Subject: District-level aggregates
Posted by Christian Bommer on Tue, 02 May 2017 17:42:58 GMT
View Forum Message <> Reply to Message

Dear DHS team,

I realize that there have been questions before on the representative of DHS data. The position of DHS on this (please correct me if I'm wrong) is that surveys were (in almost all cases) designed to be representative at the national, regional and urban/rural level. However, I was wondering why this is the case. Specifically, what I want to do is to use DHS data in order to estimate the population share of different ethnic/religious groups at the district level. I am currently unsure whether this would be really problematic to do. I can imagine two potential issues and would kindly like to ask you whether these are the only reasons why DHS advices against district-level analyses:

1) Power: There maybe a number of districts for which only very few clusters where selected such that the margin of error maybe substantial. I believe, however, that this is not necessarily an issue of representativeness but rather of power, i.e. this should be reflected in the CI (provided that a minimum number of clusters is available).

2) Relocation: To protect privacy, clusters were displaced randomly in a radius of 10 km maximum (please note that my question only refers to DHS surveys which were also geocoded). Clusters originally located close to district A may therefore end up in district B. However, if this is the only problem, I would argue that it is possible to simulate the potential bias and construct a measure of uncertainty (I probably only need to know the ethnicity shares with an margin of error of +/-5 percentage points).

Is there anything else that I am missing? I would very much appreciate your thoughts on this.

Kind regards,
Christian

Subject: Re: District-level aggregates
Posted by Bridgette-DHS on Tue, 02 May 2017 19:05:00 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

DHS surveys are representative at the stratum level, which is all the combinations of regions and urban/rural places. In order to get reliable estimates at the stratum level, there is a tendency to over-sample the smaller strata (and then weight them down) and under-sample the larger strata (and then weight them up). You are right, that for districts the problem is with confidence and power, rather than with bias. The expected values of estimates for smaller errors should be the population values, but confidence intervals of the estimates will be wider and the probability of identifying a real difference is smaller. It's always a good idea to look at the standard errors of the estimates, no matter what level you work at.

Displacement of clusters is always within districts, as well as within regions, so that does not affect district-level estimates.

Subject: Re: District-level aggregates
Posted by Christian Bommer on Wed, 03 May 2017 08:47:42 GMT
View Forum Message <> Reply to Message

Dear DHS team,

thank you very much for the quick reply. Regarding your response to the second point: My understanding is that in surveys before 2008 displacement did not respect district boundaries but in surveys after 2008 this was sometimes the case. I took this information from this DHS report (page 9):

https://dhsprogram.com/pubs/pdf/SAR7/SAR7.pdf

So to be really sure whether displacement made clusters cross district boundaries, it would be necessary to check this on a case-by-case basis. Could you kindly confirm this?

Kind regards,
Christian Bommer

Subject: Re: District-level aggregates
Posted by Bridgette-DHS on Wed, 03 May 2017 20:36:01 GMT
View Forum Message <> Reply to Message

Following is another response from Senior DHS Stata Specialist, Tom Pullum:

Yes, you are correct. I believe that in all recent surveys, displacement is within district boundaries, but this was not always required. Even when displacement across a boundary is possible, it shouldn't happen often. But to be safe, you should check.
