ADVANCING BOTSWANA POVERTY ESTIMATES

Introducing the Survey of Well-being via Instant and Frequent Tracking (SWIFT)

Private Bag 0024, Gaborone
Tel: 367 1300 Toll Free: 0800 600 200

E-mail: info@statsbots.org.bw
Website: www.statsbots.org.bw

STATISTICS BOTSWANA

ADVANCING BOTSWANA POVERTY ESTIMATES

Introducing the Survey of Well-being via Instant and Frequent Tracking (SWIFT)

Preamble

Statistics Botswana (SB) has a standing programme of household surveys, termed inter-censal surveys, with household surveys conducted in between Population and Housing Censuses since the 1981 census. Inter-censal surveys are designed to provide in-depth understanding on topics that may or may not have been covered in the censuses, collecting data on specialised topics to meet stakeholder needs. The surveys, other than the specialised topics, are also designed to have permanent modules on household demographic characteristics, and the economic activity modules carried in the immediate past population and housing census. These allow tracking of socio-demographic changes of the population and labour market indicators estimated in between censuses.

The inter-censal programme of surveys, in most instances are conducted at intervals of 10 years, leading to unavailability of more frequent data to inform policy and programmes, due to long intervals between surveys. During the third inter-censal period (2001-2011), Statistics Botswana conducted the 2002/03 Household Income and Expenditure Survey (HIES), and the 2009/10 Botswana Core Welfare Indicators Survey (BCWIS). The BCWIS as a multi modular survey was the first Multi-Topic Household Survey (MTHS) conducted by Statistics Botswana with plans to reduce the survey's interval from ten (10) years to every five (5) years. The plan was achieved through the conduct of the 2015/16 Botswana Multi-topic Household Survey (BMTHS). BCWIS and BMTHS surveys carried all the HIES modules, and other modules on, Housing, Health, Education, Immunization, Time Use, Community Activities and Access to Services. These modules were intended to capture non-income poverty.

The broad objective of the multi modular survey was to provide a comprehensive and more frequent set of poverty indicators; to inform poverty programming; and to provide household expenditure patterns to inform the Consumer Price Index (CPI) rebasing by providing a new basket of goods and services reflecting the ever changing expenditure patterns.

With pressing needs for more frequent data,Statistics Botswana introduced the Quarterly Multi-topic Survey(QMTS) in 2019 mainly to produce labour force indicators and income poverty predictors. Other modules providing topical indicators were/are included on rotational basis or as per stakeholders' needs to fill indicator gaps. The QMTS is envisaged to produce more frequent indicators that will support the National Monitoring and Evaluation System (NMES), by providing outcome indicators for National Development Plans, National Vision 2036, as well as Sustainable Development Goals (SDGs). The advent of COVID-19 and limited funds led to fewer QMTS survey quarters conducted in a particular year.

The ultimate objective of the QMTS was to have a high frequency programme of household surveys that is predictable, flexible and amenable to the ever increasing and changing data needs for government, private sector, planners and researchers. In contrast to the inter-censal programme of surveys, which to a large extent are infrequent (and the emerging stakeholder needs which have not been planned for and are done on adhoc basis) the QMTS provides a permanent platform for the collection of more frequent socio-economic data.

However, QMTS does not collect data on household consumptions to facilitate poverty estimation. Recognising the imperative for more timely and responsive data, Statistics Botswana with assistance from the World Bank has introduced the Survey of Well-Being via Instant and Frequent Tracking (SWIFT) methodology. The SWIFT methodology is employed here to provide estimates of monetary poverty leveraging on the QMTS data. The methodology represents a paradigm shift in poverty estimation, offering a dynamic and innovative solution to bridge the gap between data collection cycles.

Using the SWIFT Limited Model (which covers all the QMTS quarters conducted) the results shows an increase in poverty rate at National level from 16,1 percent in Botswana Multi Topic Household Survey (BMTHS2015/16) with SWIFT, to 17.0 percent in Q4 of 2019, reduced to 16.7 percent in Q4 of 2020 and hiked to 18.0 percent in 2021 Q4 following which there was a sharp reduction to 14.5 percent in Q4 2022. It should be noted that BMTHS was a yearlong survey, covering all the seasons, while the QMTS was covering some quarters of a year (from 2019 to 2022). The comparison across the QMTS quarters focuses on quarter 4 (Q4) to reduce variation brought in by seasonality.

