

Comment on DHS Analytical Study 66, “*The Association of Deforestation and other Environmental Variables with Child Health and Mortality*”, by Shireen Assaf, Andres Gomez, Christina Juan, and Thomas D. Fish.

Qingfeng Li, MHS PhD
Assistant Professor, Department of International Health
Johns Hopkins Bloomberg School of Public Health

November 3, 2020

Background:

The DHS Program was asked to provide an additional review of the innovative methods used in Analytical Study 66. Professor Qingfeng Li of John Hopkins University prepared the following comment.

Comment by Qingfeng Li, MHS PhD:

The analyses in AS66 were rigorously conducted. The authors of that report compiled a variety of data sources, including demographic indicators from DHS and geospatial measures from other sources.

The cross-sectional statistical models in the report are the default modeling option for cross-sectional surveys like DHS. As the authors acknowledged in the report, this modeling model is limited in making a causal inference. But overall the application of the cross-sectional model in the analysis is of scientific merit.

Two possible alternative approaches are suggested.

The first option is matching clusters in successive surveys on the basis of geographic proximity. The authors explained that they had looked into this “pseudo-panel” strategy using the 2006 and 2016 Uganda surveys. They were able to construct 336 pairs of clusters. Multi-level models were proposed to fit the paired clusters, though the report did not include model results. The authors considered accounting for sample design to be the major methodological constraint for conducting the multilevel analyses using the paired clusters.

An important methodological issue with the paired-cluster approach is the relatively small cluster size in DHS surveys. In a pseudo-panel approach, the accuracy of the summary statistics (i.e. the mean or average) of the aggregate units depends on the underlying coverage and size of the survey sample. The obtained cohort averages, however, are consistent, unbiased estimates of the true population cohort means, with the difference resulting from sampling error. The difference between population and sample measures (called measurement errors) is small and can be ignored when the cohort size is large enough. Studies suggest that aggregate measures based on 100 or more members are accurate enough to make the measurement error negligible. However, the cluster size in a typical DHS survey does not exceed that threshold. Therefore, the measurement error could not be ignored and needs to be accounted for. Adjusting for the potential bias due measurement error is an active research area and no established adjustment method is available. Combining those methodological issues with the inexact

matching and mismatching of clusters, the consultant does not think this pair-cluster is a promising approach to estimate the link between deforestation and child health using DHS surveys.

The second option is constructing aggregates for geographic areas, such as subnational administrative units. This approach may generate valid and useful insights into the research question. This approach has been used for public health studies, including a study based on DHS surveys. In a pseudo-panel model, cohorts are the units of analysis. In this case, cohorts would be defined by administrative division or geographic area. The selection of the administrative level is an important step for this approach. That requires striking a balance between the number of areas and the accuracy of area measures. Using a high-level administrative unit for the aggregation implies a large number of records per area and consequently accurate area measures. But it also means a smaller number of areas, which will reduce the sample size for the pseudo-panel models.