

AfterAccess methodology: India and Bangladesh

LIRNEASIA, 2018

The AfterAccess survey was implemented by LIRNEasia in India and Bangladesh in October and November 2017.

The AfterAccess methodology was applied with some local adjustments, considering the fact that in both countries, the most granular sample frame available was at the ward and village level. The AfterAccess sampling methodology has been developed by Research ICT Africa and applied over the past decade in numerous countries in Africa¹ to achieve nationally representative results for multiple target groups simultaneously in a cost-effective way.

In India, a further adjustment was made (detailed in the following sections) to balance the twin priorities of capturing the diversity of the population and managing fieldwork costs in the vast nation. A larger sample size in India was also implemented in consideration of these factors.

The fieldwork was conducted in Asia by competitively-procured market research companies – IPSOS Public Affairs (India) in India and Bangladesh. The companies were mainly involved in the fieldwork set-up (including scripting, translating and pilot testing the questionnaire and training of enumerators) and execution as well as dataset delivery. LIRNEasia monitored the companies in most cases being involved in the field training and monitoring the fieldwork both on the ground as well as remotely.

The details of the methodology applied in India and Bangladesh are provided in the following sections; where the methodology diverges between the two countries, separate sub-sections are used to describe the different approaches.

¹ See for example, https://researchictafrica.net/publications/Towards_Evidence-based_ICT_Policy_and_Regulation_-_Volume_1/RIA%20Policy%20Paper%20Vol%201%20Paper%201%20-%20Household%20Survey%20Methodology%20and%20Fieldwork%202008.pdf

Sample Size Determination

The desired level of accuracy was set to a confidence level of 95% and an absolute precision (relative margin of error) of 5%. The population proportion (p) was set conservatively to 0.5, which yields the largest sample size. The minimum sample size (n) was determined by the following equation:²

$$n = \left(\frac{Z_a \sqrt{p(1-p)}}{C_p} \right)^2 = \left(\frac{1.96 \sqrt{0.5(1-0.5)}}{0.05} \right)^2 = 384$$

Where:

n = Minimum sample size

Z_a = Z-value for 0.05 level of significance

C_p = Confidence level

p = Population proportion

Inserting the parameters for the survey yields the minimum sample size for simple random sampling; therefore, for our sample design (stratified with multiple levels in some cases) the minimum sample size was multiplied by the design effect variable.

In the absence of empirical data from previous surveys that would have suggested a different value, the default value of 1.5 was used as the design effect for all countries (including Bangladesh), except India. A value of 2 was used for India to consider the additional levels of selection (district and sub-district) in the sample design. This yields then, a minimum sample size of 768 per country for households and individuals. The actual sample size for the Asian countries was increased beyond the minimum requirement to compensate for clustering effects, and allow for urban/rural disaggregation of data, as well as gender-based disaggregation. Therefore, in Bangladesh the sample size was increased to 2,000. In India, the sample size was further increased to 5,000.

² Rea, L and Parker, R. (2014), *Designing and Conducting Survey Research: A Comprehensive Guide*.

Sampling

Survey target groups

The target groups of the study in India and Bangladesh were:

- a) All households
- b) The population between the ages of 15-65

Survey respondents

In a selected household, two persons were interviewed:

1. The household head, as identified by the households themselves. If the household head was not available, a suitable alternative was sought out (e.g., spouse or other) who would be able to provide sufficient detail on the household characteristics and other members.
2. A randomly selected individual from the pool of eligible household members, i.e., any household member aged 15 through 65, which could even include the household head or alternative household representative.

Sample frame and PSU definition

INDIA

The sample size in India was 5,000 households and individuals. The primary sampling unit (PSU) in India was the ward (urban) and village (rural). The method was developed using data on household numbers in the wards and villages of India available in the National Primary Census Abstract Data. Random sampling was performed in three steps PSUs resulting in the selection of 250 PSUs across 19 states in India.

BANGLADESH

The sample size in Bangladesh was 2,000 households and individuals. The PSU in Bangladesh was the ward (urban) and village (rural). The method was developed using data on household numbers in the wards and villages of India available in the National Census Data Tables. Random sampling was performed in three steps PSUs resulting in the selection of 100 PSUs across 40 Zillas in Bangladesh.

Overview of sampling procedure

The AfterAccess methodology ensures national representation through random selection of all sampling units (ward/village, household, individual, etc.), enabled by household listing at the PSU level. Since the numbers of houses in wards and villages can vary significantly

(for instance, some wards in Mumbai can have as many as 100,000 households), wards and villages over a certain size had to be segmented. Thereafter, a segment(s) had to be randomly selected. Thereafter, the listing and mapping of households in the segment could be done, enabling random selection of households. These adjustments ensure that randomness is maintained, and data collected across survey countries are still comparable.

The AfterAccess methodology involves the following steps:

- (1) Separation of ward/village sample frame into urban and rural PSUs
- (2) Sampling the required number of ward/villages from each stratum (urban and rural PSUs) using probability proportionate to size (PPS).
 - (a) Where the ward/village size exceeded a critical number, wards/villages were segmented and a segment was selected randomly, in which the remaining sampling steps were continued with.
- (3) Mapping, listing and marking all households in the selected PSU or segment – the lists serve as the sample frame for simple random selections of households; this was done with the assistance of key informants (e.g., ward / village leader, etc.)
- (4) Simple random selection of the required number of households (20-25) from each selected PSU / segment
- (5) Listing all household member or visitor aged 15-65³ staying the night at the selected household
- (6) Simple random selection of one household member or visitor for survey from household list compiled (using the Kish grid) in Step 5

PSU selection

The selection of PSUs was conducted by LIRNEasia and provided to IPSOS.

INDIA

In India, PSUs were selected by LIRNEasia in the following way; selected PSUs are depicted in Figure A3:

- Step 1: A sample (108) of district were randomly selected from the total 640 districts of India using PPS sampling
- Step 2: A sample of sub-districts were randomly selected from selected districts using PPS sampling
- Step 3: The sub-districts were split into urban (wards) and rural (villages) strata and required number of wards (108) and villages (142) were selected from the selected sub-districts using PPS.

