Subject: Re: Weights when aggregating survey results
Posted by Bridgette-DHS on Tue, 30 Jun 2020 12:51:32 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Senior Sampling Specialist, Mahmoud Elkasabi:

Thanks for posting such interesting use of DHS data on the forum. I went over your message several times to make sure that I completely understand your design. In the bullets below, I'm documenting my understanding and providing answers to several points. Please let me know if I misunderstood any of the design aspects or if you can add more details about the design and the analysis that might be helpful for the discussion:

• You used a weighted-collapse to estimate survey estimates, so now we have a dataset of Nigerian cities by survey years as unit of analysis (I would call it city-year), and weighted survey estimates for each city-year. Speaking with the multilevel models language, cities are in Level-1 and years are in Level-2.

• In your analysis, to take cities population into account, you should consider the city population as analysis weight. However, I'm not sure whether such data might be available or not. If it is available, the percentage distribution of population in each city can be declared as a weight. For each city, these percentages might change across years, and therefore each city-year might have its unique weight value.

• As an alternative for the actual cities population, the weighted percentage distribution of population in each city can be used as a proxy. I'm assuming that all these surveys have enough sample size to produce reliable population percentage distribution on the city level, including the old surveys. This can be done in the collapse step. These percentages can be used as they are as a weight for the city-year data.

• Now with a weight calculated for each city-year, is it true that the weight has to be the same within each panel? I'm not sure this is true. As in several publications (Rabe-Hesketh and Skrondal 2006; Carle 2009), the fixed weights is desirable for Level-1 weight, not for the final weight. In our case, this is the cities, which means that within each year, Level-1 weight should be the same. This can be done using several re-scaling approaches, such as using the average weight or the average of the squared weights. However, in our case, we don't even have Level-1 and Level-2 weights. So it makes sense to declare the calculated weight as Level-2 weight and declare Level-1 as 1.