Methodological aspects in the construction of a composite indicator of service delivery in Uganda

ABSTRACT

This study elaborates methodological aspects encountered in the building of composite indicators with application to service delivery in Uganda. This study amplified the three main stages, the selection of quality data, the building of the composite indicator itself and statistical approaches that may be utilized to model service delivery composite indicator (CI). This study formulated a data quality assessment framework (DQAF) to enhance the construction of a composite indicator. The DQAF was formulated with a dual orientation that prioritizes two user-oriented data quality components (DQCs) namely; relevance, and interpretability, and three producer-oriented DQCs of methodological soundness, accuracy, and statistical adequacy. The application of the DQAF to service delivery data resulted in the selection of 51 from a pool of 103 potential indicators**,** reflecting a 48.6% acceptability percentage. The composite indicator for statistical regions, which included five dimensions—education, health, water, agriculture, and roads—was developed utilizing official data from the 2021 National Service Delivery Survey conducted by the Uganda Bureau of Statistics, along with various sector performance reports from the Ministry of Health and the Ministry of Water and Environment. Additionally, the study developed an alternative composite indicator for district local governments, concentrating on the education, health, and water dimensions, which was modeled against potential covariates. The composite indicator for statistical regions indicated that Uganda achieved a score of 0.49 ($0 \leq$ composite indicator score ≤ 1) utilizing equal weighting, minimax transformation, and additive aggregation, whereas the score was 0.45 with equal weighting, distance-to-reference point transformation, and geometric aggregation. Min-max transformation yields higher scores compared to distance-to-reference point, attributable to the exogenously determined goalposts. Weights that are participatory determined were comparable with data-derived weights. Robustness tests demonstrated that the constructed composite indicator exhibited stability and can therefore be utilized. The absolute differences in ranks by region were observed, with Kampala and Lango exhibiting the lowest differences and Karamoja and Kigezi the highest, attributable to the presence of outliers and inequitable performance in the examined variables. The aggregation stage was the most sensitive accounting for nearly 60% of the total variance, primarily due to interaction with mainly the

transformation stage; this underscores the necessity to cautiously select an aggregation method, as it greatly influences the robustness of the results.

The absolute rank differences were highest in the education dimension at 2.00 and lowest in the roads and health dimension at 1.33, indicating the varying impact of excluding aspects from the composite index. In assessing the differentials of service delivery at local government level, the composite indicator scores ranged from 0.25 to 0.60, with a substantial portion of the density plot situated below 0.50, indicating inadequate service delivery levels. The scores were negatively influenced by the number of sub-counties and land area, and positively influenced by central government funding, funding from other agencies, number of town councils, and age of the district. While beta regression adeptly models bounded data, random forest regression highlights the relative importance of predictors, and generalized additive model captures non-linear covariate effects. The comparable predictive accuracy of these methods, as evaluated using root mean square error, suggests their applicability to this investigation in accordance with the analytical objectives. A dual orientation of data quality components should be developed collaboratively, ensuring that redundancies and overlaps are recognized and resolved to clarify the intrinsic qualities of data and elementary indicators. Furthermore, it is recommended to investigate penalization methods to manage the substitutability of variables during aggregation due to unequal performance, as well as to employ both parametric and non-parametric techniques to assess differentials in service delivery.

Supervisors: Dr. Yeko Mwanga, Prof. James Wokadala, Prof. Francesca Bassi, Prof. Muuela Scioni

Critical Reader: Prof. Moya Musa

Date of the PhD defense: 18th December 2025 starting at 2:00pm