STEP Survey Weighting Procedures Summary
(Based on The World Bank Weight Requirement)

Sri Lanka

October 15, 2013

Table of Contents

1   Survey Design Overview

The Sri Lanka survey firm implemented a partial literacy assessment design. The partial assessment required each selected person to attempt to complete a General Booklet comprising Reading Components and a set of Core Literacy Items. The partial assessment sampling objective was to have a minimum of about 2000 selected persons attempt the General Booklet.

The target population for the Sri Lanka STEP survey comprised all non-institutionalized persons 15 to 64 years of age (inclusive) living in private dwellings in urban and rural areas of **Sri Lanka** at the time of data collection.

Exclusions
The target population excludes,
- foreign diplomats and non-nationals working for international organizations;
- people in institutions such as hospitals or prisons;
- collective dwellings or group quarters;
- persons living outside the country at the time of data collection, e.g., students at foreign universities;
- persons who are unable to complete the STEP assessment due to a physical or mental condition, e.g., visual impairment or paralysis.

The sample frame for the selection of first stage sample units was the Census 2011/12

Table 1, based on a table of 2012 Census data provided by the Sri Lanka Department of Census & Statistics (DCS), shows the sample frame count of the number of first stage sample units, i.e., Grama Niladari (GN) divisions, the number of households, and the population size.

Table 1: Sample Frame Counts from Sri Lanka 2012 Census Frame of GN Divisions			
Stratum	GN Divisions	Households	Population (All Ages)
Urban	1,285	1,000,625	3,707,190
Rural	12,866	4,773,053	16,703,955
Total	14,151	5,773,678	20,411,145

The Sri Lanka sample design is a 5 stage stratified sample design. The stratification variable is Urban-Rural indicator.

First Stage Sample
The primary sample unit (PSU) is a Grama Niladari (GN) division.

The sampling objective was to conduct interviews in 200 GNs, consisting of 80 urban GNs and 120 rural GNs. Because there was some concern that it might not be possible to conduct any interviews in some initially selected GNs (e.g. due to war, conflict, or inaccessibility, for some other reason), the sampling strategy also called for the selection of 60 extra GNs (i.e., 24 urban GNs and 36 rural GNs) to be held in

reserve for such eventualities. Hence, a total of 260 GNs were selected, consisting of 200 'initial' GNs and 60 'reserve' GNs.

Two GNS from the initial sample of GNs were not accessible and reserve sampled GNs were used instead. Thus a total of 202 GNs were activated for data collection, and interviews were conducted in 200 GNs.

The sample frame for the selection of first stage sample units was the list of GNs from the Census 2011/12.

Note: The sample of first stage sample units was selected by the Sri Lanka Department of Census & Statistics (DCS) and provided to the World Bank. The DCS selected the GNs with probability proportional to size (PPS), where the measure of size was the number of dwellings in a GN.

Second Stage Sample
The second stage sample unit (SSU) is a GN segment, i.e., GN BLOCK. One GN Block was selected from each activated PSU (i.e., GN).

According to the Sri Lanka survey firm, each sampled GN was divided into a number of segments, i.e., GN Blocks, with approximately the same number of households, and one GN Block was selected from each sampled GN.

Third Stage Sample
The third stage sample unit is a dwelling. The sampling objective was to obtain interviews at 15 dwellings within each selected SSU.

Fourth Stage Sample
The fourth stage sample unit is a household. The sampling objective was to select one household within each selected third stage dwelling.

Fifth Stage Sample
The fourth stage sample unit is an individual aged 15-64 (inclusive). The sampling objective was to select one individual with equal probability from each selected household.

The summary of the Sri Lanka sample distribution is shown in Table 2.

Table 2: Sri Lanka Sampled Cases by Stratum & Sample Type

Stratum	Activated GNs	Number of Sampled Households		
		Initial Sample	Reserve Sample	Total
Urban	80	1200	1200	**2400**
Rural	120	1800	1800	**3600**
Total	**200**	**3000**	**3000**	**6000**

2   Data Collection Outcomes

The final survey weights depend on the outcome from the attempt to obtain an interview at each selected household. The treatment of each sampled case in the weighting process depends on the final status code assigned to the sampled household and the final status code assigned to the selected person. The final status codes used in the Sri Lanka STEP are listed in Table 3. A variable called '**Fin_Stat1**' with these code values was created during the weighting process and a '**FinStat1**' code was assigned to each sampled case.