I wish to thank the World Bank for having trained Statistics Botswana staff the SWIFT methodology. The methodology will help in the production of timely and reliable estimates, which would help in the production of effective policies, programmes and further designed to target interventions in times of crisis.

Finally, I would like to thank the respondents who provided invaluable information for the quarterly survey, and all the stakeholders who contributed to the success of the survey.

Dr. Lucky Mokgatlhe
Acting Statistician General
April 2024

Introduction

The Survey of Well-being via Instant and Frequent Tracking (SWIFT) methodology is a rapid poverty monitoring tool which applies machine learning and multiple imputation techniques to estimate household consumption expenditure and uses adjustments for sampling weights to address the sampling bias, and produce poverty rates.

The report details how the SWIFT Model was used to estimate the overall welfare of households using survey datasets that contain income/expenditure and poverty correlate variables **(see Appendix 2).** The baseline source of this data is the 2015/16 Botswana Multi-Topic Household Survey (BMTHS), which includes detailed income and expenditure information. After establishing the SWIFT model using the BMTHS dataset, it is then applied to Quarterly Multi-Topic Survey (QMTS) datasets from 2019 - 2022, which do not have income and expenditure details, but contain poverty related correlates. The poverty correlates from the QMTS datasets are then used to estimate expenditure/income and poverty rates. SWIFT makes it possible for users to obtain reliable poverty data and profile the poor, at low cost.

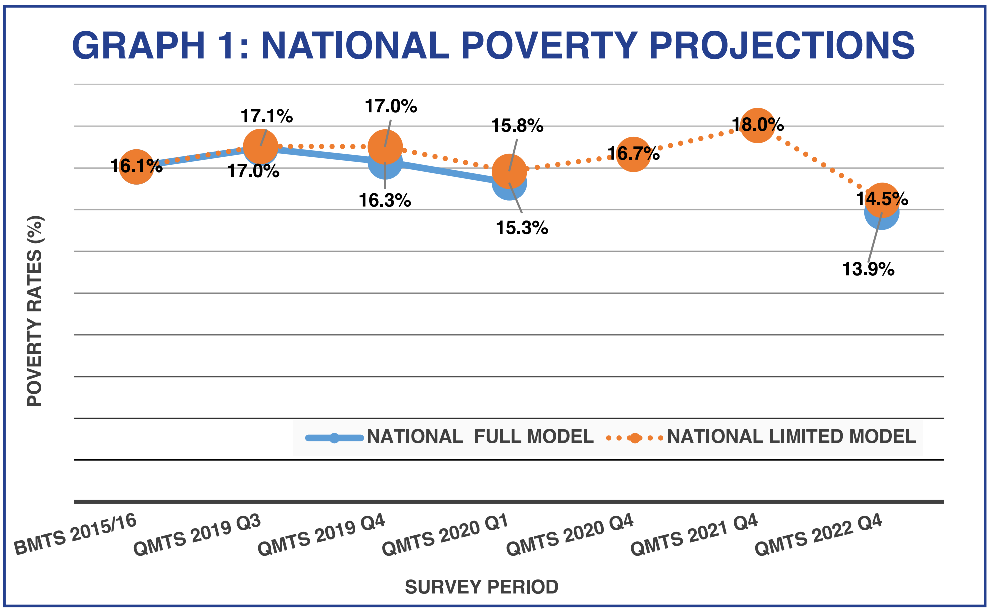
The Analysis was done using 3 (three) models. The modelling utilised a variable set restricted to elements harmonised between BMTHS 2015/16 and QMTS. The three models include Limited Model, Full Model and Full Model plus Food Insecurity Experience Scale (FIES). The limited model utilizes a smaller variable set, excluding dwelling conditions and household income sources, to enable poverty projections for specific quarters. The full model incorporates a comprehensive variable set, including dwelling conditions and household income sources harmonised between BMTHS 2015/16 and QMTS. The Full + FIES model, which was used solely for estimating 2022 Q4 poverty rates includes the full variable set along with a food security variable using data from the Food Insecurity Experience Scale (FIES).