³ Aged 15+ in Africa and Latin America

In Bangladesh PSUs were selected by LIRNEasia in the following way; selected PSUs are depicted in Figure A4:

- Step 1: The National Census Data Tables were split into urban (wards) and rural (villages)
- Step 2: The required number of wards (40) and villages (60) for each stratum were selected using PPS

Segmentation, mapping and listing within sample PSUs

For selected PSUs, PSU-level household and population data were not readily available. In order to draw a random sample of households within a PSU, all households within a PSU had to be mapped and listed.

The objective of the mapping and listing of households was to ensure that the list of households covers the entire population living in the PSU. In this way, when the sample is selected from the list of dwellings or households, all eligible households in a PSU get equal chance of being selected for the survey.

This process is mainly used for creating a list of households from which a sample would be chosen for the main survey. This process provides a complete list of occupied residential households in a selected area. This helps in defining the complete range of households, making possible the employment of systematic and equal probability random selection of the households. As an advantage, this helps in circumventing the problem of biased household selection for the main survey.

Households are found in dwellings, dwellings are located in structures and structures in clusters (PSU). The listing operation consists of visiting each PSU, recording on listing forms a description of every structure together with the names of the heads of the households found in the structure, and drawing a location map of the PSU as well as a layout sketch map of all structures residing in the PSU.

These details help and guide interviewers to relocate selected households for interview during main survey data collection.

When arriving at a PSU, the listing team met with 2-3 key informants (KIs) to verify PSU details (e.g., numbers of households, locations of households, PSU boundaries and access routes, etc.). In rural areas, a KI was either one of the following: Sarpanch/ Pradhan/ ASHA/ AWW/ Panchayat member/ Village Secretary/ Senior resident/ school teachers etc. In urban areas, KIs were: PRI members/ Ward member/ party office / Municipality officials/ local service providers etc.

The estimates of the number of households in a PSU provided by the KIs were also checked with census 2011 data. If the number varied (more or less than 10%) the figure was validated

with another available KI in the PSU. Once validated, the final figure will be reported on CAPI device.

Segmentation, mapping and listing was conducted in advance by listing teams and the survey itself was thereafter conducted by the main survey teams.

MAPPING

STUDY OF ICT ACCESS AND USE: LIRNEasia Beyond Access Asia Phase I

LOCATION MAP

(a) State Name and code: (b) District Name and code:

(c) Sub-district / Tehsil Name and code: (d) Type of PSU Urban (1) / Rural (2):

(e) Block / Town Name and Code: (f) PSU (ward/village) name and code:

W E

NOTES: How to reach PSU?

Figure A1. Mapping template used in India and Bangladesh

Three types of maps were prepared during the course of the fieldwork:

1. A **location map** was prepared when the listing team reached a PSU, to depict the main access routes to a PSU, including main roads and main landmarks in the PSU; the purpose of this was to enable the main survey team to relocate the PSU at the time of survey. The mapping layout template is shown in Figure A1.
2. A **segment map** was prepared to divide the PSU into segments and selecting the required segment as per study requirement.
3. A **layout sketch** map was produced in household listing operations, with location or marks of all structures found in the listing operation to help the interviewer to relocate the selected households, as well as ensure that PSU households are not missed out. A sketch map also contains the segment identification information, location information, access information, principal physical features and landmarks such as mountains, rivers, roads and electric poles.

SEGMENTATION

If the PSU size was greater than 250, a segmentation map was prepared. The supervisor draws an indicative map showing the approximate location of the segments and then numbers them sequentially from the north-west corner. While drawing this map uninhabited areas and isolated households of the village were included as part of a segment.

For segmenting the PSU, the team used natural boundaries as they exist in the village in form of landmarks like streets, canals, rivers, pond, footpath, major or minor roads, schools, AWC, hamlets that may exist. For each segment, the estimated number of households were collected from KIs. The segment information (name and size) were entered onto the CAPI device and then the required number of segments were automatically selected by CAPI using rand function. Information of the selected segment was then made to appear on the device screen.

The method of segmentation and determination of the number of segments to select was as per the following protocol:

1. If the PSU size was less than 250 households, no segmentation done.
2. If the PSU size was between (250-350 households), made segments of size approximately 75 to 85 households and selected 3 segments randomly to ensure listing of around 220-250 households.
3. If the PSU size was between (350 – 1,000 households), made segments of size approximately 110 to 125 households and selected 2 segments randomly to ensure listing of around 220-250 households.
4. If the PSU size was between (1,000 – 5,000 households), made segments of size approximately 220 to 250 households and selected 1 segment randomly to ensure listing of around 220-250 households.
5. If the PSU size was more than 5,000 households,
 - a) First divided the PSU in four quadrants (4 zones i.e. *North-East, East-South, South-West and West-North*) and selected one quadrant randomly.
 - b) Get an estimate of number of households in selected quadrant
 - c) In the selected quadrant make segments (as explained above depending on size of quadrant) and select segments randomly to ensure listing of around 220-250 households.