Table 3: Result Codes for Sampled Households and Selected Persons	
I. Household-level Codes	
Code	**Code Description**
01	Household Refusal
02	Household unusual circumstance (e.g., death in family, illness, fire in dwelling, etc.)
03	No knowledgeable household member found after 3 visits (only a child, non-competent adult, etc.)
04	Temporarily absent/unavailable for field period (information from other persons)
05	No competent household member to interview (because of severe illness, mental disability, etc.)
06	Language problem – no household member spoke a language understandable by survey team, no translator available
07	Dwelling could not be found / selected address has no household
08	Vacant dwelling
09	Dwelling not habitable/dwelling destroyed/dwelling converted to commercial use
10	Household moved away
21	Household Refusal to continue Household Module (i.e., begun but not complete)
22	Household unusual circumstance (e.g., death in family, illness, fire in dwelling, etc.) prevented continuance
31	No household member in the eligible range of 15-64
94	Case dropped due to data quality issue
95	Sample case justifiably not activated, e.g., reserve sample household not needed , sample goal achieved
96	General non-response; Reason unspecified by survey firm
97	Sample selection error – a wrong household was selected in the field
II. Person-level Codes	
Code	**Code Description**
	Individual modules 2-7 not begun
32	Refusal by Selected Person to begin modules 2-7
33	Unusual circumstance - Selected Person (e.g., death in family, illness, fire in dwelling, etc.)
34	Refusal - Another household member refused to let selected individual do individual modules
35	Unable to contact Selected Person after three visits during field period
36	Temporarily absence/unavailability of Selected Person for field period (info from household member)
37	Hearing Problem - Selected Person is deaf or so hard-of-hearing, no translator available
38	Learning or mental disability (includes emotional conditions like severe depression) of Selected Person
39	Speech impairment of Selected Person prevented participation, no translator available
40	Language problem – Selected Person's language not understood by survey team, no translator available
98	Sample selection error – an incorrect eligible person was selected
99	Sample selection error – an ineligible person was selected
	Individual modules 2-7 begun but not completed
51	Refusal by Selected Person to continue modules 2-7
52	Unusual circumstance - Selected Person
	General Booklet (module 9) not begun
61	Unable to read or write in language of General Booklet – Selected Person refused to begin
62	Refusal by Selected Person to begin General Booklet
63	Unusual circumstance - Selected Person
64	Blindness or visual impairment of Selected Person prevented completion of General Booklet

Table 3: Result Codes for Sampled Households and Selected Persons	
65	Physical disability of Selected Person prevented completion of General Booklet
	General booklet (module 9) looked through or attempted but not completed
71	Selected Person looked at Booklet but unable to read & write the language of Booklet - did not begin
72	Refusal by Selected Person to continue the General Booklet
73	Interruption – unacceptably long
74	Unusual circumstance - Selected Person
75	Attempted General Booklet but no answers marked
76	Marked some items but did not attempt all items in General Booklet
77	Marked some items & attempted all items in General Booklet
78	Marked all items in General Booklet

3   Weighting Overview

The Sri Lanka five-stage stratified cluster design resulted in differential probabilities of selection for the selected persons. Consequently, each selected person in the survey does not necessarily represent the same number of persons in the target population. To account for differential probabilities of selection due to the nature of the design and to ensure accurate survey estimates, the Sri Lanka STEP requires a sampling weight for each person that participated in the survey.

In general, the objectives of the Sri Lanka STEP weighting are to construct a set of survey weights to,
1) compensate for unequal probabilities of selection;
2) compensate for household-level non-response and person-level non-response;
3) adjust the weighted sample distribution for key variables of interest (for example, age, gender, education) so that it conforms to a known population distribution for these variables.

The general weighting procedure for the Sri Lanka STEP survey required the following tasks.

1) Creation of a data file to input to the weighting process;

2) **World Bank (WB) Weight Requirement**:

   Create survey weights for sampled cases of households and persons that provided sufficient data to be considered a participant in the survey. This requirement does not necessarily include the completion of an assessment General Booklet, nor does it necessarily include the completion of all household and individual questionnaire modules.
   a) Calculation of a PSU weight for 200 sampled PSUs (i.e., 80 urban PSUs, 120 rural PSUs);
   b) Calculation of a household weight for each sampled household;
      i) Calculation of a household-level non-response adjustment independently for each PSU.
   c) Calculation of a person weight for each selected person (SP);
      i) Calculation of a non-response adjustment independently for each sampled person.

3) The required output from the weighting process is a final Sri Lanka data file with the survey design weights (i.e., for each sampled PSU, household, person) appended to each data record.

4   Weighting Process Files

4.1   Input File

The key file for the weighting process is the STATA file 'STEP - Sri Lanka - Data for Weights - Jan 23 2013' provided by the WB core team. An SPSS version of this file was created for the weighting process.

Table 4 provides the frequency counts by 'Result Code of Interview', i.e., variable '*m11_q07*', from the Sri Lanka data file 'STEP - Sri Lanka - Data for Weights - Jan 23 2013'. The table reveals that, based on the 'result code of interview' assigned by the interviewers, there are 3000 acceptable interviews (i.e., codes 51 to 78) in the Sri Lanka final data file.

Table 4: Result Code of Interview by Urban-Rural Indicator - 'STEP - Sri Lanka - Data for Weights - Jan 23 2013'			
Result Code of Interview	**Urban**	**Rural**	**Total**
51-Refusal by SP to continue modules 2-7	5	2	7
52- Unusual circumstance - Selected Person	3	0	3
61-SP unable to read language of GB	1	0	1
62-SP refusal to begin GB	1	3	4
71-SP looked through GB-unable to read & write language of GB	1	6	7
72-SP refused to continue GB after starting	3	20	23
73-GB interruption	0	1	1
74-GB terminated due to unusual circumstance	0	1	1
75-SP attempted GB-did not mark any answers	0	18	18
76-SP marked some GB items-did not attempt all items	31	100	131
77-SP marked some GB items & attempted all GB items	410	698	1108
78-SP marked all GB items	745	951	1696
Total	1200	1800	3000

4.2   Revisions to Input Data File

The following revisions were made to the Sri Lanka data file:

1) Revised the value labels for the variables '*m11_q07*', the Result of Interview Code, and 'final_status', the final status code for each case, to correspond to the Codes and code descriptions in Table 3 of this document.