Results

The results presented here are projections of consumption estimates using BMTHS 2015/16 datasets. These datasets are used for multiple imputation into the QMTS 2019, 2020, 2021 and 2022, and using small area estimation theory, household consumptions were simulated in the absence of consumption data in QMTS as it focuses more on labour conditions. Quarterly poverty monitoring benefits from integrating dynamic indicators such as food consumption or food security. Challenges arose due to variations in the reference period for food security questions between BMTHS 2015/16 and QMTS 2019 Q3 – 2021 Q4, posing difficulties in data comparison. The inclusion of the food security variable (FIES) in QMTS 2022 Q4 facilitated more comparison between datasets. The report primarily presents results from the Limited model, as it consistently provides data for poverty projections across most quarters.

National Poverty Predicted Estimates

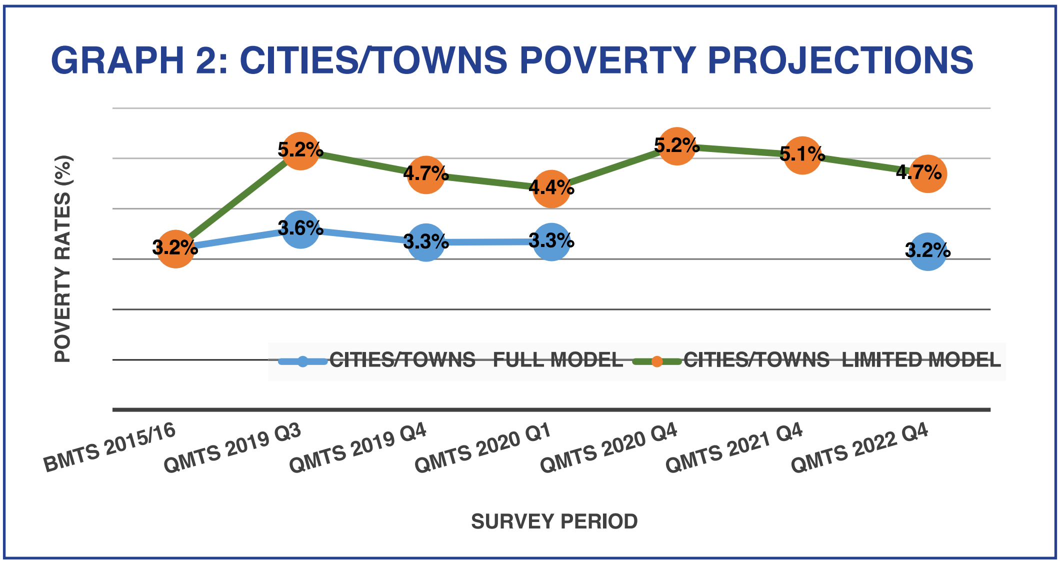
The discussion here and across strata will focus on quarter 4 (Q4) of 2019, 2020, 2021 and 2022. Other quarters are omitted in the discussion to reduce variations that maybe introduced by seasonality variations. Graph 1 Shows the National Poverty rate estimates beginning with 2025/16 at 16,1 percent. The poverty rate is predicted to have increased to 17.0 percent in Q4 of 2019, reduced to 16.7 percent in Q4 of 2020 and hiked to 18.0 percent in 2021 Q4 following which a reduction to 14.5 percent was predicted. The 18.0 percent in 2021 Q4 may be due to economic shocks that were experienced by households at the peak of Covid-19 pandemic in the country, while the decrease to 14.5 percent in 2022 may indicate improvements in economic conditions or the effectiveness of interventions in reducing poverty, offering hope for continued progress in poverty alleviation efforts.