LISTING

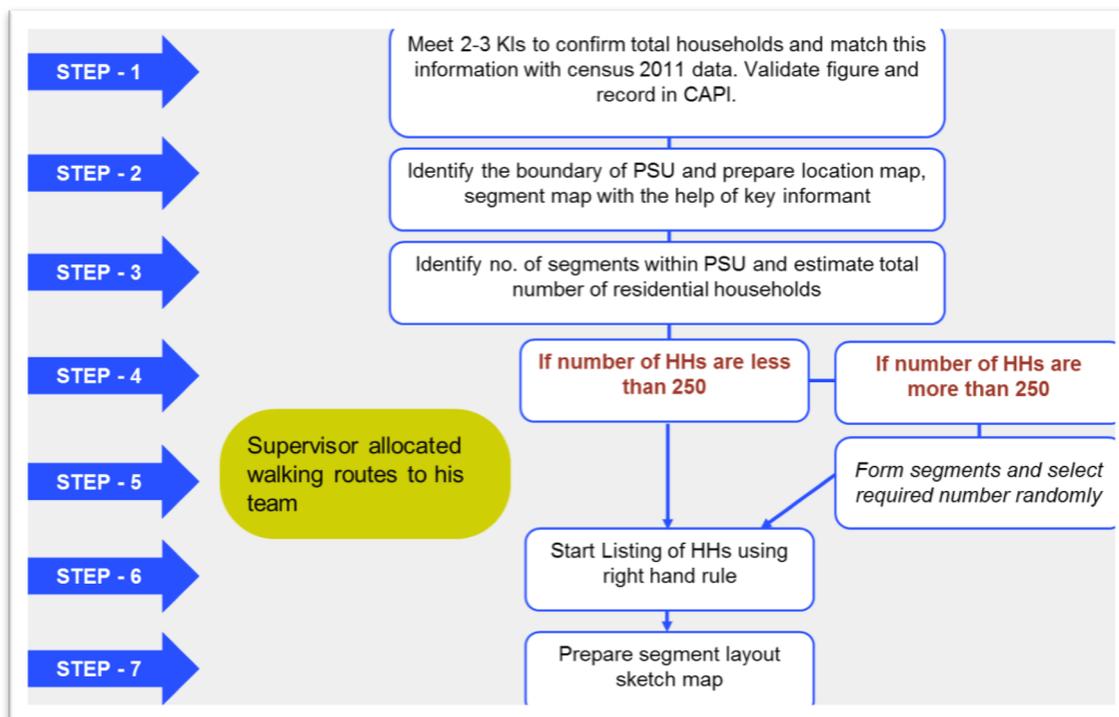


Figure A2. Listing procedure adopted in India and Bangladesh

In the selected segment/PSU all the structures and households had to be listed with the help of CAPI device as well as marked on the sketch maps to ensure no houses are missed out.

In the selected segment, the listing team supervisor allocated walking routes to his team (e.g. point AB; point CD, etc.). The lister/mapper entered the structure and household information in their CAPI device. Structure IDs used by the lister/mappers were given to correspond to their route and structure number. For instance, someone walking from A to B used structure IDs BA-AB001, BA-AB002 etc. (where in this case 'BA' refers to the internal project code – *Beyond Access*) so that there would be no duplication of structure IDs.

The lister / mapper marked the structure ID onto the gate or house wall with the consent from the household member using easily visible chalk or permanent marker.

The lister/mapper provided at least three to four structure information for each street on the map (starting-point, mid-point and end-point of street) as assigned by the supervisor.

Supervisors monitored the entire house listing/mapping operation and ensured that every structure in a segment / PSU had a structure ID. The supervisor with the help of the sketch maps prepared by the team prepared the layout sketch map of for that particular PSU/segment.

The listing procedure is depicted in Figure A2.

The listing team spent one day in each PSU to complete all the required work. The mapping of the PSU and the listing of the households was done in a systematic manner so that there are no omissions or duplications of households. If the PSU consists of several clusters, then the team finished each cluster before moving to the adjacent one. Within each cluster, the team worked using the right-hand rule.

The listing team were careful to locate hidden structures. In some areas, structures have been built so haphazardly that they can easily be missed out. If there is a pathway leading from the listed structure, team checked to see if the pathway goes to another structure. People living in the area helped in identifying any hidden structures.

It was difficult to pinpoint the exact location of the structure on the map; thus, an approximate location was used to help in finding the structure in the future.

If the household was temporary locked, information from neighbor for the locked house were taken

Team ascertained the boundaries of the sample PSU with the help of Key informants. This helped the team to understand the border area so that they should not list households of non-selected PSU.

If there was one structure with multiple households then all households were listed with same structure number

All non-residential structure was listed in the selected segment

Capturing names with surnames and landmark was most critical and teams were instructed to write full name instead of only first name like *Ram, Shyam, Anwar, Salim*, etc.

Household and individual selection

After listing of households, a frame of households was formed and 25 household in rural and 30 household in urban were randomly selected in each PSU/segment. The actual requirement was 20 but selected more considering replacement of households. The main survey team then had to identify the households, select the target respondents according to set procedures, and interview them.

IDENTIFICATION OF HOUSEHOLD DURING MAIN SURVEY

- Maps prepared by the listing team were used by the main survey team to reach the correct PSU / selected segment and identify the selected household for main survey.
- To identify the selected household, the address, landmark, name of head of household, structure number were displayed on CAPI screen.
- The maps prepared by listing mapping teams were used to locate the households, matching the structure number on the map to that on the wall of the household as well as the name of selected household.

Selection of household respondent during main survey

In a selected household, the household head was interviewed on household characteristics; if the household head was not available, a suitable alternative were sought out (e.g., spouse or other) who can provide sufficient detail on the household characteristics and other members.

Selection of individual respondent during main survey

In sample households, all eligible individuals (age 15-65 years) were listed with the help of the household head or representative and one was selected automatically (randomly) using the CAPI device. The CAPI device was programed to flash the selected individual's details (name, gender, age and relationship to the household head) on the screen and was not allowed to go back after this stage.

If the enumerator failed to reach the relevant respondent on first visit, the enumerator consulted with other members of the household/staff or neighbors about the time availability of the target respondent. In case the selected respondent was not available at the first visit, a minimum of three (3) attempts (first visit +two follow up) were made (at different times of day, or based on an appointment) to reach the target respondent.

Replacement of a selected respondent within a household was not allowed. If required, whole household were replaced by drawing additional sample from the frame.

Replacement procedures

Strict replacement procedures were set in place, to ensure the randomness of the sample at all levels was preserved. The circumstances under which replacements were allowed are detailed below.

INDIVIDUAL REPLACEMENT

If a target respondent could not be reached after a minimum of three (3) attempts were made (at different times of day, or based on an appointment) to reach the target respondent, or the target respondent refused to participate in the research, then the household had to be replaced, with a random selection from the list of houses (applying the pre-determined sampling interval). Replacement of respondents within the household (either household or individual respondents) was strictly not allowed.