2) Renamed variables as follows,
    a) RENAME VARIABLES (sector=UrbRur_Ind)
    b) RENAME VARIABLES (housingu=HhldsPSU)
    c) RENAME VARIABLES (replaced=HH_Replaced_Ind)
    d) RENAME VARIABLES (HH_Count_Listing=HHcount_Listd)
    e) RENAME VARIABLES (gnsmpid_SL=gnsmpid_LKA)

3) Revised final_status code as follows:
    a) IF (gnsmpid_LKA = 126 & HHNO=24) final_status=97.
    b) IF (gnsmpid_LKA = 126 & HHNO=27) final_status=M11_Q7

4) Created the variable 'FinStat1', a revised final status code that distinguishes non-activated cases that should have been activated from non-activated cases that should <u>not</u> have been activated, i.e., these latter cases were activated in an order that did not respect the prescribed activation order that was based on the sampling order of the reserve sample cases.
   a) With the exception of code 95, the values for 'FinStat1' are the same as the code values for 'final_status'.
   b) In the case of 'final_status'=95, the values for 'FinStat1' are as follows,
      i) Code 95 - a reserve household was not activated <u>and</u> should not have been activated.
      ii) Code 97 - a household should have been activated, but was not activated during the survey period due to an out-of-order household that was selected in its place.

   Note: The 'final status' codes 95 & 97 are treated the same way in the weighting process – they are both treated as non-participant households (eligibility unknown). So, the calculation of the sample weights is not affected by the revised 'FinStat1' status codes. However, the revision of the 'final status' codes to 'FinStat1' is relevant to the calculation of the response rate and to any subsequent non-response bias analysis and sample selection bias analysis that may be undertaken.

5) **SPSS Output File**: 'W01-LKA Final Revised Data File(23-Jan-2013)-1'
   a) This file is the primary file for input to the process of calculating the survey weights.
      i) Table 5 provides the list of key variables for weighting that are included in 'W01-LKA Final Revised Data File(23-Jan-2013)-1'
   b) This file includes all 6000 sampled cases. Table 6 provides the breakdown of the 6000 cases by Final Status Code'FinStat1', by Stratum, and by Household Sample Type.
      i) For weighting purposes, a finalized case is a sample household for which a final status code has been assigned. During the file preparations for weighting, a final status code was assigned to each of the 6000 sampled cases.

Table 5: Key Variables in Primary Input File for Weighting Process		
Variable Name	**Source**	**Description**
country	Sample File	Country abbreviation code
HHNO	Sri Lanka Data File	Household Number-assigned in field
gnsmpid_LKA	Sri Lanka Data File	PSU-SSU ID – sampled GN Blocks
HhldsPSU	Sample File (from Sample Frame)	Sample Frame # of households in PSU=GN
HHcount_Listd	Sri Lanka Data File	total # of listed households in GN Block
HHsmpSeq	Sample File	Household Sample Sequence #s 01 to 30
HHsmpType	Sample File	HH Sample Indicator
HHsmpOrder	Sample File	Sample activation order
UrbRur_Ind	Sri Lanka Data File	Stratum – Urban-Rural
GNsampled	Sri Lanka Data File	GN Sample Activation Indicator
m1c_a	Sri Lanka Data File	NUMBER A-# of eligible persons
m1c_b	Sri Lanka Data File	NUMBER B- listing # of selected person
M11_Q7	Sri Lanka Data File	Result code of interview
final_status	Sri Lanka Data File	Final Status Code for sampled case
FinStat1	Created during file preparation	Revised Final Status Code for sampled case
GN_Prob	Added to file during file preparation	Probability of Selection of PSU, i.e., GN
GN_Blocks	Added to file during file preparation	# of GN Blocks in activated GN
Smpl_Blk#	Added to file during file preparation	ID# of sampled GN Block

Table 6: Sri Lanka - Number of Finalized Cases by Final Status Code by HH Sample Type & Urban-Rural Indicator									
	Urban			Rural			Total		
Final Result Code	Initial Sample	Reserve Sample	Total	Initial Sample	Reserve Sample	Total	Initial Sample	Reserve Sample	Total
01-Refusal-Household	66	29	95	52	19	71	118	48	166
02-Unusual circumstance-Household	14	4	18	8	1	9	22	5	27
03-No household contact after 3 visits	4	4	8	30	3	33	34	7	41
04-Temporarily absent/unavailable for field period	64	40	104	72	15	87	136	55	191
05-No competent household member to interview	5	0	5	10	0	10	15	0	15
06-Language problem	1	0	1	0	0	0	1	0	1
08-household moved away	2	0	2	5	0	5	7	0	7
09- Not habitable/destroyed/converted to business	5	4	9	3	0	3	8	4	12
21- Household Refusal to continue	4	0	4	0	0	0	4	0	4
22- Household unusual circumstance	0	1	1	0	0	0	0	1	1
31-No eligible household member aged 15-64	33	9	42	23	2	25	56	11	67
32-Refusal-Refusal by SP to begin modules 2-7	6	1	7	3	0	3	9	1	10
33- Unusual circumstance - Selected Person	0	1	1	1	0	1	1	1	2
35- Unable to contact SP after three visits	11	2	13	13	0	13	24	2	26
36- Temporarily absence/unavailability of SP	2	0	2	2	0	2	4	0	4
51-Refusal by SP to continue modules 2-7	3	2	5	2	0	2	5	2	7
52- Unusual circumstance - SP (modules begun)	3	0	3	0	0	0	3	0	3
61-SP unable to read language of GB	1	0	1	0	0	0	1	0	1
62-SP refusal to begin GB	1	0	1	3	0	3	4	0	4
71-SP unable to read/write language of GB	1	0	1	6	0	6	7	0	7
72-SP refused to continue GB after starting	3	0	3	17	3	20	20	3	23
73-GB interruption	0	0	0	1	0	1	1	0	1
74-GB terminated due to unusual circumstance	0	0	0		1	1	0	1	1
75-SP attempted GB-did not mark any answers	0	0	0	15	2	17	15	2	17
76-SP marked some GB items; not all attempted	25	6	31	88	12	100	113	18	131
77-SP marked some GB items; attempted all items	335	74	409	612	84	696	947	157	1104
78-SP marked all GB items	607	135	739	827	119	946	1434	251	1685
95-Sample case not activated	0	358	358	0	733	733	0	1091	1091
97- Household not interviewed-procedure error	1	529	530	0	805	805	1	1334	1335
98-Sample selection error - incorrect eligible SP	3	1	4	7	0	7	10	1	11
Total	1200	1200	2400	1800	1800	3600	3000	3000	6000

