

# AfterAccess: ICT access and use in Sri Lanka, India, Pakistan, Bangladesh, Nepal and Cambodia

## Survey methodology note

By  
Tharaka Amarasinghe

LIRNEasia, 15 2/1, Balcombe Place, Colombo, Sri Lanka  
November 2023



LIRNEasia is a pro-poor, pro-market think tank whose mission is *catalyzing policy change through research to improve people's lives in the emerging Asia Pacific by facilitating their use of hard and soft infrastructures through the use of knowledge, information and technology.*

Contact: 15 2/1, Balcombe Place, Colombo 00800, Sri Lanka. +94 11 267 1160.  
info@lirneasia.net www.lirneasia.net

## **Table of Contents**

<b>1. Introduction.....</b>	<b>2</b>
Objectives of the study.....	2
<b>2. Survey target population and Sample size determination .....</b>	<b>4</b>
<b>3. Sampling .....</b>	<b>6</b>
Survey target groups .....	6
Survey respondents .....	6
Sample frame and Primary Sample unit (PSU) definition.....	6
Overview of the sampling procedure .....	7
PSU selection .....	8
Mapping and listing within sample PSUs.....	9
Mapping.....	10
Segmentation .....	10
Listing.....	11
Household and individual selection .....	12
Replacement procedures.....	13
Individual replacement.....	13
Household replacement.....	14
PSU replacement.....	14
<b>4. Research Instrument.....</b>	<b>16</b>
Development, translation, and scripting .....	16
Pilot Testing.....	16
<b>5. Consent .....</b>	<b>18</b>
Photography .....	18
<b>6. Quality Control (QC) Mechanisms .....</b>	<b>19</b>
Quality checks by LIRNEasia.....	20
<b>7. Weighting .....</b>	<b>21</b>

# 1. Introduction

The AfterAccess survey, conducted by LIRNEasia in Sri Lanka, India, Pakistan, Bangladesh, Nepal, and Cambodia from September 2017 to January 2019, comprised six nationally representative household and individual surveys. Additionally, two nationally representative surveys were conducted among persons with disabilities in Sri Lanka and Nepal, along with a nationally representative survey among small and medium-sized enterprises in Sri Lanka focusing on ICT access and use.

The AfterAccess methodology, developed by Research ICT Africa, was applied with local adjustments, considering the most granular sample frame available. Over the past decade, this sampling methodology has been applied in numerous countries in Africa<sup>1</sup> to attain nationally representative results for multiple target groups simultaneously in a cost-effective manner.

Further adjustments, detailed in the following sections, were implemented to balance the twin priorities of capturing the diversity of the population and managing fieldwork costs in the vast nations studied. Fieldwork in Asia was carried out by competitively procured market research companies, responsible for various aspects such as fieldwork set-up, scripting, translating, pilot testing the questionnaire, training enumerators, and dataset delivery. LIRNEasia played a monitoring role, being involved in field training and overseeing the fieldwork both on the ground and remotely.

**Table 1. Competitively procured market research companies for each country**

<b>Country</b>	<b>Competitively procured market research company</b>
<b>Sri Lanka</b>	Nielsen Lanka private limited
<b>India</b>	IPSOS India
<b>Pakistan</b>	The Dynamic research consultants
<b>Bangladesh</b>	IPSOS India
<b>Cambodia</b>	Kantar Cambodia
<b>Nepal</b>	Nielsen Nepal

## Objectives of the study

The key objective of the methodology was to ensure national representation at the desired levels of precision. The objectives of the AfterAccess Asia study were as follows:

1. To collect regionally comparable, nationally representative household and individual-level data on indicators of ICT access and use in all six countries.
2. To collect regionally comparable, nationally representative individual-level data on indicators of ICT access and use among persons with disabilities, limited to Sri Lanka and Nepal.

<sup>1</sup> [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](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)

3. To collect regionally comparable, nationally representative small and medium-sized enterprise-level data on indicators of ICT access and use, restricted to Sri Lanka.