HOUSEHOLD REPLACEMENT

A household was replaced in the following situations:

- A household head or suitable alternative refused to participate in the survey
- A household head or suitable alternative was unavailable after three (3) attempts to contact them

- A house had been found to be locked/vacant after three (3) attempts to contact the house were made

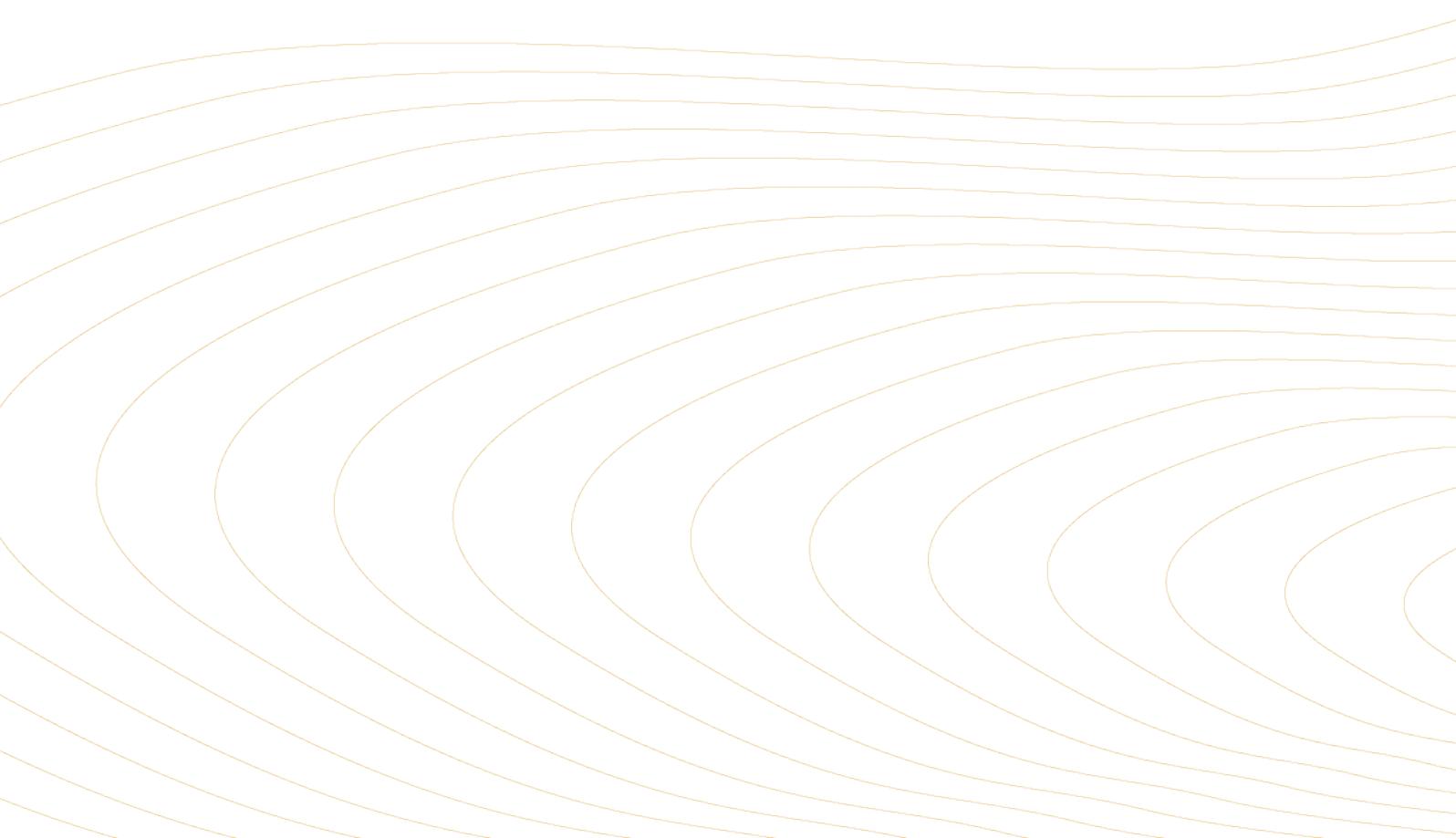
A household was replaced if the randomly **selected individuals** from a household either (*individual replacement in a household was not permitted*):

- Refused to be surveyed
- Was unwell or unable to communicate (e.g., due to sickness or disability)
- Could not be reached, after three (3) consecutive attempts to reach them were made by the enumerator.

If any of these situations arose, the house had to be replaced with a random selection from the list of houses in the PSU/segment (applying the pre-determined sampling interval).

PSU REPLACEMENT

In some cases it becomes impossible to conduct fieldwork in a selected PSU. For example, where security concerns prevent field teams from accessing the selected PSU, or local administrative leaders do not allow field teams to conduct research, or even entire villages may have migrated for a particular season. In these cases, the local fieldwork partner (IPSOS in India and Bangladesh) make requests for replacement PSUs. LIRNEasia provided random replacements, where the fieldwork partner then proceeded to carry out fieldwork in.



Research Instrument

Development, translation and scripting

A structured questionnaire was used to capture the household and individual level data. The English language questionnaire was developed by the global AfterAccess team (DIRSI, LIRNEasia and Research ICT Africa); country and regional localization was done by LIRNEasia with input from IPSOS also.

For India and Bangladesh, the questionnaire was translated into 11 languages by IPSOS, namely: Hindi, Tamil, Marathi, Gujrati, Bengali, Telegu, Assamese, Kannada, Malyalam, Meghalaya, and Oriya. LIRNEasia had the questionnaire translations checked over by external consultants to ensure accurate and context-appropriate translations.

The scripting of the questionnaire was done by IPSOS. IPSOS project teams tested the script for all logical and consistency checks before planning for pilot-test. The bilingual tools and script with login details were also shared with LIRNEasia for comments and feedback before pilot-test.

Pilot Testing

AIMS AND OBJECTIVES

Prior to the launch of the main survey, a pilot-test exercise was conducted by IPSOS. The aim of this exercise was to finalize questionnaire and to test the entire process of fieldwork starting from identifying the respondents till the final data collection to achieve a comprehensive and robust questionnaire.

Key objectives included:

1. Testing the implementation of the study methodology
 - Contacting target respondents
 - Interview with head of the household and obtaining consent
 - Selection of individual respondent randomly
 - Interview with individual respondent and taking consent.
2. Testing the research tools from the following perspectives:
 - Firming up areas of inquiry/indicators and redundancy of indicators
 - Ensuring instructions are clear and could be followed easily, by identifying:
 - Complicated/ difficult instructions
 - Unclear instructions
 - Missing instructions, if any
 - Testing out the survey questions to be administered, to
 - Identify any survey questions that are unclear, complex /difficult questions, sensitive or require rewording, which take too long to administer, etc.