4.3   Final Output File from Weighting Process

The output file from the weighting process is,

1) LKA-W01-STEP(HH) Weighted Data-Final Design Weights (15-Oct-2013)
    a) This file contains the original data variables from the Sri Lanka data file 'STEP - Sri Lanka - Data for Weights - Jan 23 2013' and the applicable PSU weights, household weights, person weights, and other variables that were created during the weighting process.
    b) The file includes 3037 participating households, each with a final household weight. The breakdown of the count of participating households by urban-rural indicator is shown below in Table 7.

Table 7: Sri Lanka Count of Participating Households by Urban-Rural Indicator			
	Urban	**Rural**	**Total**
Number of Participating Households	1219	1818	3037
Sum of Household Weights	1,000,625	4,773,053	5,773,678

    c) In addition to the Sri Lanka data file variables, the final weighted data file includes additional variables that were created during the weighting process. The essential variables that are needed for estimation of survey results are included in the final weighted data file.
        i) Table 8 provides the list of variables, created or revised during the weighting process, which are included in the final weighted data file.
        ii) Table 9 provides the list of other available variables from the weighting process that are not included in the final weighted data file.
        iii) Table 10 provides the frequency count of result codes for participating households in the final weighted data file 'LKA-W01-STEP(HH) Weighted Data-Basic Weights (25-Jan-2013)'.

Table 8: Variables Created or Revised During Weighting Process - *Added to Sri Lanka Data File*		
Variable Name	**Description**	**Values**
HhldsPSU	Sample Frame # of households in PSU (renamed from 'housingu')	n/a
HHcount_Listd	total # of listed households in PSU (renamed from 'HH_Count_Listing')	n/a
FinStat1	Final Status Code (revised from variable 'final_status')	Refer to Table 3 for code descriptions
GN_Prob	GN Probability of Selection	n/a
GN_Blocks	GN-# of Blocks in GN	n/a
Smpl_Blk	GN Block ID# - sampled Block	
hhsmpsiz	Sampled HHs in GN BLOCK	
W1_GNwt	Basic PSU Weight, PSU=GN	n/a
W2_GNBlkwt	Basic SSU Weight, SSU=GN Block	n/a
W2_GN_BLK_Fin	Final GN-GN Block Weight	n/a
HHsCntDW	Number of HHs in Selected Dwelling	n/a
FinRespCat_HH	Final Response Category-Households	1-Participant HH 2-Ineligible HH 3-Non-participant Eligible HH 4-Non-participant Unknown Eligibility HH
W4_FinHHwt	Final Household Weight Use for estimation of Household characteristics	n/a
FinRespCat_SP	Final Response Category-Selected Person	1-Participant SP 2-Non-participant, Eligible 3-Non-participant, Ineligible 9-Not applicable - non-participating HH
W5_FinSPwt	Final Person Weight Use for estimation of Person characteristics	n/a

Table 9: Variables Created During Weighting Process – Not *Added to Sri Lanka Data File*

Variable Name	Description	Values
PSU_FIN#	Count of 'FinRespCat_HH' in SSU	n/a
HH_k	# of participant HHs in SSU	n/a
HH_x	# of Ineligible HHs in SSU	n/a
HH_z	# of Non-participant Eligible HHs in SSU	n/a
HH_u	# of Non-participant Unknown Eligible HHs in SSU	n/a
W3_HH	Household Weight within SSU	n/a
W4_HHwt_basic	Basic Household Weight with adjustment for multiple households	n/a
Elig_Prop	Proportion of Eligible HHs - known eligibility	n/a
R_adj_HH	Response adjustment ratio for HHs	n/a
EligPersCnt	Eligible Person Count in Sampled HH	n/a
W5_SP	Selected Person n weight within household	n/a
R_adj_SP	Response adjustment for selected person	n/a
EP_PSU_sum	Total Eligible Persons-PSU Participant Households	n/a
EP_PSU_PPsum	Total Eligible Persons-PSU Participant SP Households	n/a

Table 10: Sri Lanka Final Weighted Data File – Frequency of Final Result Codes For participating households (i.e., FINRESPCAT_HH=1)

