

Data description for Nigeria conflict telephone-based surveys

2018

The World Bank

National Bureau of
Statistics, Nigeria

Background on the datasets

In this data manual we present background information on two datasets, the first and second round of the Nigeria emergency response survey conducted via telephone among households in three conflict affected regions of Nigeria, North East, North Central and South South.

These are two rounds of telephone based surveys collected from a subsample of households in the Nigeria General Household Survey (GHS).

The first round collected data on conflict exposure between March 22 and May 10th, 2017. The first round of the telephone-based survey had 717 completed interviews with the following geographical distribution: 175 interviews in the North East, 276 in North Central and 266 in South South. The survey first attempted to reach 742 households from the GHS panel, of which 529 could be reached and interviewed. The rest did not have phone numbers or functioning phone numbers (only 2.7 per cent refused to answer). In order to increase the sample size, an additional 288 replacement households were included in the sample also from the GHS panel. Out of these replacement households 188 could be interviewed. Therefore altogether 1030 households were attempted to be reached, with a final sample size of 717 completed interviews.

The second round collected data on food security and market access between August 15th and September 8th, 2017. During the second round of the survey 582 of the 717 households included in the conflict survey were re-interviewed on food security related issues: 147 in the North East, 219 in North Central, and 216 in South South were interviewed. The reason for a smaller sample than in the first round was that not all respondents were reached due to non-functioning phone numbers. Only 3 percent of respondents refused to answer.

NBS in collaboration with the World Bank carried out the survey using mobile phones and captured data in tablet, which was later uploaded to the server after verification. Both teams worked together to design and program the instrument in Survey Solutions.

Sample representativeness

Conflict affected areas were oversampled in order to have a large enough sample of households that in fact experienced conflict events in order to shed light on the type of events that have happened. A random sample of the zones might have given too small sample of conflict affected households and therefore restricted the analysis of the various types of conflict events. Due to the oversampling however, the sample drawn was not representative at the level of the geographical zone, as is the case in the GHS. Therefore probability weights have been created that adjust for the propensity of being in a conflict affected LGA in order to ensure that the sample is representative at the level of the geographical zone.

Description of the datasets

The questionnaires are attached, which provide information on all variables collected. Variables related to personal information have been removed from the datasets. Only data on completed interviews is included in the datasets.

The variable *hhid* is the household id of the households interviewed. It can be used to merge the telephone survey datasets to the GHS.

Round 1: Conflict exposure

1. Round_1_conflict_main

Includes basic information about the household. The unique identifier is *hhid*. Variables are as in the questionnaire. Additionally, variable *replacement* denotes whether the households were among the originally chosen sample or part of the replacement households.

2. Round_1_Conflict_Household_Sec15c

Includes variables in section 2, household conflict exposure. The dataset is at the level of the household event. Therefore, the unique identifier is a combination of the variables *hhid* and *eventId*.

3. Round_1_conflict_community_C6A

Includes variables in section 3, community conflict exposure. The dataset is at the level of the community event. The unique identifier is a combination of the variables *hhid* and *eventId*.

Round 2: Food security

1. Round_2_food_security_main

Includes information at the household level including data in sections 1, 4, 5, 7, 8 and 9. The unique identifier is *hhid*. Variables are as in the questionnaire.

2. Round_2_food_security_hhroster

Includes information at the individual level. GHS 2015-16 household roster serves as the basis of the data collection at the level of the household member. Unique identifier is a combination of the variables *hhid* and *Indiv1*. This dataset includes information on sections 2 and 3.

3. Round_2_food_security_assets

Includes information on assets, that is, section 6. Unique identifier is a combination of the variables *hhid* and *AssetId*.

Weight files

In order to analyze the data so that it is representative at the level of the geopolitical zone, weights have to be used in the analysis. The weights should be used separately for each region. In STATA one needs to set up the survey analysis in the following way:

for North Central and North East:

```
svyset UIDEA2 [pweight=wt_phone_calib], strata(stratum)
```

For South South:

```
svyset UIDEA3 [pweight=wt_phone_calib], strata(stratum)
```

The weight files are named wt_NC, wt_NE and wt_SS.