

Staff

William Blackmon

Senior Research Manager, Financial Inclusion Program

Kyla Longman

Research Manager

Jackline Namubiru

Research Associate

Northwestern
GLOBAL POVERTY
RESEARCH LAB



Productivity Costs of Quota Sampling

Case Study: Uganda, Consumer Protection Consumer Survey

IPA Uganda conducted a random digit dial (RDD) survey on consumer protection issues with a completely virtual phone bank and a quota sampling protocol meant to cover a broad selection of adults in the country. Quota sampling involves placing calls until a quota is reached for each combination of respondent characteristics, whose prevalence in the target population is believed to be known. It is a good way to achieve samples that are representative along key dimensions. In some cases, it can increase time and monetary costs substantially to meet quotas for rare combinations of respondent characteristics.

Motivation

A clear threat to the validity of phone surveys is coverage bias. This means that the survey will not be statistically representative of the population of interest because some members are underrepresented in surveying. Survey protocols can be used to construct a more representative sample. These include quota sampling, in which the researcher sets a target for respondents based on known characteristics in a population of interest. Then respondents whose combination of characteristics has been met are screened out. This strategy usually involves making many more calls than will be needed because people from over-represented groups must be screened out.

Depending on cell phone penetration, response rates, and norms around survey response, it can be costly to create representative samples. Determining strategies to build cost-effective representative samples is a vital task.

Sampling

The project team set quotas by region and education level as these characteristics were most over-represented during piloting. For example, if primary-educated adults in the Northern region represent 10 percent of the adult population, such respondents would be ineligible after 10 percent of the total sample size was reached. Each day the team tabulated how many respondents had the combination of relevant characteristics. The team updated survey skip patterns so that any group with more respondents than the maximum allowed would not qualify.

Although the main time use of a completed survey is delivering the surveys itself, contacting respondents, introductions, screening, and explaining to respondents that they don't qualify are meaningful time costs. Due to these costs, the project team determined it was more effective to aim for a minimum call size of 30 respondents, so that reweighting might be possible, instead of a maximum number of respondents. Response rates over time are displayed below to illustrate how completion rate drops and the number of ineligible respondents increases over the course of the survey.

Completion Rate by Survey Day, Among All Attempts



Note: The graph displays the percent of attempts that were completed or ineligible due to quota sampling. Success rates are lower for later attempts independent of eligibility. The average day has a lower completion rate the later it is in the survey since more cases will be later attempts.

This document was made possible by the work of William Blackmon, Kyla Longman, and Jackline Namubiru. IPA's phone survey methods case studies are part of a series on best practices on implementing surveys using computer-assisted telephone interviewing (CATI) and other remote survey modes. These case studies are made possible with the generous support from and collaboration with Northwestern University's Global Poverty Research Lab (GPRL).

Case Study: Productivity Costs of Quota Sampling in Uganda

IPA Uganda conducted a random digit dial (RDD) survey on consumer protection issues with a completely virtual phone bank and a quota sampling protocol meant to cover a broad selection of adults in the country. Quota sampling involves placing calls until a quota is reached for each combination of respondent characteristics, whose prevalence in the target population is believed to be known. It is a good way to achieve samples that are representative along key dimensions. In some cases, it can increase time and monetary costs substantially to meet quotas for rare combinations of respondent characteristics.

January 25, 2021