PROJECTED POVERTY ESTIMATES BY STRATA

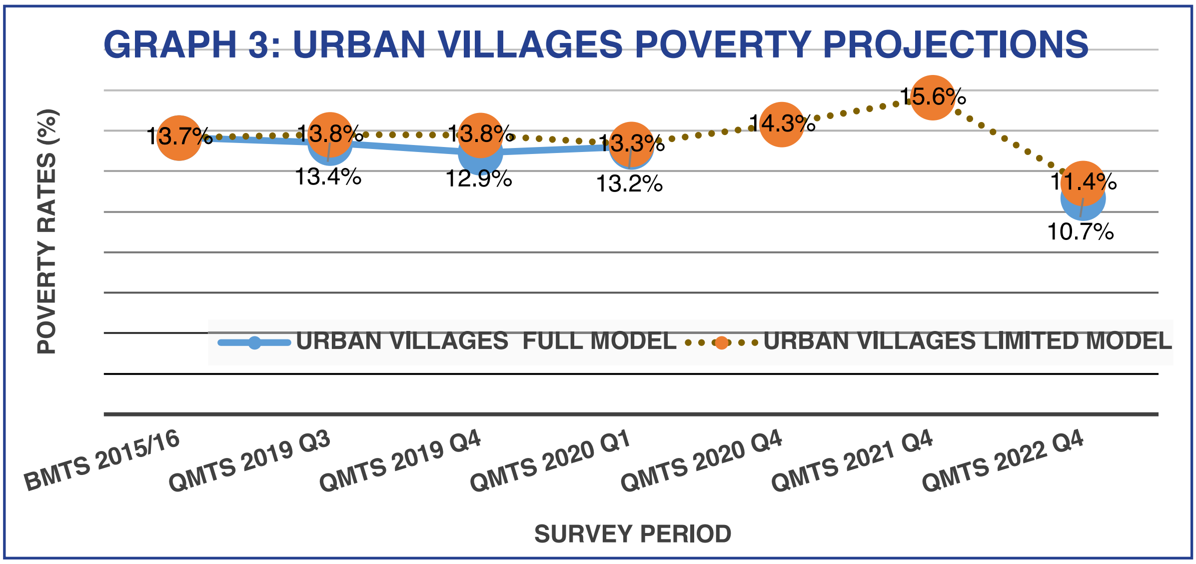
Poverty Estimates for Cities & Towns

Graph 2 shows the trend of poverty rates in cities and towns, using a limited model approach. The poverty rate estimate begins at 3.2 percent in BMTHS 2015/16 (using SWIFT). There is a notable increase in the poverty rate estimate to 4.7 percent in 2019 QMTS Q4 and to 5.3 percent in 2020 Q4, and a slight decrease to 5.1 percent in 2021 and a decrease to 4.7 percent in 2022, suggesting a partial recovery or stabilization of poverty rates. This decrease may indicate improvements in economic conditions or the effectiveness of targeted interventions in addressing urban poverty. These fluctuations may possibly reflect changing economic conditions, or the impact of external factors such as the Covid-19 pandemic on urban areas.



Poverty Estimates for Urban Villages

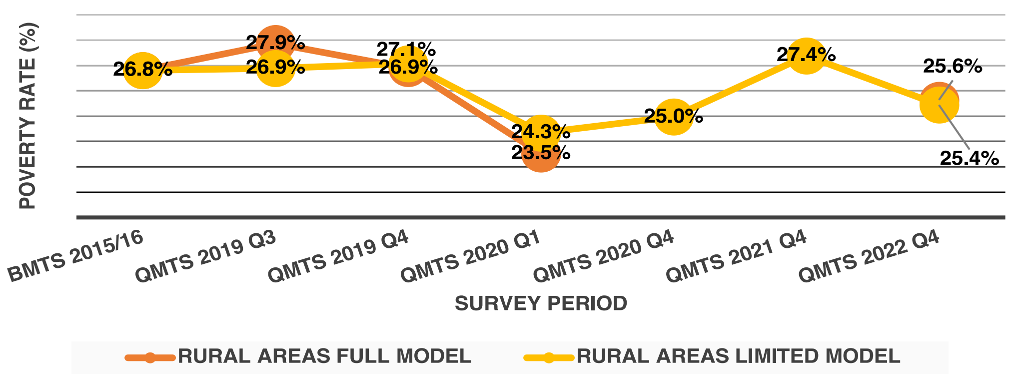
The trend observed in the UV LIM (Urban Villages Limited Model), as reflected in **Graph 3** indicates fluctuations in poverty rates over time within urban village settings. From 13.7 percent in BMTHS 2015/16 to 13.8 percent in 2019 QMTS Q4, increasing to 14.0 percent and 15.6 percent in 2020 and 2021 respectively. A sharp decrease to 11.4 percent is predicted for 2022 Q4. This decrease suggests a positive development, and may reflect improvement in economic conditions within urban village settings, effective interventions or economic resilience within urban villages.



