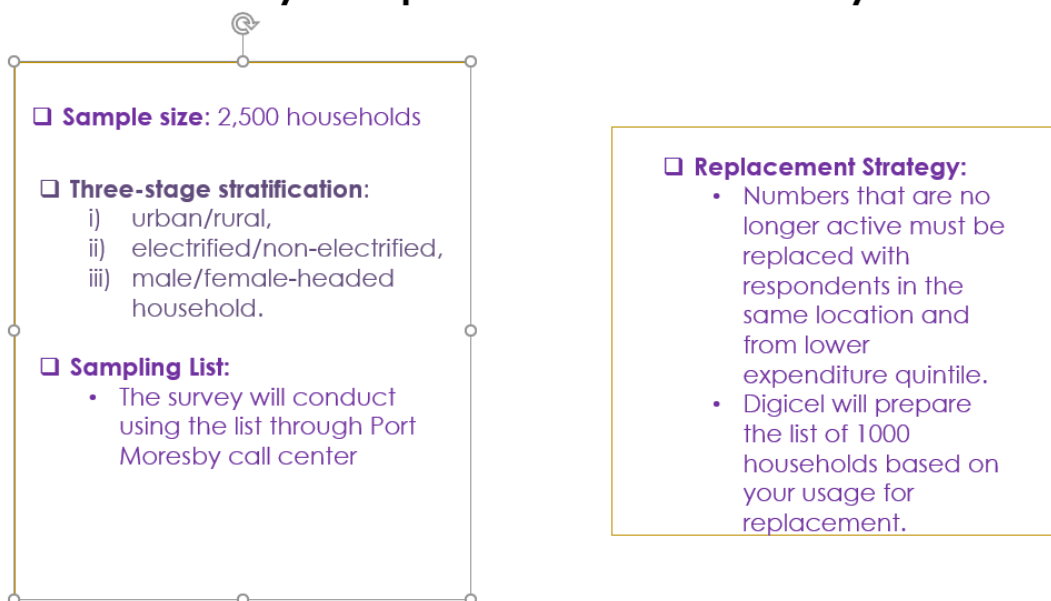


Sampling Strategy in Papua New Guinea

SAMPLING

The total sample was 2,500 households. The Survey Solutions (SUSO) computer software package was used for data collection and management. Based on test questionnaire's results, a lead list of 20,000 random digital dialing numbers (households) were generated for call center enumerators to perform a survey. The total number of interviews recorded on the SUSO application was 4,227, from which 1,560 were rejected due participation refusal, disengagement, or an inability to reach respondents over the phone (calls directly led to voicemails). Accordingly, 2,667 successfully completed surveys were conducted. Sampling weights were used to mitigate the impact of selection bias. The weights were based on information from the 2016–18 PNG DHS and included adjustments for household location, size, and wealth and the respondent's sex and education level. After cleaning and incorporating the wealth index, the sample size was reduced to 2,635. Interviews were conducted in a total of 87 districts. The survey was reweighted based on the 2016–18 DHS nationally representative dataset (World Bank 2020). PNG is divided into four regions, which are its broadest administrative divisions.

MTF Survey in Papua New Guinea– Survey Details



Wealth Index

It is important to note that mobile phone surveys tend to be biased toward wealthier population groups, who are able to use (charged) mobile phones at the time of the survey and live in areas with mobile phone coverage. Therefore, there is a likelihood that households from the lowest expenditure quintiles are under-represented, affecting the accuracy of results. To avoid this bias, a wealth index was calculated to mitigate the potential bias as much as possible.

⁵ Principal component analysis, or PCA, is a dimensionality reduction method that is often used to reduce the dimensionality of large data sets, by transforming a large set of variables into a smaller one that still contains most of the information in the large set.

⁶ Margin of error: To put it simply, the margin of error means the degree of uncertainty that survey results might have. The larger the margin of error is, the more likely it is for results to be further away from the “true figures” for the whole population.

The MTF survey in 2021 re-created the wealth index which allows comparison between the MTF survey and the Demographic and Health Survey (DHS) from 2016 to 2018. The 2016–18 DHS dataset in PNG (DHS 2019) used a wealth index based on household assets and housing characteristics. If households have steadily acquired more assets, they will appear higher in the wealth distribution in this MTF survey than they would in the DHS. According to the Papua New Guinea High Frequency Survey on COVID-19: Results from Round 1 (World Bank 2020), the DHS wealth index (DHS n.d.) is calculated using principal components analysis.⁵ Not all variables of the full DHS wealth index were considered in the mobile phone survey due to the limited survey length. In the recalculation of the wealth index of the phone survey there is more than 98 percent correlation between the recalculated and original measurements. To get the identical measurement, the pooled data was used with a single set of codes. To infer at the population level instead of mobile phone holders, it was essential to re-weight the survey data. To address the potential upward bias, several methodologies were used, such as correlation, logit model, and raking, among others. The detailed methodology of wealth index can be found in Kastelic et al. (n.d.). And the general sample design can be found in Papua New Guinea High Frequency Phone Survey on COVID-19 (Kastelic et al. n.d.), which presents the results from round 2. In the round 2 survey, to address a skew to higher deciles of wealth distribution in round 1, a different stratification mechanism was applied based on the subscribers' characteristics. For example, if households did not send text messages, they were assumed to be less literate. If they received only incoming calls, or credit transferred from other subscribers, the household was assumed to be poorer. The round 3 survey used a similar methodology in wealth index calculation. For the quality check of the cleaned final survey data the margin of error⁶ of the grid access rate is calculated, which is about 2.6 percent.

⁵ Principal component analysis, or PCA, is a dimensionality reduction method that is often used to reduce the dimensionality of large data sets, by transforming a large set of variables into a smaller one that still contains most of the information in the large set.

⁶ Margin of error: To put it simply, the margin of error means the degree of uncertainty that survey results might have. The larger the margin of error is, the more likely it is for results to be further away from the "true figures" for the whole population.