Final Result Code	Urban n	Urban % of Total n	Rural n	Rural % of Total n	Total n	Total % of Total n
32-Refusal-Refusal by SP to begin modules 2-7	7	0.2%	3	0.1%	10	0.3%
33-Interview termination-SP unusual circumstance	1	0.0%	1	0.0%	2	0.1%
35-Unable to contact SP after 3 attempts	13	0.4%	13	0.4%	26	0.9%
36-SP absent for entire survey period	2	0.1%	2	0.1%	4	0.1%
51-Refusal by SP to continue modules 2-7	5	0.2%	2	0.1%	7	0.2%
52-Interview termination-SP unusual circumstance	3	0.1%	0	0.0%	3	0.1%
61-SP unable to read language of GB	1	0.0%	0	0.0%	1	0.0%
62-SP refusal to begin GB	1	0.0%	3	0.1%	4	0.1%
71-SP looked through GB-unable to read & write language of GB	1	0.0%	6	0.2%	7	0.2%
72-SP refused to continue GB after starting	3	0.1%	20	0.7%	23	0.8%
73-GB interruption	0	0.0%	1	0.0%	1	0.0%
74-GB terminated due to unusual circumstance	0	0.0%	1	0.0%	1	0.0%
75-SP attempted GB-did not mark any answers	0	0.0%	18	0.6%	18	0.6%
76-SP marked some GB items-did not attempt all items	31	1.0%	100	3.3%	131	4.3%
77-SP marked some GB items & attempted all GB items	409	13.4%	696	22.9%	1105	36.3%
78-SP marked all GB items	742	24.4%	946	31.1%	1688	55.5%
98-Sample selection error – incorrect eligible person selected	4	0.1%	7	0.2%	11	0.4%
Total	1223	40.2%	1819	59.8%	3042	100.0%

5 Weight Calculations based on World Bank (WB) Weight Requirement

The input file for the process of calculating the weights based on the World Bank (WB) weight requirement is 'W01-LKA Final Revised Data File(23-Jan-2013)-1', which was created from the provided file 'STEP - Sri Lanka - Data for Weights - Jan 23 2013'.

5.1 First-stage Weight (GN Weight)

The basic first stage weight, i.e., the GN weight, is calculated separately for the selected urban GNs and the selected rural GNs. The urban GN weight is based on a sample size of 80 urban GNs, while the rural GN weight is based on 120 rural PSUs. Although 260 PSUs were initially selected, 202 GNs were attempted during data collection; 2 initial GNs were replaced due to accessibility difficulty, resulting in only 200 GNs actually activated during data collection. Therefore, the Sri Lanka STEP GN survey weight is based on the probability of selection for each of the 200 <u>activated</u> PSUs.

For the Sri Lanka PPS sample, the probability of selection of the jth GN in geographic area 'g', say p_{gj}, may be represented as,

$$1) \qquad p_{gj} = \frac{y_g * m_{gj}}{M_g}$$

where,

$$g = \begin{cases} 1, Urban\ GN \\ 2, Rural\ GN \end{cases}$$

y_g - the number of sampled GNs in geographic area g (i.e., y_1=80 urban PSUs, y_2=120 rural PSUs) in the case of the Sri Lanka STEP Survey),

m_{gj} - the measure of size for the jth GN in geographic area 'g', in the case of the Sri Lanka STEP Survey the measure of size for the first stage selection of GNs was the sample frame number of households in each GN in the target population,

M_g - the total number of households, based on the sample frame information, in the target population in geographic area 'g', is represented by

$$2) \qquad M_g = \sum_{g=1}^{2} \sum_{j=1}^{Y_g} m_{gj}$$

where,

Y_g - the total number of GNs in the target population.
M_1 = 1,000,625 is the sample frame count of *urban* households for the Sri Lanka STEP survey,
M_2 = 4,773,053 is the sample frame count of *rural* households for the Sri Lanka STEP survey.

NOTE: The Sri Lanka DCS selected the GNs at the first stage of sample selection. The value for the probability of selection of each first-stage GN was provided by the Sri Lanka DCS.

The basic first stage weight, W_{1gj}, for the jth GN in geographic area 'g' was calculated as the inverse of its probability of selection:

$$3) \qquad W_{1gj} = \frac{1}{p_{gj}} = \frac{M_g}{y_g * m_{gj}}$$

➢ **The basic first stage GN weight is appended to the STEP data file.**

Example 1: First Stage Weight

The basic first stage weight for the *urban* (i.e., g=1) Sri Lanka PSU # 165 which contains 3294 sample frame households is calculated as follows:

$$W_{11j} = \frac{1}{p_{1j}} = \frac{M_1}{y_1 * m_{1j}} = \frac{1,000,625}{80 * 3294} \approx 3.797$$

The basic first stage weight is similarly calculated for each of the 200 PSUs.

5.2   Second-stage Weight (GN BLOCK Weight)

The second stage unit (SSU) is a GN Block. The 200 PSUs, i.e., GNs, were segmented into several GN Blocks, and one GN Block was selected from each set of SSU segments.

Note: The second stage weight is calculated under the assumption that each sampled GN was divided into a number of segments, i.e., GN Blocks, <u>with the same number of households in each GN Block</u>, and one GN Block was selected from each sampled GN.

The basic second stage weight, W_{2js} , for the sth GN Block in GN 'j' is calculated as follows:

4) $$W_{2gjs} = \frac{s_{gj}}{n_{gjs}}$$

where,

s_{gj} -   the number of **_GN Blocks_** in the jth GN in stratum 'g',

n_{gjs} -   the number of GN Blocks sampled in the jth GN (i.e., n_{sj} =1 for all activated GNs).

