ABSTRACT

This study examined the development of infant mortality factors from 1995 to 2016. Specifically, the study examined the periodic specific variation in the contribution of individual, household and community factors towards infant mortality between 1995 and 2016; assessed the relative contribution of the individual, household and community factors towards the urban-rural differences in infant mortality in short ran and long run.

The study used secondary data from the Measure DHS data set for Uganda. Individual, household, neighborhood, and national-level characteristics were among the variables examined. The analysis was done at univariate, bivariate (unadjusted logistic regression model), and multivariate (Oaxaca Brenda multi-level logistic decomposition; vector error correction model; Bayesian vector auto regression; and vector auto regression) levels, eleven machine learning techniques were researched and used to produce prediction algorithms for newborns at risk. Multilevel decomposition of the mixed logistic regression model was used to quantify the influence of grouped variables to time variation in the risks of infant mortality.

The effect differences represented a larger share of the total difference in infant mortality in both short and long run. Household-level factors contributed the highest share of the total difference in infant mortality. The effects of maternal factors (C), which include marital status, education level, birth interval, and intended pregnancy, were generally responsible for a greater portion of the overall difference in infant mortality between the years 2001 and 2016 (82.4%) and 2001 to 2011 (31.01%), respectively. To the contrary, proximate factors overall accounted for a larger share of the total difference in infant mortality due to compositional differences in the long run (2001–2016, 11.53%) compared to the short run (2001–2011, 2.77%). Child, father, maternal, and household factors made up a larger share of the explained variance in the risk of infant mortality across communities (PCV) in rural areas. In urban settings, most of the variations were explained by community-level and proximate-level variables. The effect (C) and compositional (E) differences led to significant changes in infant mortality in both rural and urban areas in the short and long run. Irrespective of residence, the effect (C) differences represented a larger share of the total difference in infant mortality during the two periods as compared to compositional differences. However, there were greater variations in the effects between urban and rural areas; they went from 226.11% to 177.33% in the former and from 104.29% to 92.38% in the latter.

The CatBoost algorithm was evaluated using the F1 score, ROC-AUC, accuracy, and MCC and performed best when 28 variables were used, followed by when 8 and 10 variables were used, respectively, based on variable importance as compared to the other 11 models. The model best suited for forecasting IMR and NMR in the long run was VECM since it outperformed BVAR and VAR models based on RMSE, MAPE, etc. In the short run (6 years), more than 80% of the forecast error variance for IMR was explained by itself, just like 89% of the NMR, and very little influence was seen from GDP and GDPP. In the long run (10 to 15 years), the influence of LGDP on IMR increases while that of the GDPP remains stagnant, the influence of GDPP on NMR becomes stronger, and that of GDP remains weak all through the period. In contrast to the VAR and BVAR models, the VECM model provided the most precise long-term projections for both IMR and NMR. The NMR and IMR projections for the short term performed well using univariate time series (ARIMA). The VECM model specifically asserted that GDPP had a longer lasting impact on NMR than GDP and that GDP had a longer lasting impact on IMR.

Today's policymakers should design and implement mechanisms that protect vulnerable populations from the effects of fluctuating national income. In rural areas, increasing coverage and consideration of the health-system context are needed more. In contrast to the VAR and BVAR models, the VECM model provided the most precise long-term projections for both IMR and NMR. The future forecast should be done using VECM due to its long-term performance. The government should focus on interventions that would target an increase in household income to help reduce infant mortality by 2030.