- Ensure a logical flow of questions
 - Ensure routing instructions (e.g., skip patterns, etc.) are correct
 - Identify terminologies or concepts which are difficult for target respondents to comprehend
 - Ensure no duplication
 - Perform logical checks, etc.
 - Identify un-defined or long reference periods
 - Identify responses that may have been: incorrect, overlapping, missing, unclear, etc.
 - Ensure appropriateness of the language, terminologies
 - Identify scripting errors
3. Identify operational issues relating to:
- Permission from gatekeepers
 - Time it takes to administer single questionnaire
 - Productivity issues
 - Work load
 - Need of repeat and multiple visit
 - Replacement of household etc.

PILOT TEST CONDUCT

The pilot-test was conducted by the IPSOS team using tablet based questionnaire in the following languages:

- Hindi
- Tamil
- Bangla

The pilot-test exercise was performed by a team of field supervisors (4) and data collectors (16) who were experienced in conducting social surveys, particularly relating to technology and Internet. Two researchers from the IPSOS project team (core-team members) and three field executives in-charge of Delhi, Chennai and Bangladesh participated and took the active cognizance of the on-ground themes. Following was the pilot-test plan: -

Table A1. India pilot test locations

Area	Urban - Uttam Nagar, New Delhi	Rural - Arokye Medv, Tiruvallur District
Date	17-18 October, 2017	25-27 October, 2017
Day 1	Field Briefing in Delhi field office	Field Briefing in Chennai field office
Day 2	Pilot-test on Field (Urban)	Pilot-test on Field (Rural)

Table A2. Bangladesh pilot test locations

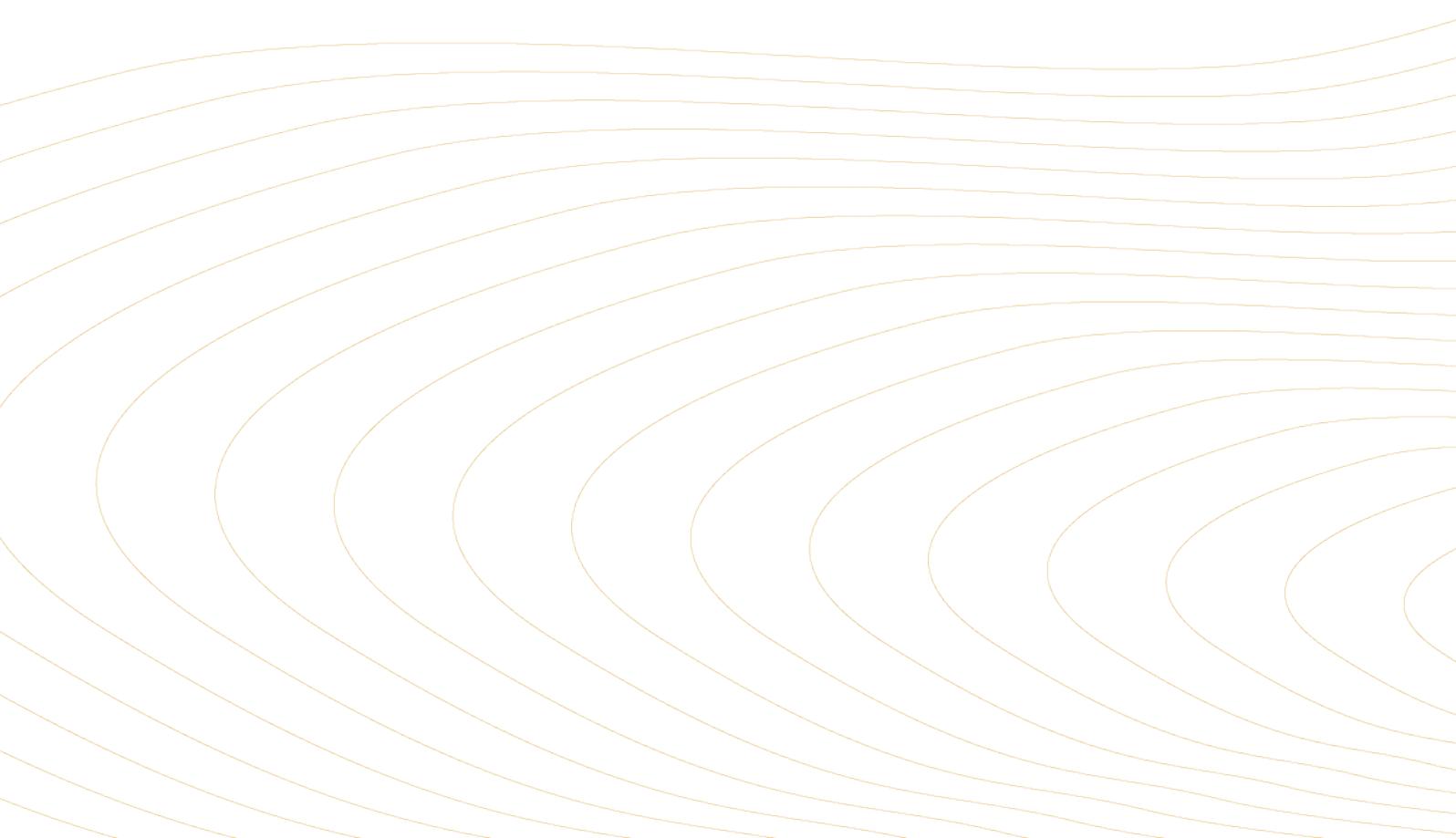
Area	Urban – Azimpur & Dhanmondi, Dhaka	Rural- Atibajar, Keraniganj
Date	31 October – 1 November, 2017	2-3 November, 2017
Day 1	Field Briefing in Dhaka field office	Field Briefing in Dhaka field office
Day 2	Pilot-test on Field (Urban)	Pilot-test on Field (Rural)

The pilot-test exercise was conducted in actual field setting. The pilot-tests were conducted in non-sample locations, and among a range of respondents to ensure those with different literacy levels and both genders were included.