The study aimed to collect regionally comparable household-level data on indicators of ICT access and use, enabling estimates to be made within a 95% confidence interval with the desired margin of errors, as will be explained in the next section.

## 2. Survey target population and 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.

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

Were,

n = Minimum sample size

Za= Z-value for 0.05 level of significance

Cp = 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, a value of 2 was used as the design effect to consider the additional levels of selection in the sample design. This yields then, a minimum sample size of 768 for households.

The actual sample size was increased beyond the minimum requirement to compensate for clustering effects and allow for the disaggregation of data into urban, rural, and estate categories, as well as gender-based disaggregation. This increase in sample size contributed to a reduction in the margin of errors for the estimates, as explained in table 2.

**Table 2. Target population, sample sizes and margin of errors of the surveys**

Country	Target population	Sample size (actual achieved)	Margin of Error at 95% Confidence Level (National Estimates)
<b>Sri Lanka</b>	• All households	2,017	±3.3%
<b>India</b>	• Age 15-65 population	5,069	±2.0%
<b>Pakistan</b>		2,002	±3.3%

<b>Bangladesh</b>		2,020	±3.3%
<b>Cambodia</b>		2,123	±3.3%
<b>Nepal</b>		2,008	±3.3%
<b>Sri Lanka</b>	<ul style="list-style-type: none"> <li>Persons with disabilities aged 15-65</li> </ul>	400	±9.8%
<b>Nepal</b>		400	±9.8%
<b>Sri Lanka</b>	<ul style="list-style-type: none"> <li>All small and medium sized enterprises</li> </ul>	402	±9.8%

This data was collected through face-to-face surveys conducted using Computer Assisted Personal Interviewing (CAPI) and according to the methodology outlined in this document.

## 3. Sampling

### Survey target groups

The target groups of the study were as follows,

Country	Target population
Sri Lanka	All households Age 15-65 population
India	
Pakistan	
Bangladesh	
Cambodia	
Nepal	
Sri Lanka	Persons with disabilities aged 15-65
Nepal	
Sri Lanka	All small and medium sized enterprises

### Survey respondents

**For household and individual surveys, two persons were interviewed in a selected household:**

1. The household head, as identified by the households themselves. If the household head was unavailable, a suitable alternative (e.g., spouse or other) was sought out to provide sufficient details on household characteristics and other members.
2. A randomly selected individual from the pool of eligible household members, aged 15 through 65, which could include the household head or an alternative household representative.

**For surveys among persons with disabilities:**

1. The household head, as identified by the households themselves. If unavailable, a suitable alternative was sought to provide sufficient details on household characteristics and other members.
2. A randomly selected individual with a disability from the pool of eligible household members. In cases where the person with a disability couldn't communicate effectively with the enumerator, a proxy was used.

### Sample frame and Primary Sample unit (PSU) definition

Different sample frames were used to select the primary sample units during the sampling process based on the availability of the sample frames.

**Table 3. Coverage and sample frame related information on 2017-2018 AfterAccess survey and 2021 COVID impact surveys**

Country	Coverage	Sample frame used	PSU	Level of representation
Sri Lanka	100 Grama Niladhari Divisions (GNDs) covering all nine provinces	GND-level data from the National Census of Population and Housing 2012	GND	National, and urban-rural level
India	250 wards and villages covering 19 states and 108 districts	Ward (urban) and village (rural) level data from the 2011 National Primary Census Abstract Data	Wards and Villages	National, and urban-rural level
Pakistan	100 Enumerator areas (EAs). The AJK, FATA and Gilgit-Baltistan provinces – amounting to approximately 2% of the were excluded from the sample frame due to practical and security considerations	2017 Census of Pakistan EA level sampling frame.	EA	National, and urban-rural level
Bangladesh	100 wards and villages covering 40 Zillas	Ward (urban) and village (rural) level data from the 2011 National Census Data	Wards and Villages	National, and urban-rural level
Cambodia	100 villages covering 25 provinces	village-level data from the 2014 intercensal survey	Villages	National, and urban-rural level
Nepal	100 wards covering 7 provinces	ward-level data from the National Population and Housing Census 2011 based on the new structure of 753 local units	Wards	National, and urban-rural level

