

Nigeria - COVID-19 National Longitudinal Phone Survey 2020

National Bureau of Statistics (NBS) - Federal Government of Nigeria

Report generated on: August 25, 2021

Visit our data catalog at: <https://www.nigerianstat.gov.ng/nada/index.php>

Overview

Identification

ID NUMBER

NGA-NBS-2020-NLPS-v12-M

Version

VERSION DESCRIPTION

Version 12: Edited, anonymized dataset for public distribution

Overview

ABSTRACT

Nigeria was among the first few countries in Sub-Saharan Africa to identify cases of COVID-19. Reported cases and fatalities have been increasing since it was first identified. The government implemented strict measures to contain the spread of this virus (such as travel restrictions, school closures and home-based work). While the Government is implementing these containment measures, it is important to understand how households in the country are affected and responding to the evolving crises, so that policy responses can be designed well and targeted effectively to reduce the negative impacts on household welfare.

The objective of Nigeria COVID-19 NLPS is to monitor the socio-economic effects of this evolving COVID-19 pandemic in real time. These data will contribute to filling critical gaps in information that could be used by the Nigerian government and stakeholders to help design policies to mitigate the negative impacts on its population. The Nigeria COVID-19 NLPS is designed to accommodate the evolving nature of the crises, including revision of the questionnaire on a monthly basis.

The households were drawn from the sample of households interviewed in 2018/2019 for Wave 4 of the General Household Survey—Panel (GHS-Panel). The extensive information collected in the GHS-Panel just over a year prior to the pandemic provides a rich set of background information on the Nigeria COVID-19 NLPS households which can be leveraged to assess the differential impacts of the pandemic in the country.

Each month, the households will be asked a set of core questions on the key channels through which individuals and households are expected to be affected by the COVID-19-related restrictions. Food security, employment, access to basic services, coping strategies, and non-labour sources of income are channels likely to be impacted. The core questionnaire is complemented by questions on selected topics that rotate each month. This provides data to the government and development partners in near real-time, supporting an evidence-based response to the crisis.

KIND OF DATA

Sample survey data [ssd]

UNITS OF ANALYSIS

- Households

- Individuals

Scope

NOTES

The Nigeria COVID-19 National Longitudinal Phone Survey 2020 covered the following topics:

- Household Roster (Rounds 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)
- Knowledge Regarding the Spread of COVID-19 (Round 1)
- Behaviour and Social Distancing (Round 1)

- Access to Basic Services (Rounds 1, 2, 3, 4, 7, 9, 10, 11)
- Housing (Round 3)
- Credit (Rounds 4, 11)
- Education (Rounds 5, 6, 9, 11)
- Employment (Rounds 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11)
- Income Loss (Round 1)
- Other Income (Rounds 2, 3, 5, 6)
- Income Changes (Rounds 4, 9)
- Food Security (Rounds 1, 2, 4, 7)
- Concerns (Rounds 1, 4, 7, 10)
- Coping/Shocks (Rounds 1, 3, 8)
- Social Safety Nets (Rounds 1, 2, 3, 4, 7, 11)
- Survey of Well-being via Instant and Frequent Tracking (SWIFT) (Round 6)
- COVID Testing and Vaccination (Rounds 6, 10, 12)
- Early Childhood Development (Round 9)
- Phone Signal (Round 11)
- Youth Aspirations and Employment (Round 12)

Coverage

GEOGRAPHIC COVERAGE

National

UNIVERSE

The survey covered all de jure households excluding prisons, hospitals, military barracks, and school dormitories.

Producers and Sponsors

PRIMARY INVESTIGATOR(S)

Name	Affiliation
National Bureau of Statistics (NBS)	Federal Government of Nigeria

OTHER PRODUCER(S)

Name	Affiliation	Role
The World Bank		Collaborated in the implementation of the survey

FUNDING

Name	Abbreviation	Role
Bill and Melinda Gates Foundation	BMGF	Funded the study
Federal Government of Nigeria	FGN	Funded the study

Name	Abbreviation	Role
United States Agency for International Development	USAID	Funded the study

Metadata Production

METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
Development Economics Data Group	DECDG	The World Bank	Documentation of the DDI
NBS ICT Department	NBS-ICT	National Bureau of Statistics	Documentation and Publishing of the DDI

DATE OF METADATA PRODUCTION

2021-08-19

DDI DOCUMENT VERSION

Version 12 (August 2021). This is an update to the Nigeria COVID-19 National Longitudinal Phone Survey documentation with the Round 12 data and documents.

DDI DOCUMENT ID

DDI-NGA-NBS-2020-NLPS-v12-M

Sampling

Sampling Procedure

Wave 4 of the GHS-Panel conducted in 2018/19 served as the frame for the Nigeria COVID-19 NLPS survey. The GHS-Panel sample includes 4,976 households that were interviewed in the post-harvest visit of the fourth wave in January/February 2019. This sample of households is representative nationally as well as across the 6 geopolitical Zones that divide up the country. In every visit of the GHS-Panel, phone numbers are collected from interviewed households for up to 4 household members and 2 reference persons who are in close contact with the household in order to assist in locating and interviewing households who may have moved in subsequent waves of the survey. This comprehensive set of phone numbers as well as the already well-established relationship between NBS and the GHS-Panel households made this an ideal frame from which to conduct the COVID-19 monitoring survey in Nigeria.

