
Subject: Matching Data from DHS 1995 to DHS 2016 by district

Posted by [hata](#) on Fri, 09 Jun 2023 00:31:36 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hi there!

For my bachelor's thesis I am trying to estimate the effect of the Free Primary Education policy introduced in 1997 on total years of education for the first stage of a 2sls model to estimate the impact of the policy on other outcomes of interest.

I am using a Difference in Difference specification for the first stage, comparing cohorts affected by the policy (born in 1991:1995) to a cohort unlikely to have profited (born in 1976-1981) and using regional variance in pre-policy levels of enrollment to define regional treatment intensity.

To calculate the pre-policy levels of enrollment per district, I am using the PR file of the 1995 DHS. I am now facing the problem, that in the 1995 DHS survey Uganda consisted of 38 administrative districts that were reported in hv023, while in the mean time, due to decentralisation efforts by the governments, the number of administrative districts has grown to 112 in the DHS 2016 (variable sdistrict and smdistrict in the IR and MR files of the DHS 2016). This means I cannot directly match the calculated treatment intensity measure of the districts in 1995 to my data of 2016.

I've tried to find data on how the districts evolved over time to find out whether I can aggregate the 2016 districts to match the 1995 districts but I have been unsuccessful. Thus, I am reaching out to ask you whether the DHS has any sort of matching variable for this, or any tips on where to find data on how the administrative districts evolved over time and how to aggregate them to make them comparable.

Thank you in advance and best regards,
Hannah

Subject: Re: Matching Data from DHS 1995 to DHS 2016 by district

Posted by [Janet-DHS](#) on Tue, 20 Jun 2023 15:27:19 GMT

[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS staff member, Tom Pullum:

This sounds like an innovative and useful strategy. Very ambitious for a bachelor's thesis.

When countries modify their administrative boundaries and subdivide areas, it is very difficult to establish continuity across successive surveys. Uganda has gone through a major redesign of its level 1 and level 2 administrative units. I think it would be virtually impossible to follow these areas over time.

Have you looked at the GE files, such as UGGE53FL.dta? These have sample clusters as cases and give the latitude and longitude of each cluster. There is some displacement of coordinates to anonymize the clusters, but it is possible that you could construct a grid with lines of latitude and longitude and then, in effect, construct your own areas--squares or rectangles--to

follow over time. You could assign each cluster in each survey to a cell of the grid and then follow those cells over time. That may your best option.
