Subject: Comparing coefficients across years in the same country
Posted by Ryan on Mon, 27 May 2013 17:29:19 GMT
View Forum Message <> Reply to Message

I am using four waves of male and female DHS data from Ghana (1993-2008), and my intention is to examine how different independent variables influenced education (v133) over time. I have 3 questions, which are in italics following a short explanatory statement.

First, after reading through the forum I am confused about if I need to denormalize my survey weights if I am comparing across time within the same country. If I do need to denormalize weights then I will use the formula: denormalized weight = weight * total size of survey population / size of population sampled. Can someone advise on if I need to denormalize the weights? If I do, is this formula correct?

Second, the stratification procedures in Ghanaian DHS surveys changed over time. In 1993, the survey was stratified by three ecological zones and then rural/urban. It was also self-weighting. From 1998 to 2008, I believe that the sample was stratified by region and then urban/rural. In the latter three waves, the northern regions were oversampled. I am uncertain as to how I should set up Stata (or R) to deal with these changes. I also don't know if I need to do anything special to the male recodes, as they were sampled similarly.

Third, to conduct the analysis I will add dummies for each DHS wave (e.g. all 1993 observations are 1993Dummy=1, else 0) and a male dummy (all male observations are MaleDummy=1, else 0) and then I'll combine all of the individual datasets. This involves renaming the male variables, but that is all fine. To make the next part easier to explain, I'll pretend that I'm only merging and comparing the 1993 and 1998 portion of the data and I'll use only a few variables. To make valid comparisons over time, I would have a regression formula like:
v133 ~ MaleDummy + UrbanDummy + 1998Dummy + 1998Dummy * (MaleDummy + UrbanDummy)

This will tell me if the effect of being male or urban on educational attainment is changing over time and will provide confidence intervals for the effect of time. I realize this final question is more statistical in nature, but can anyone foresee problems with using the DHS data to conduct this kind of analysis?

Thank you for any thoughts or recommendations.

Ryan

Information on DHS surveys:
1993: http://www.measuredhs.com/pubs/pdf/FR59/FR59.pdf
1998: http://www.measuredhs.com/pubs/pdf/FR106/FR106.pdf
2003: http://www.measuredhs.com/pubs/pdf/FR152/FR152.pdf
2008: http://www.measuredhs.com/pubs/pdf/FR221/FR221%5B13Aug2012%5 D.pdf