Among the 4,976 households interviewed in the post-harvest visit of the GHS-Panel in 2019, 4,934 (99.2%) provided at least one phone number. Around 90 percent of these households provided a phone number for at least one household member while the remaining 10 percent only provided a phone number for a reference person. Households with only the phone number of a reference person were expected to be more difficult to reach but were nonetheless included in the frame and deemed eligible for selection for the Nigeria COVID-19 NLPS.

To obtain a nationally representative sample for the Nigeria COVID-19 NLPS, a sample size of approximately 1,800 successfully interviewed households was targeted. However, to reach that target, a larger pool of households needed to be selected from the frame due to non-contact and non-response common for telephone surveys. Drawing from prior telephone surveys in Nigeria, a final contact plus response rate of 60% was assumed, implying that the required sample households to contact in order to reach the target is 3,000.

3,000 households were selected from the frame of 4,934 households with contact details. Given the large amount of auxiliary information available in the GHS-Panel for these households, a balanced sampling approach (using the cube method) was adopted. The balanced sampling approach enables selection of a random sample that still retains the properties of the frame across selected covariates. Balancing on these variables results in a reduction of the variance of the resulting estimates, assuming that the chosen covariates are correlated with the target variable. Calibration to the balancing variables after the data collection further reduces this variance (Tille, 2006). The sample was balanced across several important dimensions: state, sector (urban/rural), household size, per capita consumption expenditure, household head sex and education, and household ownership of a mobile phone.

Response Rate

BASELINE (ROUND 1): All 3,000 households were contacted in the baseline round of the phone survey. 69 percent of sampled households were successfully contacted. Of those contacted, 94 percent or 1,950 households were fully interviewed. These 1,950 households constitute the final successful sample and will be contacted in subsequent rounds of the survey.

ROUND 2: Interviewers attempted to contact and interview all 1,950 households that were successfully interviewed in the baseline of the COVID-19 NLPS. 1,852 households (95% of the 1,950 attempted) were contacted and 1,820 (93.3%) were successfully interviewed in the second round. Of those contacted, 22 households refused outright to be interviewed and 10 were partially interviewed.

ROUND 3: Interviewers attempted to contact and interview all 1,925 households that were successfully interviewed in the Baseline of the COVID-19 NLPS, excluding 25 households that had refused in Round 2. Thus, the sample included households that were not successfully interviewed in Round 2, in an effort to maintain the sample size. 1,837 households (95.4% of the 1,925 attempted) were contacted and 1,790 (93%) were successfully interviewed in the third round. Of those contacted, 28 households refused outright to be interviewed and 18 were partially interviewed. Of the 1,790 successfully interviewed households, 1,737 were households that have been successfully interviewed in all three rounds of the survey so far. These are the households that form a complete panel across the three rounds.

ROUND 4: Interviewers attempted to contact and interview all 1,881 households that were successfully interviewed in the Baseline of the COVID-19 NLPS, excluding 69 households that had refused in Round 2 or Round 3. Thus, the sample included households that were not successfully interviewed subsequent to the baseline in an effort to maintain the sample size. As shown in 7-11, 1,819 households (96.7% of the 1,881 attempted) were contacted and 1,789 (95.1%) were successfully interviewed in the fourth round. Of those contacted, 19 households refused outright to be interviewed and 9 were partially interviewed. Of the 1,789 successfully interviewed households, 1,691 were households that have been successfully

interviewed in all four rounds of the survey so far. These are the households that form a complete panel across the four rounds.

ROUND 5: Interviewers attempted to contact and interview all 1,856 households that were successfully interviewed in the Baseline of the COVID-19 NLPS, excluding 94 households that had refused in previous rounds of the survey. Thus, the sample included households that were not successfully interviewed subsequent to the baseline in an effort to maintain the sample size. As shown in Table 7-15 of the BID, 1,794 households (96.7% of the 1,856 attempted) were contacted and 1,774 (95.6%) were successfully interviewed in the fifth round. Of those contacted, 13 households refused outright to be interviewed and 5 were partially interviewed. Of the 1,774 successfully interviewed households, 1,656 were households that have been successfully interviewed in all five rounds of the survey so far. These are the households that form a complete panel across the five rounds.

ROUND 6: Interviewers attempted to contact and interview all 1,839 households that were successfully interviewed in the Baseline of the COVID-19 NLPS, excluding 111 households that had refused in previous rounds of the survey. Thus, the sample included households that were not successfully interviewed subsequent to the baseline in an effort to maintain the sample size. As shown in Table 7-19 of the BID, 1,781 households (96.8% of the 1,839 attempted) were contacted and 1,762 (95.8%) were successfully interviewed in the sixth round. Of those contacted, 8 households refused outright to be interviewed and 11 were partially interviewed. Of the 1,762 successfully interviewed households, 1,640 were households that have been successfully interviewed in all six rounds of the survey so far. These are the households that form a complete panel across the six rounds.