Example 2: Second Stage Weight

The basic second stage weight for the *urban* (i.e., g=1) Sri Lanka PSU # 165 which contains 7 sample frame GN Blocks is calculated as follows:

$$W_{2gjs} = \frac{1}{p_{gjs}} = \frac{7}{1} = 7$$

5.3 Household Weight - Third-stage & Fourth-stage

The basic third stage weight for the kth household in GN Block 's' in GN 'j', W_{3jsk} , is calculated as follows:

$$5) \qquad W_{3gjsk} = \frac{m'_{gjs}}{m_{gjs}}$$

where,

      m'_{gjs} -   the number of **_listed_** households in GN Block 's' in GN 'j' in stratum 'g',

      n_{gjsh} -   the number of households sampled in GN Block 's' in GN 'j' in stratum 'g'.

Example 2: Third Stage Basic Weight

The basic third stage weight for each of the sampled households in the Sri Lanka PSU # 165, which contains 101 **_listed_** households, is calculated as follows:

$$W_{31jsk} = \frac{m'_{1js}}{n_{1jsh}} = \frac{101}{30} \approx 3.367$$

The basic third stage weight is similarly calculated for each of the 200 PSUs.

5.4 Household Weight – Adjustment for Multiple Households

Some selected households were found to include more than one household. In such cases, one household was randomly selected for an interview. Therefore, for each household 'k', the third stage weight is adjusted by the following factor,

$$6) \qquad MH_{gjsk} = \frac{h_{gjsk}}{n_{gjsk}}$$

where,

      h_{gjsk} -   the number of actual households in household 'k',

      n_{gjsk} -   the number of households sampled in household 'k', (i.e., n_{kh} =1 for all sampled cases).

Example 3: Third Stage Basic Weight – Adjustment for Multiple Households

The basic third stage weight for sampled household # 28 in the Sri Lanka PSU # 165, which contains 3 households, is calculated as follows:

$$W_{3gjsk} = \frac{m'_{gsj}}{n_{gsj}} * MH_{gjsk} = \frac{101}{30} * 3 \approx 10.1$$

5.4.1 Household Weight – Adjustment for Non-Response

For treatment in the weighting process the data collection outcomes for the sampled households may be summarized in four possible categories: participant, ineligible, non-participant with known eligibility, non-participant with unknown eligibility.

Participant - A participating household '**k**' is an <u>eligible</u> sampled household that provides sufficient data to be considered a participant in the survey. The sampled households '**k**' with person-level result codes 32-40, 51-78 and 98-99 satisfy the requirement for classification of the household as a survey participant. The result codes and corresponding code descriptions for participating households '**k**' are as follows:

	I. Household-level Codes – Participant, Eligible Household
Code	**Code Description**
	Individual modules 2-7 <u>not</u> begun
32	Refusal by Selected Person to begin modules 2-7
33	Unusual circumstance - Selected Person (e.g., death in family, illness, fire in dwelling, etc.)
34	Refusal - Another household member refused to let selected individual do individual modules
35	Unable to contact Selected Person after three visits during field period
36	Temporarily absence/unavailability of Selected Person for field period (info from household member)
37	Hearing Problem - Selected Person is deaf or so hard-of-hearing, no translator available
38	Learning or mental disability (includes emotional conditions like severe depression) of Selected Person
39	Speech impairment of Selected Person prevented participation, no translator available
40	Language problem – Selected Person's language not understood by survey team, no translator available
	Individual modules 2-7 begun but not completed
51	Refusal by Selected Person to continue modules 2-7
52	Unusual circumstance - Selected Person
	General Booklet (module 9) not begun
61	Unable to read or write in language of General Booklet – Selected Person refused to begin
62	Refusal by Selected Person to begin General Booklet
63	Unusual circumstance - Selected Person
64	Blindness or visual impairment of Selected Person prevented completion of General Booklet
65	Physical disability of Selected Person prevented completion of General Booklet
	General booklet (module 9) looked through or attempted but not completed
71	Selected Person looked at Booklet but unable to read & write the language of Booklet - did not begin
72	Refusal by Selected Person to continue the General Booklet
73	Interruption – unacceptably long
74	Unusual circumstance - Selected Person
75	Attempted General Booklet but no answers marked
76	Marked some items but did not attempt all items in General Booklet
77	Marked some items & attempted all items in General Booklet
78	Marked all items in General Booklet
98	Sample selection error – an incorrect eligible person was selected by the interviewer
99	Sample selection error – an ineligible person was selected by the interviewer

Ineligible – A sampled household '**x**' is categorized as <u>ineligible</u> if there is no household member in the eligible 15-64 age range. The result codes for ineligible sampled households '**x**' are as follows:

II. Household-level Codes – Ineligible Household	
Code	**Code Description**
07	Dwelling could not be found / selected address has no household
08	Vacant dwelling
09	Dwelling not habitable/dwelling destroyed/dwelling converted to commercial use
10	Household moved away
31	No household member in the eligible range of 15-64

Non-participant - Eligibility Known: A non-participating household '**z**' results when an eligible household does not participate due to refusal, unusual circumstance, or the household was selected in error. The result codes for sampled households '**z**' with known eligibility are as follows:

III. Household-level Codes – Non-participants, Eligible Household	
Code	**Code Description**
21	Household Refusal to continue Household Module (i.e., begun but not complete)
22	Household unusual circumstance (e.g., death in family, illness, fire in dwelling, etc.) prevented continuance
94	Case dropped due to data quality issue