Poverty Estimates for Rural Areas (RA)

The trend observed in **Graph 4,** Rural Areas Limited Model (RA LIM) shows fluctuations in poverty rates within rural areas settings over the surveyed periods from 26.8 percent in BMTHS 2015/16 there is an increase to 27.1 percent in Q4 2019, reducing to 25.0 percent in Q4 2020, and increasing to 27.4 percent in Q4 2021, followed by a decrease to 25.4 percent in Q4 2022. While there are fluctuations, there is a notable high poverty rate estimate in rural areas, which suggests a potential worsening of poverty levels during this period. However, the notable decrease in poverty rates from 27.4 percent to 25.4 percent, potentially indicates improvements or stabilization in economic conditions. Overall, the trend in poverty rates within rural areas appears to be volatile, characterized by fluctuations rather than a clear directional movement. This suggests that rural areas may be experiencing dynamic and complex socio-economic conditions, influenced by various factors such as economic changes, demographic shifts, and external events like the Covid-19 pandemic. These fluctuations underscore the dynamic nature of poverty in rural areas and the need for targeted interventions to address economic challenges and promote sustainable development.

GRAPH 4: RURAL AREAS POVERTY PROJECTIONS



Conclusion

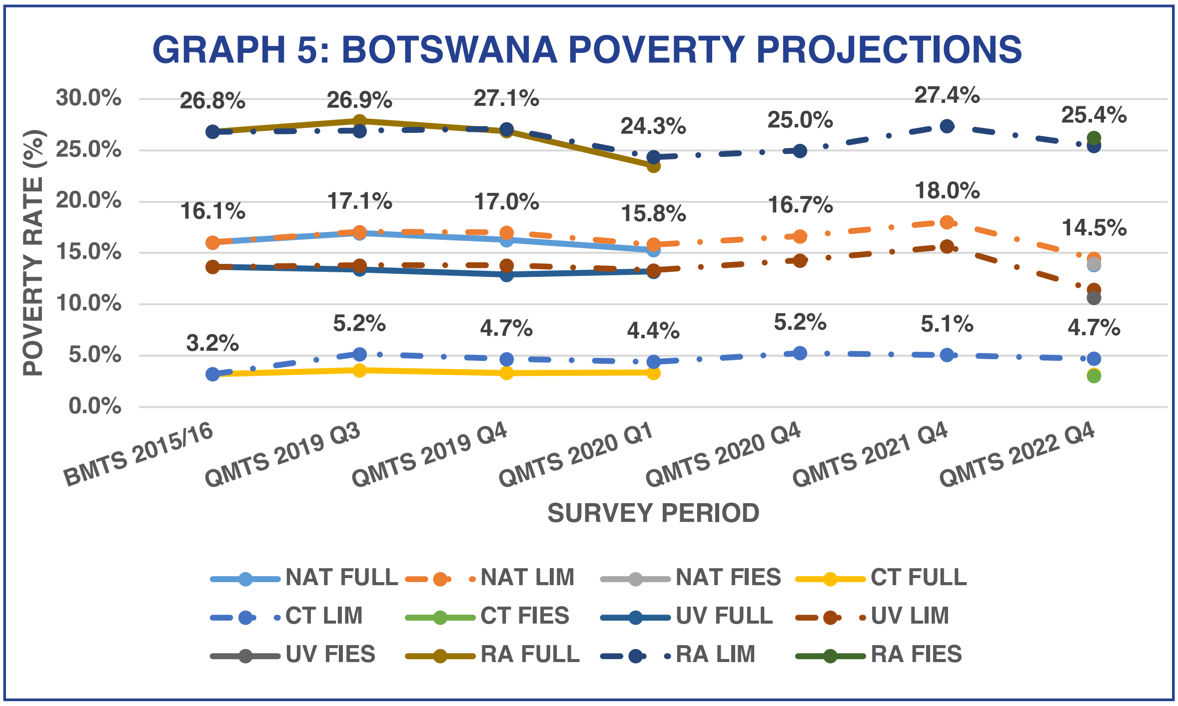
The trend across strata generally show a reduction in poverty in the fourth quarter of 2022. The SWIFT limited model allow for poverty estimation with limited data from the quarterly survey. Going forward, to address indicator gaps, the rising needs of stakeholders for more frequent poverty estimates to inform targeted interventions and policies aimed at reducing poverty, Statistics Botswana will be using SWIFT to estimate poverty in between the major poverty surveys using data from the frequent quarterly multi-topic surveys. The frequent tracking of poverty with SWIFT will reduce the time lag in production of poverty indicators.