A full day briefing (training) was conducted for data collectors prior to the pilot test. During the training, data collectors were briefed on process to be adopted for administering household head and Individual components of the interviews, replacement of household, need of repeat visits, permission from gatekeeper etc. The objectives of pilot-testing as mentioned earlier were briefed to data collectors and they were requested to make a note of their findings. The training was given by the core-team member of the project team.

The second day was utilized for actual pilot-test. All interviews were supervised by supervisors in the presence of core team member and executive in-charge. All team members including data collectors in the evening had debrief session for feedback and making pilot-testing notes.

Based on the observations made during the pilot tests, various improvements to the field process as well as the research tools were made.



Fieldwork Implementation

Coverage Maps

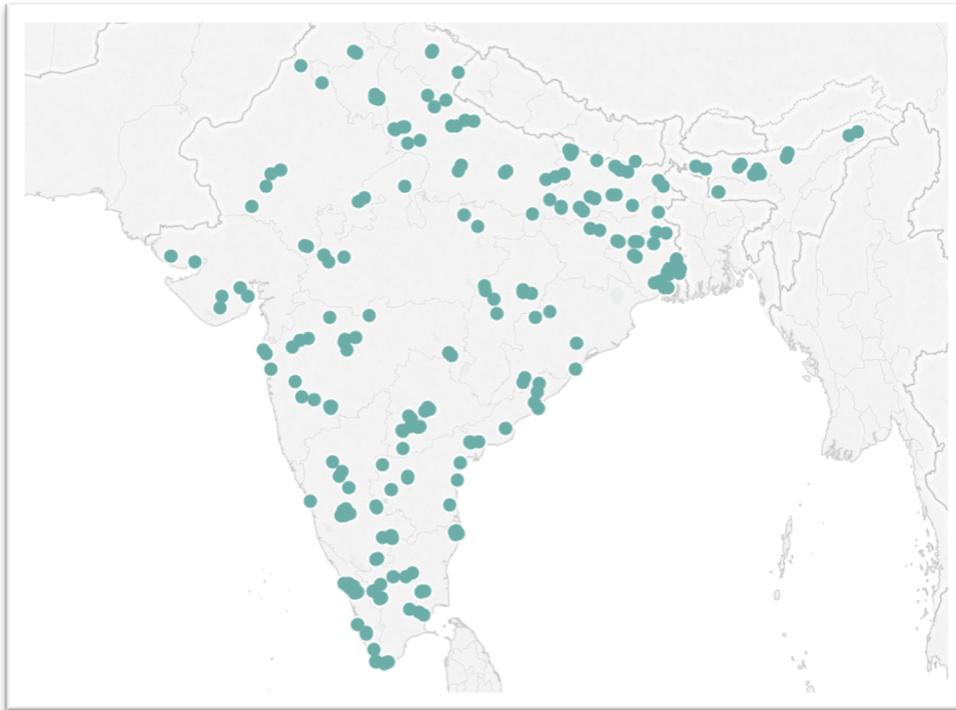


Figure A3. India sample locations

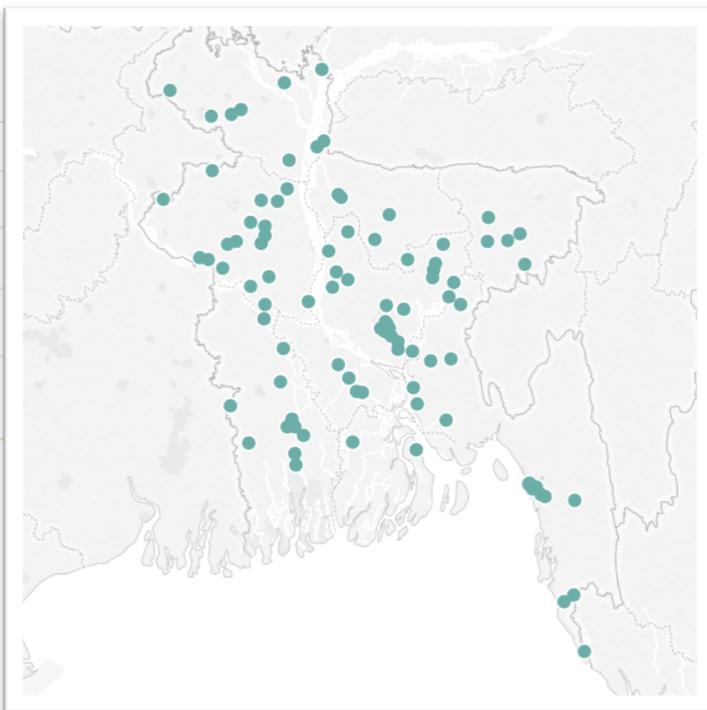


Figure A4. Bangladesh sample locations

Consent

Prior to the commencement of an interview, the respondent was informed of the following:

1. The objectives of the research
2. That his/her participation was voluntary
3. That he/she could choose to end the interview at any point

Once informed of these, the respondent's written or verbal consent was obtained from the respondent to (1) proceed with

the interview; (2) be photographed [if applicable]; and (3) be contacted again for further research. Respondents who refused to be photographed could still be interviewed if they agreed to participate.