ROUND 7: Interviewers attempted to contact and interview all 1,811 households that were successfully interviewed in the Baseline of the COVID-19 NLPS, excluding 139 households that had refused in previous rounds of the survey. Thus, the sample included households that were not successfully interviewed subsequent to the baseline in an effort to maintain the sample size. As shown in Table 7-23 of the BID, 1,740 households (96.8% of the 1,811 attempted) were contacted and 1,726 (95.3%) were successfully interviewed in the seventh round. Of those contacted, 4 households refused outright to be interviewed and 10 were partially interviewed. Of the 1,726 successfully interviewed households, 1,573 were households that have been successfully interviewed in all seventh rounds of the survey so far. These are the households that form a complete panel across the seventh rounds.

ROUND 8: Interviewers attempted to contact and interview all 1,810 households that were successfully interviewed in the Baseline of the COVID-19 NLPS, excluding 139 households that had refused in previous rounds of the survey. Thus, the sample included households that were not successfully interviewed subsequent to the baseline in an effort to maintain the sample size. As shown in Table 7-27 of the BID, 1,738 households (96.0% of the 1,810 attempted) were contacted and 1,723 (95.2%) were successfully interviewed in the eighth round. Of those contacted, 9 households refused outright to be interviewed and 6 were partially interviewed. Of the 1,723 successfully interviewed households, 1,547 were households that have been successfully interviewed in all eight rounds of the survey so far. These are the households that form a complete panel across the eight rounds.

ROUND 9: Interviewers attempted to contact and interview all 1,789 households that were successfully interviewed in the Baseline of the COVID-19 NLPS, excluding 161 households that had refused in previous rounds of the survey. Thus, the sample included households that were not successfully interviewed subsequent to the baseline in an effort to maintain the sample size. As shown in Table 7-31, 1,712 households (95.7% of the 1,789 attempted) were contacted and 1,706 (95.4%) were successfully interviewed in the ninth round. Of those contacted, 4 households refused outright to be interviewed and 2 were partially interviewed. Of the 1,706 successfully interviewed households, 1,533 were households that have been successfully interviewed in all nine rounds of the survey so far. These are the households that form a complete panel across the nine rounds.

ROUND 10: Interviewers attempted to contact and interview all 1,785 households that were successfully interviewed in the Baseline of the COVID-19 NLPS, excluding 165 households that had refused in previous rounds of the survey. Thus, the sample included households that were not successfully interviewed subsequent to the baseline in an effort to maintain the sample size. As shown in Table 7-35, 1,716 households (96.1% of the 1,785 attempted) were contacted and 1,699 (95.2%) were successfully interviewed in the tenth round. Of those contacted, 7 households refused outright to be interviewed and 9 were partially interviewed. Of the 1,699 successfully interviewed households, 1,497 were households that have been successfully interviewed in all ten rounds of the survey so far. These are the households that form a complete panel across the ten rounds.

One new feature of the tenth round was more extensive individual-level data collection on employment of household members, similar to round five⁶. For the employment module, information on working age (15-64 years) members of the household was targeted, including respondents that fall into this age range. However, information was not captured for all members aged 15-64. In order to limit the burden for respondents and interviewers in cases where the number of working age members is large, a maximum of 6 household members were selected (in addition to the main respondent) to capture

information on employment. Therefore, for households with less than 6 working age members, all eligible members were included.

However, 94 percent of interviewed households had 6 or less working age members and only 6 percent had more than six. For the 6 percent with more than 6 working age members, 6 members were randomly selected from among the pool of eligible members. The selection was stratified by sex with an equal split of 3 male and 3 females was targeted, depending on the pool of eligible males and females. However, the application of selection was relatively rare. Of 4,592 working age members of the household in round 10 (excluding the main respondent), information was collected on 4,259 or about 93 percent of eligible individuals. In addition, 1,508 main respondents aged 15-64 were also interviewed bringing the final sample of working age members with employment information to 5,767.

ROUND 11: Interviewers attempted to contact and interview all 1,777 households that were successfully interviewed in the Baseline of the COVID-19 NLPS, excluding 173 households that had refused in previous rounds of the survey. Thus, the sample included households that were not successfully interviewed subsequent to the baseline in an effort to maintain the sample size. As shown in Table 7-39, 1,695 households (95.4% of the 1,777 attempted) were contacted and 1,680 (94.5%) were successfully interviewed in the eleventh round. Of those contacted, 5 households refused outright to be interviewed and 9 were partially interviewed. Of the 1,680 successfully interviewed households, 1,472 were households that have been successfully interviewed in all eleventh rounds of the survey so far. These are the households that form a complete panel across the eleven rounds.

One new feature of the eleventh round was more extensive individual-level data collection on education of household members. For the education module, information on household members 5-18 years was targeted. In order to limit the burden for respondents and interviewers in cases where the number of school-aged members is large, a maximum of 6 household members were selected to capture information on their school attendance and other education variables. Therefore, for households with less than 6 school-aged members, all eligible members were included. However, 92 percent of interviewed households had 6 or less school-aged members and only 8 percent had more than six. For the 8 percent with more than 6 school-aged members, 6 members were randomly selected from among the pool of eligible members. The selection was stratified by sex with an equal split of 3 male and 3 females was targeted, depending on the pool of eligible males and females. However, the application of selection was relatively rare. Of the 4,353 school-aged members of the household in round 6, information was collected on 3,992 or about 91.7 percent of eligible individuals.

