Subject: Re: Calculating a Representative Wealth Index for Clusters Using DHS Sample Weights
Posted by Anonymous on Mon, 13 Jan 2025 05:14:59 GMT
View Forum Message <> Reply to Message

Thank you so much for your response and the helpful guidance regarding calculating the Wealth Index (WI) at the cluster level. Based on some challenges I encountered during my analysis, I have a couple of follow-up questions.

1. Overlapping Cluster Areas:
Given that GPS coordinates provided for clusters can have up to a 10km positional error, we often use a 10km × 10km square area centered around the given coordinates to ensure that the actual sampling points fall within this boundary. However, after defining these 10km squares, I observed that many clusters have spatially very close coordinates, resulting in significant overlap--sometimes as high as 80-90% between their corresponding areas. Despite this overlap, these clusters' mean WI values often differ significantly. This raises concerns about the representativeness of using the mean WI as the cluster-level indicator. I want to ask if DHS considered such overlapping cluster areas during the survey design. If so, how such scenarios are typically handled to ensure the validity of cluster-level WI values?

2. Weighted vs. Unweighted Mean:
In my work on predicting cluster-level WI using remote sensing data, I noticed that using the weighted mean of WI values often leads to better results than the unweighted mean. However, you mentioned previously that the unweighted mean is the recommended approach for cluster-level WI calculations. Could this observation be a coincidence, or does it suggest that the weighted WI might still have some relevance or utility at the cluster level, despite being theoretically less representative in this context?

I greatly appreciate any insights or advice you could provide on these issues. Thank you for your time and support in helping researchers like me better understand and utilize DHS data.

File Attachments

1) overlap_summary.txt, downloaded 82 times
2) cluster_means_corrected.csv, downloaded 101 times