Appendix 1: Botswana Poverty Projections from 2015 to 2022 – All Models

Graph 5 presents poverty rate estimates using the SWIFT methodology across various strata, including National Full Model (NAT), National Limited Model (NAT LIM), Cities and Towns Full Model (CT), Cities and Towns Limited Model (CT LIM), Urban Villages Full Model (UV), Urban Villages Limited Model (UV LIM), Rural Areas Full Model (RA), and Rural Areas Limited Model (RA LIM), (NAT FIES)National Full Model + FIES, Cities and Towns Full Model (CT),), Urban Villages Full Model (UV) + FIES, Rural Areas Full Model (RA) + FIES.

The estimates are derived from different survey periods, covering BMTHS 2015/16 and QMTS 2019 Q3, Q4, 2020 Q1, 2020 Q4 and QMTS 2022 Q4. For the limited and full model, at national level, the results shows that poverty rates remained relatively stable from the BMTHS 2015/16 to QMTS 2020 Q1, with a slight increase noted in QMTS 2021 Q4. However, there was a significant decrease observed in QMTS 2022 Q4, indicating potential improvements in economic conditions.

In Cities and Towns, poverty rates have generally shown consistency over the surveyed periods, with minimal fluctuations observed in both the full and limited models. In contrast, poverty rates in Urban Villages and Rural Areas displayed more varied trends. While there were marginal fluctuations observed in urban villages, a notable decrease in poverty rates was evident in QMTS 2022 Q4 compared to QMTS 2021 Q4 for the limited model. Conversely, rural areas exhibited fluctuations, with increases noted in QMTS 2021 Q4, followed by a notable decrease in QMTS 2022 Q4 in a limited model. Overall, the FIES models highlights improvements in poverty rates in quarter 4 2022, across all strata.



Note: data labels in the figure pertain exclusively to the limited models

Appendix 2: Basics and Assumptions of the SWIFT Methodology

The SWIFT methodology utilised core variables on poverty correlates like Districts and regions, household demographics, education, employment, sources of household income, dwelling conditions, food security etc. These were used to estimate household income or expenditure via a model. Poverty statistics were then calculated based on this estimated data. Before modelling, datasets were reviewed for outliers. The models were trained using Botswana Multi-topic household survey 2015/16 which encompassed both income/expenditure and poverty correlates, referred to as the "training dataset." It operates on the assumption of a consistent linear relationship between income/expenditure and poverty correlates over time, with a residual element to account for imperfections.

The SWIFT modelling process involved several steps aimed at refining the estimation of coefficients and the distributions of both the coefficients and the residuals to enhance the model's accuracy. The SWIFT methodology integrates Cross-Validation and Multiple Imputation to project household income or expenditure.