ROUND 12: Round 12 focused on interviewing one randomly selected youth in the household and thus, the sample for the round was restricted to households that have at least one person between 15-25 years from previous rounds. Overall, 1238 households that were successfully interviewed in the Baseline of the COVID-19 NLPS had at least one youth 15-25 years of age. This formed the frame for the Round 12 interviews. In each of these 1238 households, one youth 15-25 years was randomly selected to be interviewed. Thus, the sample included households that were not successfully interviewed subsequent to the baseline in an effort to maintain the sample size. During the interview, these selected youths were asked to confirm their age to ensure that the analysis uses information from persons within the 15-25 years cohort.

As shown in Table 7-43, 995 youths (80.4% of the 1,238 attempted) were contacted and 967 (78.1%) were successfully interviewed in the twelfth round. Of the 967 successfully interviewed, 841 (67.9%) of them confirmed to be 15-25 years while 126 (10.1%) are outside the age cohort. Of those contacted, 9 youths refused outright to be interviewed and 19 were partially interviewed. Of the 967 successfully interviewed youths, 635 were members of the household during the GHS-Panel Wave 4 Post-Harvest visit. These members can therefore be traced back to the GHS-Panel data from Wave 4 to form a panel.

Weighting

In order to produce national estimates from the successfully interviewed sample, weights must be applied to the information provided by sampled households. Weights for the GHS-Panel serve as the basis for the Nigeria COVID-19 NLPS, but the weights must be adjusted to reflect the selection and interviewing process. The weights for the Nigeria COVID-19 NLPS were calculated in several stages.

1. Begin with the GHS-Panel full sample household weights.
2. Apply an adjustment factor for the selection into the frame (GHS-Panel households that have contact details). A ratio adjustment was applied at the Zone-level (the strata for the GHS-Panel) to preserve the sum of household weights within each Zone between the full GHS-Panel sample and the NLPS frame.
3. Apply an adjustment for selection into the NLPS sample. The adjustment is a simple expansion factor that is the inverse of the selection probability from the frame for each sampled unit.
4. Apply an adjustment factor for non-contact of sampled households. This was again performed with a ratio adjustment at the Zone-level.
5. Apply an adjustment factor for non-response of contacted households through a ratio adjustment at the Zone-level.
6. Calibrate the weights (following adjustments 2-5) according to the properties of the full weighted GHS-Panel sample. This

calibration step adjusts the weights such that the estimates obtained from the final NLPS sample will match the weighted means of the full GHS-Panel sample for specified characteristics. The calibration was performed using only information obtained from the GHS-Panel interview and thus will only reflect changes in the sample composition and not changes over time. The calibration applied here aims to correct for selection bias that is introduced at any point between identification of the frame and the final successfully interviewed sample. Selection bias is of particular concern in phone surveys since some segment of the population does not have access to a phone and there are more difficult barriers to successfully reach and interview households over the phone. The calibration was applied using the ReGenesee package in R. The characteristics that were considered in the calibration were the same factors included in the balanced sample selection described in 3.1 above. The weights were also applied to the total number of households in the population given by the GHS-Panel weights. 7. Trim the weights. Outlier weights were trimmed using the ReGenesee package in R which adjusts the weights to given bounds while minimizing the deviation from the estimates obtained from the calibration in step 6.

In subsequent rounds of the survey, steps 4, 5, and 6 will be applied to the final baseline weights.

BASELINE (ROUND 1): The weights can be found in the household-level data file (r1_sect_a_3_4_5_6_8_9_12). The variable name is wt_baseline.

ROUND 2: The baseline weights were adjusted for noncontact and nonresponse as well as calibrated following the same procedures outlined above (steps 4, 5 and 6). The round 2 weights can be found in the household-level data file (r2_sect_a_2_5_6_8_12). The variable name is wt_round2.

ROUND 3: In Round 3, two different weights are provided: cross section and panel weights. The cross section weights are applicable to the entire round 3 sample while the panel weights are only applicable to round 3 sample households that have been successfully interviewed in all three rounds of the survey so far. For both of these weights, the baseline weights were adjusted for noncontact and nonresponse as well as calibrated following the same procedures outlined above (steps 4, 5 and 6). The round 3 weights can be found in the household-level data file (r3_sect_a_2_5_6_12). The cross section weight is contained in the variable named wt_round3 while the panel weight is contained in the variable named wt_r3panel.

ROUND 4: In Round 4, two different weights are provided: cross section and panel weights. The cross section weights are applicable to the entire round 4 sample while the panel weights are only applicable to round 4 sample households that have been successfully interviewed in all four rounds of the survey so far. For both of these weights, the baseline weights were adjusted for noncontact and nonresponse as well as calibrated following the same procedures outlined above (steps 4, 5 and 6). The round 4 weights can be found in the household-level data file (r4_sect_a_2_5_5b_6_8_9_12). The cross section weight is contained in the variable named wt_round4 while the panel weight is contained in the variable named wt_r4panel.