## Overview of the sampling procedure

### Overview of the Sampling Procedure:

The AfterAccess methodology ensures national representation through the random selection of all sampling units (GNDs, EAs, ward/village, household, individual, etc.), facilitated by household listing at the PSU level. To account for variations in the number of houses in wards and villages, especially in larger urban areas (e.g., some wards in Mumbai with up to 100,000 households), wards and villages exceeding a certain size were segmented. Subsequently, a segment(s) was randomly selected, and the listing and mapping of households in the segment were conducted to enable the random selection of households. These adjustments ensure the maintenance of randomness, allowing for comparable data collection across surveyed countries.



The AfterAccess household and individual survey methodology involves the following steps:

1. Separation of PSU sample frame into urban and rural PSUs.
2. Sampling the required number of PSUs from each stratum (urban and rural PSUs) using probability proportionate to size (PPS).
  - a. Where the PSU size exceeded a critical number, wards/villages were segmented, and a segment was randomly selected. The remaining sampling steps were then continued within the selected segment.
3. Mapping, listing, and marking all households and structures in the selected PSU or segment. These lists serve as the sample frame for simple random selections of households and enterprises, assisted by key informants (e.g., ward/village leaders).
4. Simple random selection of the required number of households (20-25) from each selected PSU/segment.
5. Listing all household members or visitors aged 15-65 staying the night at the selected household.
6. Simple random selection of one household member or visitor for the survey from the household list compiled (using a random selection program embedded in the CAPI program).

For the AfterAccess persons with disability survey methodology, steps 4, 5, and 6 differed as follows:

4. Simple random selection of the required number of households with persons with disabilities (about 4) from each selected PSU/segment.
5. Listing all household members with disabilities aged 15-65 staying the night at the selected household.
6. Simple random selection of one household member with disabilities from the household list compiled (using a random selection program embedded in the CAPI program) in Step 5.

For the AfterAccess small and medium-sized enterprises survey methodology, steps 4, 5, and 6 differed as follows:

4. Simple random selection of the required number of enterprises (about 4) from each selected PSU/segment.
5. Conducting the interview with the enterprise owner or manager.

## PSU selection

LIRNEasia conducted the selection of Primary Sampling Units (PSUs) and provided the chosen PSUs to the respective fieldwork vendors for execution. A predetermined number of PSUs were selected using Probability Proportionate to Size (PPS) sampling techniques.

## Mapping and listing within sample PSUs.

For the selected PSUs, household and population data at the PSU level were not readily available. To draw a random sample of households within a PSU, mapping and listing of all households and structures within the PSU were essential.

The objective of mapping and listing structures was to ensure that the list of households covered the entire population residing in the PSU. This approach ensured that when the sample was selected from the list of dwellings or households, all eligible households in a PSU had an equal chance of being chosen for the survey. This process created a comprehensive list of occupied residential households in a selected area, allowing for systematic and equal probability random selection of households and mitigating the issue of biased household selection.

For the persons with disability survey, the availability of persons with disabilities was also checked and recorded during the listing. Similarly, small and medium-sized enterprises were marked in the listing.

Households are situated in dwellings, dwellings are located in structures, and structures are grouped in clusters (PSUs). The listing operation involved visiting each PSU, recording on listing forms a description of every structure along with the names of the heads of households found in the structure, and creating location and layout sketch maps of the PSU.

These details served as a guide for interviewers to locate selected households during the main survey data collection. Upon arriving at a PSU, the listing team met with 2-3 key informants (KIs) to verify PSU details such as the number of households, locations of households, PSU boundaries, access routes, etc. In rural areas, KIs included Grama Niladhari officers, Sarpanch/Pradhan/ASHA/AWW/Panchayat member/Village Secretary/Senior resident/school teachers, while in urban areas, KIs were PRI members/Ward members/party office/Municipality officials/local service providers.