Photography

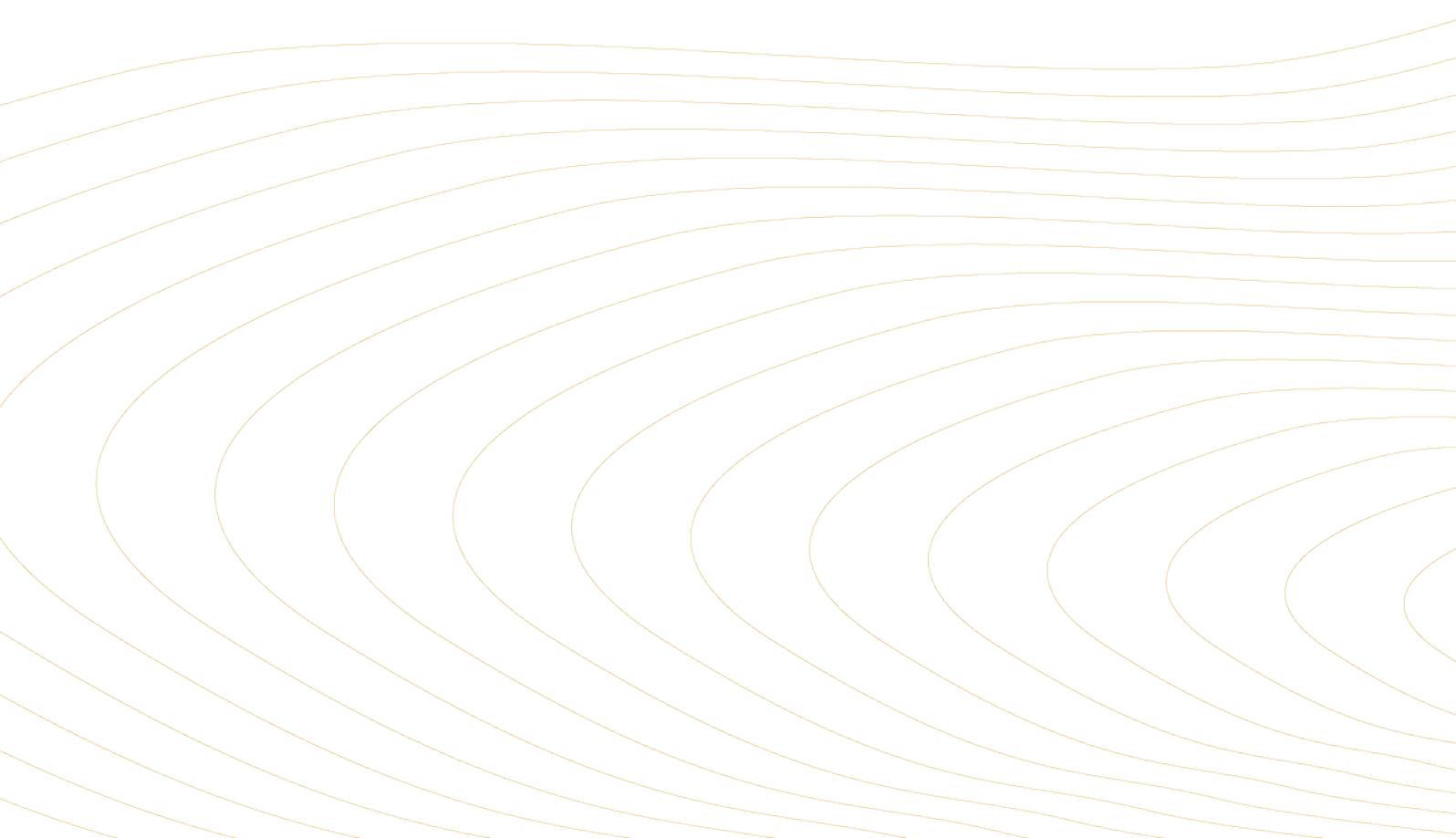
In every PSU, while working high resolution digital photographs were captured. Proper consent of the relevant community members was taken before taking photograph.

Two-three teams captured photographs during the listing / mapping exercise (structure number on walls, listing household, KI interview or permission) using CAPI devices

Photographs and scanned copies of the three types of maps (location, segmentation and layout sketch) were taken for documentation and verification purposes.

Collection of network quality data

GIS-tagged network quality metrics were collected in all sample locations by the Netradar application which was installed and running on all CAPI devices. The network quality data was automatically uploaded via the mobile network that the CAPI device was connected to, and collected from the back-end by LIRNEasia for analysis.



Quality control

Quality control by IPSOS

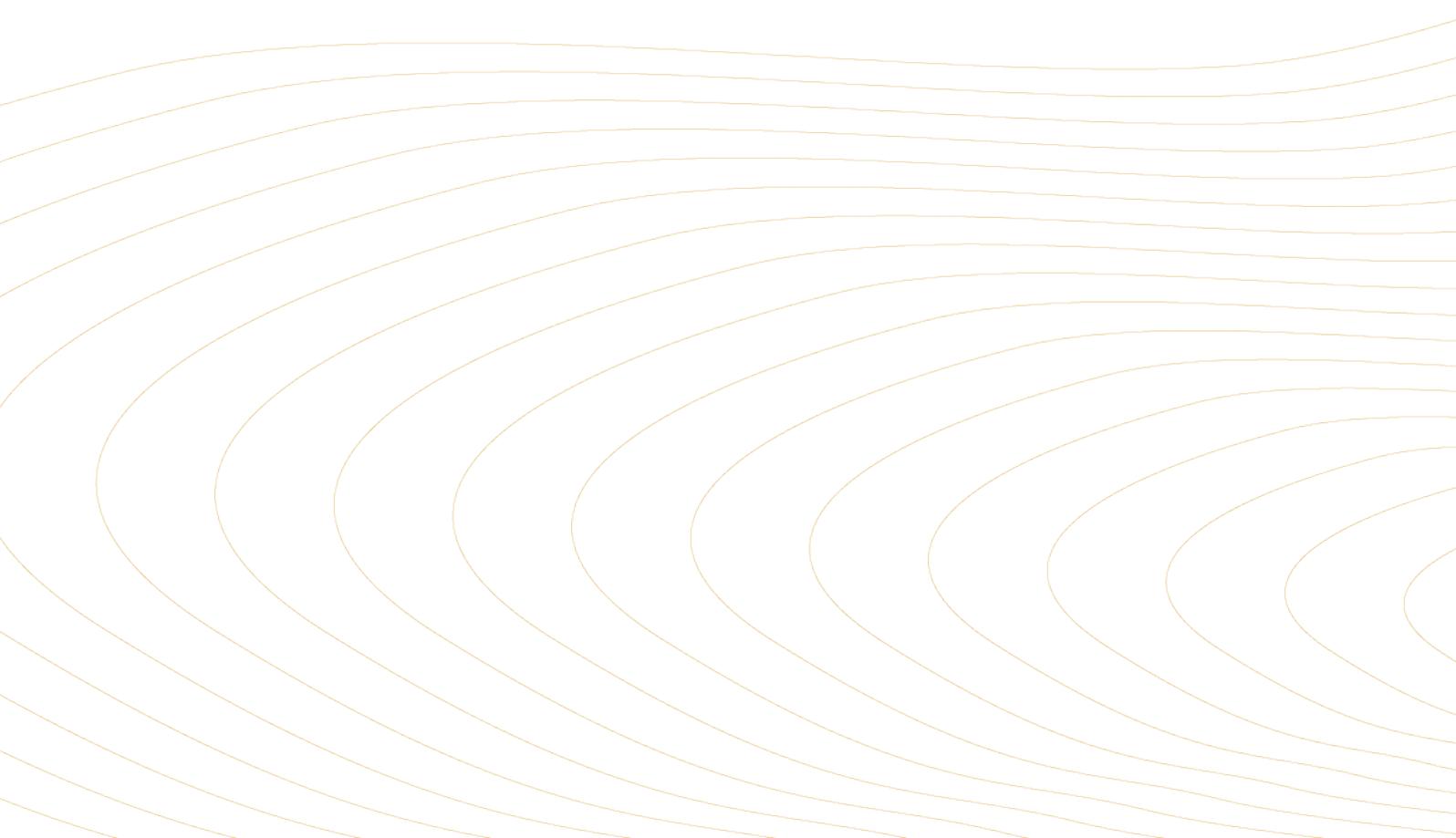
Ensuring collection of good quality - that is, complete, reliable and valid data was a priority and commitment. Quality assurance steps were taken at each stage of research to ensure high-quality of data generated and processed.