ROUND 5: In Round 5, several different weights are provided: two at the household-level and three at the individual-level. The two household weights are the same as have been provided in previous rounds, that is cross section and panel weights. The cross section weights are applicable to the entire round 5 sample while the panel weights are only applicable to round 5 sample households that have been successfully interviewed in all five rounds of the survey so far. For both of these weights, the baseline weights were adjusted for noncontact and nonresponse as well as calibrated following the same procedures outlined in section 2.2 (steps 4, 5 and 6). The round 5 household weights can be found in the household-level data file (r5_sect_a_2_5c_6_12). The cross-section weight is contained in the variable named wt_round5 while the household panel weight is contained in the variable named wt_r5panel.

Given the focus on individual employment information in round 5 and the selection steps outlined above for the sample of working age members, an additional three individual-level weights were calculated and provided in the round 5 data. The individual weights for the employment module were calculated according to:

$$w_{ish} = w_h \times (n_{hs} / N_{hs})^{(-1)}$$

Where w_{ih} is the sampling weight for individual i who is sex s (male or female) in household h , w_h is the final household level weight (i.e. wt_round5), N_{hs} is the total number of eligible household members (aged 15-64) of sex s in household h and n_{hs} is the equivalent number of selected eligible individuals in the household. The individual weights were then calibrated to correspond to the sex and age distribution of the total working age population according to the post-harvest visit of the GHS-Panel.

The basic individual weight described above is the cross section individual weight that considers all individuals that employment information was collected on. This weight is called wt_employ_r5 and can be found in the individual-level employment data file (r5_sect_6b). However, an additional two weights are provided for the panel of individuals interviewed in the GHS-Panel wave 4 and round 5 of the NLPS (i.e. excluding individuals added in the NLPS). The first weight (wt_employ_r5_pp_panel) contains the weight for individuals interviewed in the post-planting visit of the GHS-Panel wave 4 and the second (wt_employ_r5_ph_panel) contains the weight for individuals interviewed in the post-harvest visit of the GHS-Panel wave 4.

ROUND 6: In Round 6, several different weights are provided: two at the household-level and three at the individual-level. The two household weights are the same as have been provided in previous rounds, that is cross section and panel weights. The cross section weights are applicable to the entire round 6 sample while the panel weights are only applicable to round 6 sample households that have been successfully interviewed in all six rounds of the survey so far. For both of these weights, the baseline weights were adjusted for noncontact and nonresponse as well as calibrated following the same procedures outlined in section 2.2 (steps 4, 5 and 6). The round 6 household weights can be found in the household-level data file (r6_sect_a_2_3a_6_9a_12). The cross section weight is contained in the variable named wt_round6 while the household panel weight is contained in the variable named wt_r6panel.

Given the focus on individual education information in round 6 and the selection steps outlined above for the sample of school-aged members (5-18 years), an additional three individual-level weights were calculated and provided in the round 6 data. The individual weights for the employment module were calculated according to:

$$w_{ish} = w_h \times (n_{hs} / N_{hs})^{(-1)}$$

Where w_{ih} is the sampling weight for individual i who is sex s (male or female) in household h , w_h is the final household level weight (i.e. wt_round6), N_{hs} is the total number of eligible household members (aged 5-18) of sex s in household h and n_{hs} is the equivalent number of selected eligible individuals in the household. The individual weights were then calibrated to correspond to the sex and age distribution of the total school-aged population according to the post-harvest visit of the GHS-Panel.

The basic individual weight described above is the cross section individual weight that considers all individuals that employment information was collected on. This weight is called wt_educ_r6 and can be found in the individual-level education data file (r6_sect_5c.dta). However, an additional two weights are provided for the panel of individuals interviewed in the GHS-Panel wave 4 and round 6 of the NLPS (i.e. excluding individuals added in the NLPS). The first weight (wt_educ_r6_pp_panel) contains the weight for individuals interviewed in the post-planting visit of the GHS-Panel wave 4 and the second (wt_educ_r6_ph_panel) contains the weight for individuals interviewed in the post-harvest visit of the GHS-Panel wave 4.

ROUND 7: In Round 7, two different weights are provided: cross section and panel weights. The cross section weights are applicable to the entire round 7 sample while the panel weights are only applicable to round 7 sample households that have been successfully interviewed in all four rounds of the survey so far. For both of these weights, the baseline weights were adjusted for noncontact and nonresponse as well as calibrated following the same procedures outlined in section 2.2 (steps 4, 5 and 6). The round 7 weights can be found in the household-level data file (r7_sect_a_5_6_8_9_12). The cross section weight is contained in the variable named wt_round7 while the panel weight is contained in the variable named wt_r7panel.

ROUND 8: In Round 8, two different weights are provided: cross section and panel weights. The cross section weights are applicable to the entire round 8 sample while the panel weights are only applicable to round 8 sample households that have been successfully interviewed in all eight rounds of the survey so far. For both of these weights, the baseline weights were adjusted for noncontact and nonresponse as well as calibrated following the same procedures outlined in section 2.2 (steps 4, 5 and 6). The round 8 weights can be found in the household-level data file (r8_sect_a_2_6_12). The cross section weight is contained in the variable named wt_round8 while the panel weight is contained in the variable named wt_r8panel.