The estimates of the number of households in a PSU provided by the KIs were cross-checked with the relevant latest census data for each country. If the numbers varied (more or less than 10%), the figure was validated with another available KI in the PSU. Once validated, the final figure was reported on the CAPI device.

Segmentation, mapping, and listing were conducted in advance by listing teams, and the main survey was subsequently carried out by the main survey teams.

## Mapping

LIRNEasia Beyond Access: Asia Phase I


**STUDY OF ICT ACCESS AND USE:**

**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:



NOTES: How to reach PSU?

**Figure 1. A mapping template used in India and Bangladesh**

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

1. A location map was created upon reaching a PSU, illustrating main access routes, main roads, and landmarks in the PSU. This map aimed to assist the main survey team in relocating the PSU during the survey. The mapping layout template is shown in Figure 1.
2. A segment map was prepared to divide the PSU into segments and select the required segment as per the study's needs.
3. A layout sketch map was produced during household listing operations, marking the location of all structures to aid interviewers in relocating selected households. This map also contained segment identification information, location details, access information, and principal physical features and landmarks.

## Segmentation

If the PSU size exceeded 250, a segmentation map was prepared. The supervisor drew an indicative map showing the approximate location of segments and numbered them sequentially from the northwest corner. Uninhabited areas and isolated households in the village were included in a segment. The team used natural boundaries such as streets, canals, rivers, ponds, footpaths, major or minor roads, schools, and hamlets to segment the PSU. The estimated number of households in each segment was collected from Key Informants (KIs), and the required number of segments was selected using a random function in CAPI.

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

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 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 Sarath, Ram, Shyam, Anwar, Salim, etc.

For persons with disability surveys in Sri Lanka and Nepal, a few questions were asked from each household to identify the households with persons with disability residing in households or not. Similarly in Sri Lanka, the listing questionnaire consisted of questions to identify the small and medium sized enterprises.

## Household and individual selection

### **Household and individual survey:**

After listing of households and structures, a frame of households was formed and about 25 household 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.

### **Persons with disability survey:**

After the listing of households and structures, a frame of households with persons with disabilities was established, and about four households were selected in each PSU for the main survey.

### **Small and medium sized enterprise survey:**

Following the listing of households and structures, a frame of enterprises was generated, and approximately four enterprises were selected for the main survey.

### 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 respondents during main survey

### **Household and individual survey:**

In a selected household, the interview focused on the household head regarding household characteristics. If the household head was unavailable, an alternative, such as the spouse or another suitable member, was sought to provide necessary details about the household and its members. In sample households, all eligible individuals aged 15-65 were listed, and one was automatically (randomly) selected using the CAPI device. The CAPI device was programmed to display the selected individual's details (name, gender, age, and relationship to the household head) on the screen, and no modifications were allowed after this stage.

### **Persons with Disability Survey:**

Similar to the household and individual survey, in a selected household, the interview focused on the household head regarding household characteristics. If the household head was unavailable, a suitable alternative, such as the spouse or another member, was sought. In sample households, all eligible persons with disabilities were listed, and one person with a disability was randomly selected. There were a few rare cases of multiple disabled individuals in the same household.

### **Small and Medium-Sized Enterprises Survey:**

In a selected small and medium-sized enterprise, the interview was conducted with the enterprise owner or manager as the respondent.

## 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
- Disabled respondent was unable to communicate

A similar replacement procedure was used for the small and medium sized enterprises as well.

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).

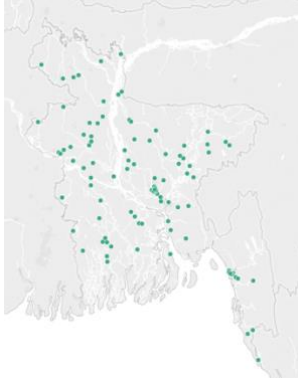
### PSU replacement

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 make requests for replacement PSUs. LIRNEasia provided random replacements, where the fieldwork partner then proceeded to carry out fieldwork in.