Non-participant - Eligibility Unknown: A non-participating household '**u**' results when a household does not participate due to non-contact with a responsible household member, inconclusive information about the eligibility status of the sampled household, or sample case was not activated (this usually pertains to the reserve sample cases). The result codes for non-participating households '**u**' with unknown eligibility are as follows:

IV. Household-level Codes – Non-participants, Unknown Eligibility Household	
Code	**Code Description**
01	Household Refusal
02	Household unusual circumstance (e.g., death in family, illness, fire in dwelling, etc.)
03	No knowledgeable household member found after 3 visits (only a child, non-competent adult, etc.)
04	Temporarily absent/unavailable for field period (information from other persons)
05	No competent household member to interview (because of severe illness, mental disability, etc.)
06	Language problem – no household member spoke a language understandable by survey team, no translator available
95	Sample case not activated
96	General non-response; Reason unspecified by survey firm
97	Sample selection error – a wrong household was selected in the field

The total number of sampled households in each PSU-SSU combination may be represented by the following equation.

$$7) \qquad n_{gjs} = n_{gjsk} + n_{gjsx} + n_{gjsz} + n_{gjsu}$$

where,

n_{gjs} - the number of households sampled in the jth PSU and s^{th} SSU,

n_{gjsk} - the number of households 'k' that participate in the jth PSU and s^{th} SSU,

n_{gjsx} - the number of ineligible households 'x' in the jth PSU and s^{th} SSU,

n_{gjsz} - the number of non-participating households 'z' with known eligibility in the jth PSU and s^{th} SSU,

n_{gjsu} - the number of non-participating households 'u' with unknown eligibility in the jth PSU and sth SSU.

Table 11 provides the STEP counts from the Sri Lanka data file for the outcome categories described above.

Table 11: Sample Count of Households by Final Response Category						
Final Response Category-Households	**Urban**		**Rural**		**Total**	
		% of Total		**% of**		**% of**
	n_{1js}	n_{1js}	n_{2js}	**Total** n_{2js}	n_{gjs}	**Total** n_{gjs}
1-Participant HH (n_{gjsk})	1223	20.4%	1819	30.3%	**3042**	**50.7%**
2-Ineligible HH (n_{gjsx})	53	0.9%	33	0.5%	**86**	**1.4%**
3-Non-participant, Eligible HH (n_{gjsz})	5	0.1%	0	0.0%	5	0.1%
4-Non-participant, Unknown Eligibility HH (n_{gjsu})	1119	18.6%	1748	29.1%	**2867**	**47.8%**
Total (n_{js})	**2400**	**40.0%**	**3600**	**60.0%**	**6000**	**100.0%**

5.4.2 Household Weight – Adjustment for Unknown Eligibility

The STEP assumes that the households 'u' with unknown eligibility in the sth SSU are comprised of some proportion, say ε_g, that are eligible households and the complementary proportion, 1 - ε_g, that are ineligible households. Note that the calculation of ε_g is carried out independently for the urban and rural PSUs.

Amongst the n_{jsu} non-participating households 'u' with unknown eligibility the proportion of households with known eligibility is estimated to be

$$8) \quad \varepsilon_g = \frac{Total\ \#\ of\ known\ eligible\ cases}{Total\ \#\ of\ known\ eligible\ \&\ known\ ineligible\ cases} = \frac{n_{gjsk} + n_{gjsz}}{n_{gjsk} + n_{gjsx} + n_{gjsz}}$$

➢ **Note that the eligible proportion is the same for all eligible participating households in a sampled GN Block.**

Example 4: Eligible Proportion

The following response category counts for PSU # 165 provide input for calculation of the proportion of households with known eligibility for this PSU.

Participant HH	Ineligible HH	Non-Participant HH (Eligible)	Non-Participant HH (Eligibility Unknown)
15	1	0	14

For the *urban* PSU # 165, the calculation of the eligibility proportion is as follows,

$$\varepsilon_{165} = \frac{Total\ \#\ of\ known\ eligible\ cases}{Total\ \#\ of\ known\ eligible\ \&\ known\ ineligible\ cases}$$

$$= \frac{n_{jsk} + n_{jsz}}{n_{jsk} + n_{jsx} + n_{jsz}}$$

$$= \frac{15 + 0}{15 + 1 + 0} = \frac{15}{16} \approx .9375$$

5.4.3 Household Weight - Adjustment for Non-participation

A household-level response adjustment is necessary to compensate for households that did not participate in the survey. The household-level response adjustment for the sth SSU in the 'j' PSU in geographic area 'g', R_{gjs}, is calculated as follows:

$$9) \quad R_{gjs} = \frac{Total\ eligible\ households\ in\ SSU\ 's'}{Total\ particpating\ households\ in\ SSU\ 's'} = \frac{n_{jsk} + n_{jsz} + (\varepsilon * n_{jsu})}{n_{jsk}}$$

- ➢ The response adjustment ratio, R_{gj}, incorporates the adjustment for unknown eligibility, ε_g.
- ➢ The households that are categorized as ineligible, n_{jx}, are excluded from the calculation of the household-level response adjustment.

Example 5: Second Stage Response Adjustment for sampled Households

For example, the following results were obtained for PSU #165

n_{gjs}	n_{gjsk}	n_{gjsx}	n_{gjsz}	n_{gjsu}
30	15	1	0	14

Therefore,

$$R_{1,165} = \frac{Total\ eligible\ households\ in\ SSU\ 's'}{Total\ particpating\ households\ in\ SSU\ 's'} = \frac{n_{jsk} + n_{jsz} + (\varepsilon_1 * n_{jsu})}{n_{jsk}}$$

$$= \frac{15 + 0 + (0.9375 * 14)}{16}$$

$$\approx 1.875$$

The response adjustment ratio, R_{gj}, is similarly calculated for all sampled PSUs.