ROUND 9: In Round 9, two different weights are provided: cross section and panel weights. The cross section weights are applicable to the entire round 9 sample while the panel weights are only applicable to round 9 sample households that have been successfully interviewed in all nine rounds of the survey so far. For both of these weights, the baseline weights were adjusted for noncontact and nonresponse as well as calibrated following the same procedures outlined in section 2.2 (steps 4, 5 and 6). The round 9 weights can be found in the household-level data file (r9_sect_a_2_5_5c_5d_6_12). The cross section weight is contained in the variable named wt_round9 while the panel weight is contained in the variable named wt_r9panel.

ROUND 10: In Round 10, several different weights are provided: two at the household-level and three at the individual-level. The two household weights are the same as have been provided in previous rounds, that is cross section and panel weights. The cross section weights are applicable to the entire round 10 sample while the panel weights are only applicable to round 10 sample households that have been successfully interviewed in all ten rounds of the survey so far. For both of these weights, the baseline weights were adjusted for noncontact and nonresponse as well as calibrated following the same procedures outlined in section 2.2 (steps 4, 5 and 6). The round 10 household weights can be found in the household-level data file (r10_sect_a_2_5_6_9a_12). The cross-section weight is contained in the variable named wt_round10 while the household panel weight is contained in the variable named wt_r10panel.

Given the focus on individual employment information in round 10 and the selection steps outlined above for the sample of working age members, an additional three individual-level weights were calculated and provided in the round 10 data. The individual weights for the employment module were calculated according to:

$$w_{ish} = w_h \times (n_{hs} / N_{hs})^{-1}$$

Where ????? is the sampling weight for individual ?? who is sex ?? (male or female) in household ?, ??? is the final household level weight (i.e. wt_round10), ????? is the total number of eligible household members (aged 15-64) of sex ?? in household ? and ????? is the equivalent number of selected eligible individuals in the household. The individual weights were then calibrated to correspond to the sex and age7 distribution of the total working age population according to the post-harvest visit of the GHS-Panel.

The basic individual weight described above is the cross section individual weight that considers all individuals that employment information was collected on. This weight is called wt_employ_r10 and can be found in the individual-level employment data file (r10_sect_6b). However, an additional two weights are provided for the panel of individuals interviewed in the GHS-Panel wave 4 and round 10 of the NLPS (i.e. excluding individuals added in the NLPS). The first weight (wt_employ_r10_pp_panel) contains the weight for individuals interviewed in the post-planting visit of the GHS-Panel wave 4 and the second (wt_employ_r10_ph_panel) contains the weight for individuals interviewed in the post-harvest visit of the GHS-Panel wave 4.

ROUND 11: In Round 11, several different weights are provided: two at the household-level and three at the individual-level. The two household weights are the same as have been provided in previous rounds, that is cross section and panel weights. The cross section weights are applicable to the entire round 11 sample while the panel weights are only applicable to round 11 sample households that have been successfully interviewed in all eleven rounds of the survey so far. For both of these weights, the baseline weights were adjusted for noncontact and nonresponse as well as calibrated following the same procedures outlined in section 2.2 (steps 4, 5 and 6). The round 11 household weights can be found in the household-level data file (r11_sect_a_2_5_5b_6_12b_12). The cross section weight is contained in the variable named wt_round11 while the household panel weight is contained in the variable named wt_r11panel.

Given the focus on individual education information in round 11 and the selection steps outlined above for the sample of school-aged members (5-18 years), an additional three individual-level weights were calculated and provided in the round 11 data. The individual weights for the education module were calculated according to:

$$w_{ish} = w_h \times (n_{hs} / N_{hs})^{-1}$$

Where ????? is the sampling weight for individual ?? who is sex ?? (male or female) in household ?, ??? is the final household level weight (i.e. wt_round11), ????? is the total number of eligible household members (aged 5-18) of sex ?? in household ? and ????? is the equivalent number of selected eligible individuals in the household. The individual weights were then calibrated to correspond to the sex and age8 distribution of the total school-aged population according to the post-harvest visit of the GHS-Panel.

The basic individual weight described above is the cross section individual weight that considers all individuals that employment information was collected on. This weight is called wt_educ_r11 and can be found in the individual-level education data file (r11_sect_5c.dta). However, an additional two weights are provided for the panel of individuals interviewed in the GHS-Panel wave 4 and round 11 of the NLPS (i.e. excluding individuals added in the NLPS). The first weight (wt_educ_r11_pp_panel) contains the weight for individuals interviewed in the post-planting visit of the GHS-Panel wave 4 and the second (wt_educ_r11_ph_panel) contains the weight for individuals interviewed in the post-harvest visit of the GHS-Panel wave 4.