Sri Lanka



India



Pakistan



Bangladesh

Cambodia

Nepal

**Figure 2. 2017-18 AfterAccess survey sample locations based on GPS coordinates recorded during fieldwork.**



## 4. 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 local fieldwork vendors also.

The survey questionnaire was translated into local languages by relevant fieldwork vendors. 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 the local fieldwork vendors. Fieldwork vendor's 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

Prior to the launch of the main survey, a pilot-test exercise was conducted by local fieldwork vendors. 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.

Separate pilot tests were carried out for persons with disability survey, and enterprise survey.

Key objectives included:

1. Testing the implementation of the study methodology
  - a. Contacting target respondents
  - b. Interview with head of the household and obtaining consent (enterprise owner or manager for the enterprise survey)
  - c. Selection of individual respondent randomly and person with disability responded
  - d. Interview with individual respondent and taking consent.
2. Testing the research tools from the following perspectives:
  - a. Firming up areas of inquiry/indicators and redundancy of indicators
  - b. Ensuring instructions are clear and could be followed easily, by identifying:
    - i. Complicated/ difficult instructions
    - ii. Unclear instructions
    - iii. Missing instructions, if any
    - iv. 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.
    - v. Ensure a logical flow of questions

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

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.

### Pilot Goals

The goal of the pilot survey included:

- Estimate realistic median Length of Interview (LOI).
- Test Skipping and Routing.
- Test Language (phrasing).
- Test Translation.
- Test Understanding of the Questions, Cognitive Difficulties and Questions Sensitivities.
- Test Tablet Functionality.

## 5. 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

## 6. Quality Control (QC) Mechanisms

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 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.

## 7. 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 Sri Lanka, India, Pakistan , Bangladesh, Cambodia and Nepal. The weighting methodology was adjusted for India as per the design

### Sri Lanka

Household weight:

$$HH_W = DW \frac{1}{P_{GN} * P_{SEG} * P_{HH}}$$

Individual weight:

$$IND_W = DW \frac{1}{P_{GN} * P_{SEG} * P_{HH} * P_I}$$

Disabled individual weight:

$$DIS\_IND_W = DW \frac{1}{P_{GN} * P_{SEG} * P_{DIS\_HH\_SEG} * P_{DIS\_HH} * P_{DIS\_IND}}$$

Enterprise weight:

$$ENT_W = DW \frac{1}{P_{GN} * P_{SEG} * P_{ENT}}$$

GN selection probability:

$$P_{GN} = m \frac{HH_{GN}}{HH_{STRATA}}$$

Segment selection probability:

$$P_{SEG} = \frac{HH_{SEG}}{HH_{WARD}}$$

Household selection probability:

$$P_{HH} = \frac{n}{HH_{SEG}}$$

Probability of household with disabled individual in a selected segment:

$$P_{DIS\_HH\_SEG} = \frac{HH_{DIS}}{HH_{SEG}}$$

Household with disabled individual selection probability:

$$P_{DIS\_HH} = \frac{q}{HH_{DIS}}$$

Enterprise selection probability:

$$P_{ENT} = \frac{r}{ENT_{SEG}}$$

Individual selection Probability:

$$P_I = \frac{1}{HH_{m15+}}$$

Disabled Individual selection Probability:

$$P_{DIS\_IND} = \frac{1}{HH_{m15+}}$$

When  $HH_{GN} \leq 250$ ,  $HH_{GN} = HH_{SEG}$ . Thus,  $P_{SEG}=1$ . (i.e., if a GN 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<sub>GN</sub>**= number of households in selected GN based on information of last census or updated listing by field team;

**HH<sub>STRATA</sub>**=number of households in strata (urban, rural);

**HH<sub>DIS</sub>**=number of households with disabled individuals in a segment;

**ENT<sub>SEG</sub>** =number of medium sized enterprises in a segment;

**HH<sub>m15+</sub>**=number of household members or visitors aged 15 -65 years;

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

n = target number of households in a segment;

q = target number of households with disabled individuals in a segment

r = target number of medium-sized enterprises in a segment

The target number of households in each PSU (GN/segment) for the national sample will be 20 and for the booster sample will be 4.