- Surprise Checks, in addition to Spot and Back-Checks by core team and enumerators in charge (EICs).
- Rigorous Monitoring by core team and EICs
- Back check by supervisors and quality assurance teams
- Checking of data for completeness and accuracy of recorded data by data processing team
- Coverage of 20 complete sample size with minimum refusals/shortfalls
- Completeness and correctness of information captured
- At the time when the data collectors are in the household, supervisors were instructed to do accompaniment check
- Supervisors were instructed to ensure that the interviewer behavior is appropriate and that the respondent is comfortable to respond throughout the interview
- The supervisors were instructed to not interrupt the interviewer during an interview; instead give feedback after completion
- Interviews were discarded in case there were fatal errors in the fieldwork. These fatal errors primarily include:
 - Replacement of individual respondent in a household (i.e., rather than replacing the household)
 - Non-eligible respondents interviewed
 - Incorrect household has been surveyed (discrepancies between GIS coordinates recorded at listing and interview of selected households exceeding 25 meters)
 - Interviews conducted in non-sample locations
- Replacement of discarded sample was performed by the teams doing sampling centrally (i.e., not at the field level)
- Recording of consent and checking of same by quality assurance team centrally
- CAPI were not allowed to go back once the individual respondent get selected
- Randomly recording of few questions and checking of the same by quality assurance team
- Automated GPS coordinates to get digital attendance of the survey team and survey location
- Obtaining Daily feedback to the team based on the feedback received from EICs, core team, DP team and quality assurance team

Quality checks by LIRNEasia

LIRNEasia conducted different types of quality checks during the fieldwork:

- **Monitoring the listing process in the field** – These monitoring exercises covered the end-to-end listing procedure of randomly selected PSUs. This includes meeting with KIs, Identifying the ward/village boundaries, segmentation of large villages and wards, random selection procedure of the blocks (segments), listing activity, marking of the structures etc.
- **HCC (Household count check)** – This was done after finalizing the listing by the IPSOS team. The LIRNEasia team followed the maps created by the IPSOS team after verifying the maps with the KIs from the PSU. The LIRNEasia team conducted a complete re-count of the households in some randomly selected PSUs.
- **In-house checks using GPS** – LIRNEasia's Big Data team conducted some quality checks using GPS coordinates captured during the listing exercise. The team visually inspected the GPS coordinates plotted on the Google maps. The purpose of these checks was to look at whether the listing is conducted in the correct PSU location. This check went down to the enumerator level.



Weighting

Weights were calculated and applied to the data collected in each country in order to gross up the data to the national level. Two weights were constructed, one for household and one for individuals. The weights are based on inverse selection probabilities. This method was used for weight the data in India, Bangladesh, Pakistan, Cambodia and Nepal. The weighting methodology was adjusted for India as per the design.

Household weight: $HH_W = DW \frac{1}{P_{AU} * P_{SEG} * P_{HH}}$

Individual weight: $IND_W = DW \frac{1}{P_{AU} * P_{SEG} * P_{HH} * P_I}$

Administrative unit (AU) selection probability: $P_{AU} = m \frac{HH_{GN}}{HH_{STRATA}}$

Segment selection probability: $P_{SEG} = \frac{HH_{SEG}}{HH_{WARD}}$

Household selection probability: $P_{HH} = \frac{n}{HH_{SEG}}$

Individual selection Probability: $P_I = \frac{1}{HH_{m15-65}}$

The administrative units in India and Bangladesh were the ward (urban) and village (rural).
The above-mentioned weights were applied for Bangladesh.

For India, two additional levels were introduced:

District Selection Probability: $P_{DIST} = d \frac{HH_{DIST}}{HH_{COUNTRY}}$

Sub-District Selection Probability: $P_{SUB-DIST UR} = sd \frac{HH_{SUB-DIST-UR}}{HH_{DIST-UR}}$

Hence the overall weights for India were as follows:

Household weight: $HH_W = DW \frac{1}{P_{DIST} * P_{SUB-DIST UR} * P_{AU} * P_{SEG} * P_{HH}}$

Individual weight: $IND_W = DW \frac{1}{P_{DIST} * P_{SUB-DIST UR} * P_{AU} * P_{SEG} * P_{HH} * P_I}$

When $HH_{AU} \leq 250$, $HH_{AU} = HH_{SEG}$.

Thus, $P_{SEG}=1$. (i.e.: if an AU has less than 250 households, it will be treated as a PSU)

DW = design weight compensation for over-sampling of urban PSUs and under-sampling of rural PSUs;

HH_{AU} = number of households in selected AU based on information of last census or updated listing by field team;

HH_{STRATA} = number of households in strata (urban, rural);

HH_{m15-65} = number of household members or visitors aged 15 -65 years;

m = target number of AUs for each strata, (urban, rural);

n = target number of households in a segment;

The target number of households in each PSU (AU/segment) was 20.