ROUND 12: In Round 12, two different weights are provided: cross section and panel weights. The cross section weights are applicable to the entire round 12 sample of individuals while the panel weights are only applicable to round 12 sample individuals that were also members of the household at the time of the GHS-Panel Wave 4 Post-Harvest visit. For both of these weights, the baseline household weights were adjusted for (1) selection among youths in the household with only one youth being randomly selected, (2) noncontact of the selected youth, (3) nonresponse of the contacted youth, and (4) ineligibility of interviewed individual. After these simple ratio adjustments were performed, the weights were then calibrated according to the demographic breakdown of youths (by sex and 3 age groups: 15-17, 18-21, and 22-25 years) in the GHS-Panel Wave 4 sample. The round 12 individual weights can be found in the data file (r12_sect_a_12.dta). The cross-section weight is contained in the variable named wt_youth_r12 while the panel weight is contained in the variable named wt_youth_r12_panel.

Questionnaires

Overview

BASELINE (ROUND 1): One questionnaire, the Household Questionnaire, was administered to all households in the sample. The Household Questionnaire provides information on demographics; knowledge regarding the spread of COVID-19; behaviour and social distancing; access to basic services; employment; income loss; food security; concerns; coping/shocks; and social safety nets.

ROUND 2: One questionnaire, the Household Questionnaire, was administered to all households in the sample. The Household Questionnaire provides information on demographics; access to basic goods and services; employment (including non-farm enterprise and agricultural activity); other income; food security; and social safety nets.

ROUND 3: One questionnaire, the Household Questionnaire, was administered to all households in the sample. The Household Questionnaire provides information on demographics; access to basic goods and services; housing; employment (including non-farm enterprise and agricultural activity); other income; coping/shocks; and social safety nets.

ROUND 4: One questionnaire, the Household Questionnaire, was administered to all households in the sample. The Household Questionnaire provides information on demographics; access to basic goods and services; credit; employment (including non-farm enterprise, crop farming and livestock); food security; income changes; concerns; and social safety nets.

ROUND 5: One questionnaire, the Household Questionnaire, was administered to all households in the sample. The Household Questionnaire provides information on demographics; education; employment (including non-farm enterprise and agricultural activity); and other income.

ROUND 6: One questionnaire, the Household Questionnaire, was administered to all households in the sample. The Household Questionnaire provides information on demographics; education; employment (including non-farm enterprise); COVID testing and vaccination; and other income.

ROUND 7: One questionnaire, the Household Questionnaire, was administered to all households in the sample. The Household Questionnaire provides information on demographics; access to basic services; employment (including non-farm enterprise); food security; concerns; and safety nets.

ROUND 8: One questionnaire, the Household Questionnaire, was administered to all households in the sample. The Household Questionnaire provides information on demographics; employment (including non-farm enterprise and agriculture); and coping/shocks.

ROUND 9: One questionnaire, the Household Questionnaire, was administered to all households in the sample. The Household Questionnaire provides information on demographics; education; early childhood development, access to basic services, employment (including non-farm enterprise and agriculture); and income changes.

ROUND 10: One questionnaire, the Household Questionnaire, was administered to all households in the sample. The Household Questionnaire provides information on demographics; access to basic services; employment (including non-farm enterprise and agricultural activity); concerns and COVID testing and vaccination.

ROUND 11: One questionnaire, the Household Questionnaire, was administered to all households in the sample. The Household Questionnaire provides information on demographics; credit; access to basic services; education; employment (including non-farm enterprise); safety nets; youth contact details; and phone signal.

ROUND 12: One questionnaire, the Household Questionnaire, was administered to all households in the sample. The Household Questionnaire provides information on youth aspirations and employment; and COVID vaccination.

Data Collection

Data Collection Dates

Start	End	Cycle
2020-04-20	2020-05-11	Baseline (Round 1)
2020-06-02	2020-06-16	Round 2
2020-07-06	2020-07-20	Round 3
2020-08-09	2020-08-24	Round 4
2020-09-07	2020-09-21	Round 5
2020-10-09	2020-10-24	Round 6
2020-11-07	2020-11-23	Round 7
2020-12-05	2020-12-21	Round 8
2021-01-09	2021-01-25	Round 9
2021-02-06	2021-02-22	Round 10
2021-03-13	2021-03-30	Round 11
2021-04-10	2021-04-28	Round 12

Data Collection Mode

Computer Assisted Telephone Interview [cati]

DATA COLLECTION NOTES

ORGANIZATION OF FIELDWORK: Data were collected by trained NBS interviewers who individually made phone calls from their respective homes. Since the country was on lockdown during the preparation and data collection exercise for the baseline round, interviewers were not allowed to be in the office. Therefore, all interviews were conducted from interviewers' homes. Although the lockdown restrictions were partially lifted following the baseline, interviewers will continue to conduct interviews from home in subsequent rounds until it is deemed safe for them to return to the office. In addition, all other correspondence to the interviewers were made through WhatsApp, phone and emails.

PRE-LOADED INFORMATION: Basic information on every household was pre-loaded in the CATI assignments for each interviewer. The information was pre-loaded to (1) assist interviewers in calling and identifying the household and (2) ensure that each pre-loaded person is properly addressed and easily matched to the most recent interviews. Basic household information (location, household head name, phone number, etc.) was pre-loaded. The list of individuals from the previous interview and their basic characteristics were uploaded. This helped maintain the panel of individuals and ensured the status of each individual in the subsequent round of the survey.