## India

Household weight:

$$HH_W = DW \frac{1}{P_{DIST} * P_{SUB-DISTUR} * P_{VILL/WARD} * P_{EA} * P_{HH}}$$

Individual weight:

$$IND_W = DW \frac{1}{P_{DIST} * P_{SUB-DISTUR} * P_{VILL/WARD} * P_{EA} * P_{HH} * P_I}$$

District Selection Probability:

$$P_{DIST} = d \frac{HH_{DIST}}{HH_{COUNTRY}}$$

Sub-District Selection Probability:

$$P_{SUB-DIST-UR} = sd \frac{HH_{SUB-DIST-UR}}{HH_{DIST-UR}}$$

Village/Ward Selection Probability:

$$P_{VILL/WARD} = m \frac{HH_{VILL/WARD}}{HH_{SUB-DIST-UR}}$$

EA selection Probability:

$$P_{EA} = e \frac{HH_{EA}}{HH_{VILL/WARD}}$$

Household Selection Probability:

$$P_{HH} = \frac{n}{HH_{EA}}$$

Individual selection Probability:

$$P_I = \frac{1}{HH_{m15+}}$$

When  $HH_{VILL/WARD} \leq 250$ ,  $HH_{VILL/WARD} = HH_{EA}$ . Thus,  $P_{EA}=1$ . (i.e., If a village or a ward has less than 250 households, it will be treated as an EA)

$HH_{COUNTRY}$  = number of households in the country;

$HH_{DIST}$  = number of households in a district;

$HH_{DIST-UR}$  = number of households in a district at urban rural level;

$HH_{SUB-DIST-UR}$  = number of households in strata (Sub districts at urban, rural level);

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

$HH_{VILL/WARD}$  = number of households in selected EA based on information of last census or updated listing by field team;

$HH_{m15+}$  = number of household members or visitors 15 years or older;

$d$  = target number of districts;

$sd$  = target number of sub-districts;

$m$  = target number of Villages/Wards for each strata, (urban, rural);

$e$  = target number of EAs in a Village/Ward;

$n$  = target number of households in an EA;

$q$  = target number of businesses in an EA;

The target number of households in each Village/Ward will be 20.



## Pakistan

Household weight:

$$HH_W = DW \frac{1}{P_{EA} * P_{HH}}$$

Individual weight:

$$IND_W = DW \frac{1}{P_{EA} * P_{HH} * P_I}$$

EA Selection Probability:

$$P_{EA} = m \frac{HH_{EA}}{HH_{STRATA}}$$

Household Selection Probability:

$$P_{HH} = \frac{n}{HH_{EA}}$$

Individual selection Probability:

$$P_I = \frac{1}{HH_{m15-65}}$$

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

**HH<sub>EA</sub>**= number of households in selected EA based on information of last census or updated listing by field team;

**HH<sub>STRATA</sub>**=number of households in strata (urban, rural);

**HH<sub>m15-65</sub>**=number of household members or visitors 15-65 years;

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

n = target number of households in a EA;

The target number of households in each EA will be 20.

## Bangladesh

Household weight:

$$HH_W = DW \frac{1}{P_{VILL/WARD} * P_{EA} * P_{HH}}$$

Individual weight:

$$IND_W = DW \frac{1}{P_{VILL/WARD} * P_{EA} * P_{HH} * P_I}$$

Village/Ward Selection Probability:

$$P_{VILL/WARD} = m \frac{HH_{VILL/WARD}}{HH_{STRATA}}$$

EA selection Probability:

$$P_{EA} = e \frac{HH_{EA}}{HH_{VILL/WARD}}$$

Household Selection Probability:

$$P_{HH} = \frac{n}{HH_{VILL/WARD}}$$

Individual selection Probability:

$$P_I = \frac{1}{HH_{m15+}}$$

When  $HH_{VILL/WARD} \leq 250$ ,  $HH_{VILL/WARD} = HH_{EA}$ . Thus,  $P_{EA}=1$ . (i.e., If a village or a ward has less than 250 households, it will be treated as an EA)

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

**HH<sub>VILL/WARD</sub>**= number of households in selected EA based on information of last census or updated listing by field team;

**HH<sub>STRATA</sub>**=number of households in strata (urban, rural);

**HH<sub>m15+</sub>**=number of household members or visitors 15 years or older;

**m** = target number of Villages/Wards for each strata, (urban, rural);

**e** = target number of EAs in a Village/Ward;

**n** = target number of households in a EA;

**q** = target number of businesses in an EA;

The target number of households in each Village/Ward will be 20.

