates. In that case, you may (and may not) want to include household fixed effects (or dummy variables for each household). This would mean that your estimate is based off within-household differences. It would also limit your effective sample to households with multiple children. When I say "control for household characteristics", I'm usually referring to this, which is about getting the right identifying variation for your model.

But what I was talking about before was "accounting for within-household and within-cluster similarities" in your standard errors (your estimated precision). In that case, you want to cluster at the PSU level because you get all of the benefits of clustering at HH level and the benefits of accounting for the sampling design (though not stratification, which could technically shrink your SEs back down a bit).

What the multi-level models do is, depending on what you choose, something closer to what I call "Random Effects". That is, the V/C matrix on the error terms is parametric in some way, whereas it is "nonparametric" in the clustering case. This gets technical real fast. So let me just repeat the main point:

Household fixed-effects would deal with things like selection (some parents are good, some are bad) or other omitted variable bias problems. Clustering will get the standard errors right. Two distinct problems.

That help? I can try again if it doesn't.