RESPONDENTS: Each round of the Nigeria COVID-19 NLPS has ONE RESPONDENT per household. The respondent was the household head or a knowledgeable adult household member. The respondent must be a member of the household. Unlike many other household surveys, interviewers were not expected to seek out other household members to provide their own information. The respondent may still consult with other household members as needed to respond to the questions, including to provide all the necessary information on each household member.

Interviewers were instructed to make every effort to reach the same respondent in subsequent rounds of the survey, in order to maintain the consistency of the information collected. However, in cases where the previous respondent was not available, interviewers would identify another knowledgeable adult household member to interview.

DATA MONITORING AND EVALUATION: As an additional aid to ensuring good quality data, extensive monitoring was performed throughout the fieldwork for each round of the survey. Three monitoring exercises were implemented during data collection. First, Survey Solutions' audio recording functionality was activated for 15 percent of the sample. These interview recordings were audited by 3 trained monitors, though not all recorded interviewers were able to be reviewed due to personnel constraints. On a daily basis, the monitors will listen to these recordings and fill in a structured questionnaire with their observations on interviewer performance. The feedback from these audio audits are then filtered to the respective interviewers.

The second quality check implemented were call backs to contacted households. The call backs were conducted by trained interviewers who are not part of the main data collection interviewers. Each day, up to 36 households that were contacted by the interviewing team are called by these call back interviewers. The call back interviewers conduct a short interview with the household to confirm that the interviewer did indeed conduct the interview, that certain key elements were clearly stated to the respondent, that the interviewer conducted themselves in a professional manner, and other details on the interview process. Further, the call back team asked several time-invariant questions of the respondent to further confirm

the interview was fully conducted and the interviewer captured the information correctly. Feedback from call backs were routed to the respective interviewers to improve on identified areas. Further, the call back interviewers also called households that were not successfully contacted by the main interviewer. In some cases, the call back interviewer was able to reach the household. In such cases, the case was sent back to the interviewer to conduct the interview.

The third quality check was interviews with “mystery respondents”. These were interviews conducted with the monitoring team without the interviewer’s knowledge. Interviewers were given an assignment with pre-filled details from a household not selected for the NLPS but where the prefilled contact details routed the call to a member of the monitoring team. The mystery respondents were given pre-determined answers to questions in the questionnaire such that when the interviewer calls, they should provide those responses. A short questionnaire was also prepared for the mystery respondents to fill during or immediately after the interview to share their feedback on the interviewer’s performance. The feedback from this exercise were routed to the interviewers to improve on areas highlighted by the monitoring team.

As a result of these quality checks, some of the interviewers were dropped from participating in the survey. There were also regular check-ins to address questions and issues the interviewers might have.

Data Collectors

Name	Abbreviation	Affiliation
National Bureau of Statistics	NBS	Federal Government of Nigeria

Data Processing

Data Editing

COMPUTER ASSISTED TELEPHONE INTERVIEW (CATI): The Nigeria COVID-19 NLPS exercise was conducted using Computer Assisted Telephone Interview (CATI) techniques. The household questionnaire was implemented using the CATI software, Survey Solutions. The Survey Solutions software was developed and maintained by the Data Analytics and Tools Unit within the Development Economics Data Group (DECDG) at the World Bank. Each interviewer was given two tablets, which they used to conduct the interviews. Overall, implementation of survey using Survey Solutions CATI was highly successful, as it allowed for timely availability of the data from completed interviews.

DATA COMMUNICATION SYSTEM: The data communication system used in the Nigeria COVID-19 NLPS was highly automated. Each interviewer was given a mobile modem allowing for internet connectivity and daily synchronization of their tablet. This ensured that head office in Abuja has access to the data in real-time. Once the interview is completed and uploaded to the server, the data is first reviewed by the Supervisors, and then routed for call back or audio audit if selected. A feedback questionnaire was also designed in Survey Solutions where interviewers receive respective feedback on their tablet from the various monitoring stages. This activity is done on a daily basis throughout the duration of the data collection.

DATA CLEANING: The data cleaning process was done in three main stages. The first stage was to ensure proper quality control during the fieldwork. This was achieved in part by incorporating validation and consistency checks into the Survey Solutions application used for the data collection and designed to highlight many of the errors that occurred during the fieldwork.

The second stage cleaning involved the use of Supervisors in Survey Solutions. As indicated above, once the interview is completed and uploaded to the server, the Supervisors reviewed completed interviews for inconsistencies and extreme values. Depending on the outcome, they can either approve or reject the case. If rejected, the case goes back to the respective interviewer's tablet upon synchronization. The supervisor will provide general and question-specific comments when rejecting a particular completed interview. These errors were then corrected based on another call to the household on the instruction of the supervisor. The data that had gone through this first stage of cleaning and has no issues is then approved by the Supervisor.

The third stage of cleaning involved a comprehensive review of the final raw data following the first and second stage cleaning. Every variable was examined individually for (1) consistency with other sections and variables, (2) out of range responses, and (3) formatting. Some minor errors remain in the data where the diagnosis and/or solution were unclear to the data cleaning team.

CONFIDENTIAL INFORMATION: For purposes of maintaining the confidentiality of the data, all names, phone numbers, and addresses have been removed from the datasets.

Data Appraisal

No content available