## Cambodia

Household weight:

$$HH_W = DW \frac{1}{P_{VILL} * P_{EA} * P_{HH}}$$

Individual weight:

$$IND_W = DW \frac{1}{P_{VILL} * P_{EA} * P_{HH} * P_I}$$

Village/Ward Selection Probability:

$$P_{VILL/WARD} = m \frac{HH_{VILL}}{HH_{STRATA}}$$

EA selection Probability:

$$P_{EA} = e \frac{HH_{EA}}{HH_{VILL}}$$

Household Selection Probability:

$$P_{HH} = \frac{n}{HH_{VILL}}$$

Individual selection Probability:

$$P_I = \frac{1}{HH_{m15+}}$$

When  $HH_{VILL} \leq 250$ ,  $HH_{VILL} = HH_{EA}$ . Thus,  $P_{EA}=1$ . (i.e., If a village or a ward has less than 250 households, it will be treated as an EA)

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

**HH<sub>VILL</sub>**= number of households in selected EA based on information of last census or updated listing by field team;

**HH<sub>STRATA</sub>**=number of households in strata (urban, rural);

**HH<sub>m15+</sub>**=number of household members or visitors 15 years or older;

$m$  = target number of Villages/Wards for each strata, (urban, rural);

$e$  = target number of EAs in a Village/Ward;

$n$  = target number of households in a EA;

$q$  = target number of businesses in an EA;

The target number of households in each Village/Ward will be 20.

## Nepal

Household weight:

$$HH_W = DW \frac{1}{P_{WARD} * P_{SEG} * P_{HH}}$$

Individual weight:

$$IND_W = DW \frac{1}{P_{WARD} * P_{SEG} * P_{HH} * P_I}$$

Disabled individual weight:

$$DIS\_IND_W = DW \frac{1}{P_{WARD} * P_{SEG} * P_{DIS\_HH\_SEG} * P_{DIS\_HH} * P_{DIS\_IND}}$$

Ward Selection Probability:

$$P_{WARD} = m \frac{HH_{WARD}}{HH_{STRATA}}$$

Segment selection Probability:

$$P_{SEG} = \frac{HH_{SEG}}{HH_{WARD}}$$

Household Selection Probability:

$$P_{HH} = \frac{n}{HH_{SEG}}$$

Probability of Household with Disabled individual in a selected segment:

$$P_{DIS\_HH\_SEG} = \frac{HH_{DIS}}{HH_{SEG}}$$

Household with Disabled individuals Selection Probability:

$$P_{DIS\_HH} = \frac{q}{HH_{DIS}}$$

Individual selection Probability:

$$P_I = \frac{1}{HH_{m15+}}$$

Disabled Individual selection Probability:

$$P_{DIS\_IND} = \frac{1}{HH_{m15+}}$$

When  $HH_{WARD} \leq 250$ ,  $HH_{WARD} = HH_{SEG}$ . Thus,  $P_{SEG}=1$ . (i.e., if a ward has less than 250 households, it will be treated as an EA)

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

**HH<sub>WARD</sub>**= number of households in selected ward based on information of last census or updated listing by field team;

**HH<sub>STRATA</sub>**=number of households in strata (urban, rural);

**HH<sub>DIS</sub>**=number of households with disabled individuals in a segment;

**HH<sub>m15+</sub>**=number of household members or visitors aged 15 -65 years;

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

n = target number of households in a segment;

q = target number of households with disabled individuals in a segment

The target number of households in each PSU (ward/segment) for the national sample will be 20 and for the booster sample will be 4.