Cross-validation (CV) is a widely used empirical test in machine learning to evaluate model fitness and avoid overfitting. It involves splitting the dataset into training and testing datasets, ensuring that evaluation results are unaffected by overfitting. SWIFT employed a 10-fold CV process, where the BMTHS dataset was divided into ten subsets, or K-folds. A model was trained on nine folds using stepwise Ordinary Least Square (OLS) regression, and its performance was assessed on the remaining fold. This process was repeated with different significance levels to optimize model performance. The optimal p-value identified by the CV exercise was 0.3 percent. Finally, the selected optimal p-value was applied to the full BMTHS 2015/16 dataset to finalize the model **(see Appendix 3 for more details).**

After selecting the optimal significance level, a stepwise OLS regression was performed on the full dataset to estimate the models. To ensure stability, additional regressions were conducted on testing datasets to confirm consistent coefficients and address collinearity issues. Variables affected by collinearity or coefficient changes were removed, and the final models were assessed for reasonableness of coefficients. Imputation of household consumption into the target data (Quarterly Multi-Topic Survey (QMTS), Q3 2019 – Q4 2022) was performed, focusing on the distinct population subset- Cities & towns, urban villages, rural villages. This process involved a stepwise regression to estimate household consumption based on variables derived from the final model created using the training dataset (BMTHS 2015/16) and applied to the target data (QMTS).

The performance of the models was assessed by projecting poverty rates using the imputed household expenditure or income from the training dataset and comparing them to the actual poverty rates observed in the same dataset. This comparison allowed for a comprehensive evaluation of the models' accuracy and effectiveness in predicting poverty rates.

Distinct models were devised for cities and towns, urban villages, and rural villages in Botswana to tackle variations in poverty levels. Additionally, three model types were crafted to adapt to differences in available variables across all QMTS surveys. These models include (1) full, (2) limited, and (3) full + FIES. The full model incorporates a comprehensive variable set, including dwelling conditions and household income sources, harmonized between BMTHS 2015/16 and QMTS. The limited model utilizes a smaller variable set, excluding dwelling conditions and household income sources, to enable poverty projections for specific quarters. The full + FIES model, used solely for estimating 2022 Q4 poverty rates, includes the full variable set along with a food security variable using data from the Food Insecurity Experience Scale (FIES).

Appendix 3: Cross-Validation

The following process was carried out:

Initially, the BMTHS 2015/16 dataset was divided into K subsets, or folds. Each fold underwent a rigorous process: For each fold k = 1,2,...K, a subset of "good" variables were identified using all available data except those within the current fold. With these variables, outcomes were predicted within the current fold and prediction errors were calculated. This process was repeated across all folds to understand the distribution of prediction errors comprehensively.

Adaptation of K-fold cross-validation:

Additionally, the K-fold CV method was adapted with the BMTHS 2015/16 dataset. Following the initial division of data into K folds, each fold underwent the following steps:

Step 1:

a. A stepwise regression was utilized, employing a predetermined significance level (p-value), to select the model using all of the samples except those in fold k.

b. The selected variables were then used to predict outcomes within the current fold (k-fold), and mean squared error was computed.

c. This process (a. and b.) was iterated for all folds k = 1, 2...K, and the average mean squared error was calculated.

Step 1 was repeated for different p-values to determine the optimal one, based on the minimum average absolute difference in poverty rate. The optimal p-value identified by the CV exercise was 0.3 percent. Mean Squared Errors (MSE) metric measures the average squared differences between predicted and actual values, providing insight into the accuracy of the model's predictions. Mean Squared Errors (MSE) metric measures the average squared differences between predicted and actual values, providing insight into the accuracy of the model's predictions. Finally, the selected optimal p-value was applied to the full BMTHS 2015/16 dataset to finalize the model.

Private Bag 0024,
Gaborone
Tel: 367 1300
Toll Free: 0800 600 200

Private Bag F193,
Francistown
Tel: 241 5848 Fax: 241 7540

Private Bag 47
Maun
Tel: 371 5716 Fax: 686 4327

E-mail: info@statsbots.org.bw
Website: www.statsbots.org.bw

STATISTICS BOTSWANA