5.4.4 Household Weight - Final

The final household weight is the weight used for producing household-level estimates, such as the proportion of households with a specific attribute. This weight is appended to the STEP data file.

The final household weight is calculated as the product of the PSU weight, the basic household weight, and the response adjustment for the basic household weight. Therefore, for household 'k' in PSU 'j', the final household weight, W_{4H}, is calculated as follows:

$$10) \quad W_{4H} = W_{1gj} * W_{2gjs} * W_{3gjsk} * MH_k * R_{gjs}$$

➤ **The final second stage household weight is appended to the STEP data file.**

Example 6: Second Stage – Final Household Weight

For PSU #165,

$$
\begin{aligned}
W_{4H} &= W_{1gj} * W_{2gjs} * W_{3gjsk} * MH_k * R_{gjs} \\
&= 3.797 \times 7 \times 3.367 \times 3 \times 1.875 \\
&\approx 503.36
\end{aligned}
$$

The final household weight is similarly calculated for all sampled households.

5.4.5 Weighted Household Counts

Based on a table of 2012 Census data provided by the Sri Lanka Department of Census & Statistics (DCS), there is a total of 5,773,678 households in the target population, comprised of 1,000,625 urban households and 4,773,053 rural households. The numbers of households on the sample frame includes households that no longer existed at the time of the STEP survey, e.g. vacant dwellings, demolished dwellings, no eligible person for STEP in household, etc..

The total unweighted number of households in the sampled PSUs is provided in Table 12. The counts in Table 12 are unweighted counts that have not been adjusted for ineligible households found in the STEP sample.

Table 12: Number of Households in Sampled PSUs by Source of Count by Urban-Rural Indicator			
Source of Count	Urban	Rural	Total
Sample Frame Count	113,647	73,433	187,080
Listing Operation (Est.)	78,868	53,179	132,047

Table 13 provides,
  i.   the total number of households based on the sample frame count of households;
  ii. the weighted count of households based on the listed households from the Sri Lanka listing operation.

Table 13: Sri Lanka – Weighted Number of Households – Sample Frame vs. STEP Data File						
	Urban		Rural		Total	
Source of PSU Household Count	# of Sampled PSUs/SSUs	Weighted Sum of Households	# of Sampled PSUs/SSUs	Weighted Sum of Households	# of Sampled PSUs/SSUs	Weighted Sum of Households
1) Based on # of Hhlds in Sample Frame from DCS Information	80	1,000,625[*]	120	4,773,053[*]	200	5,773,678[*]
2) Based on # of Listed Hhlds in STEP Data File 'STEP - Sri Lanka - Data for Weights - Jan 23 2013' [**all households**, i.e., participants, non-participants, ineligibles]	80	779,734	120	4,188,859	200	4,968,594
3) Based on # of Listed Hhlds in STEP Data File 'STEP - Sri Lanka - Data for Weights - Jan 23 2013' [**eligible households only**]	80	777,843	120	4,182,837	200	4,960,680

* Unweighted sum of PSU households

Notes for Table 13
  1. For source 1, the 'sum of households' is an unweighted total - it is the actual total number of households for all PSUs in the sample frame.

  2. For source 2, the decrease in the weighted number of households based on the listed household counts is due to the decreased count of households from the listing operation for the sampled SSUs. These weighted estimates of the number of households are not adjusted for sampled ineligible households.

  3. For source 3, the decrease in the weighted number of households is due to the adjustment of the household weights for ineligible households found in the STEP sample; the source 3 figures are also adjusted for the multiple households.
     a. *The 'Weighted Sum of Households' is the estimate of the number of households in the STEP target population.*

5.5 Person Weight

The basic third stage weight for the ith person, W_{5gjski}, is calculated as follows:

$$11) \quad W_{5gjski} = \frac{a_{gjsk}}{b_{gjsk}} = \frac{a_{gjsk}}{1} = a_{gjsk}$$

where,

a_{gjsk} - the number of eligible persons in the kth household in the sth SSU,

b_{gjsk} - the number of persons sampled in the kth household in the sth SSU, (i.e., n_{jk} =1 person in each Sri Lanka sampled household).

N.B. $A_{gjsk} = \sum_{k=1}^{n_{gjsk}} a_{gjsk} = total \text{ } \# \text{ } of \text{ } selected \text{ } persons \text{ } in \text{ } the \text{ } s^{th} \text{ } SSU$

where,

n_{gjsk} - the number of households 'k' that participate in the sth SSU

Example 7: Third Stage – Basic Person Weight within sampled Household

For PSU #165, selected household # 28 has 6 eligible persons. Therefore,

$$W_{5gjki} = \frac{a_{gjsk}}{b_{gjsk}} = \frac{6}{1} = 6$$

5.5.1 Person Weight - Response Adjustment

For eligible participating households, there are three possible person-level outcomes, a selected eligible person participates in the survey, a selected eligible person does not participate, a selected person is ineligible.

Participant SP - A participating selected person '**i**' is an eligible sampled person that provides sufficient data to be considered a participant in the survey. The sampled persons '**i**' with result codes 51-78 satisfy the requirement for classification of the person as a survey participant. The result codes for participating persons '**i**' are as follows:

I. Person-level Codes – Participant in Eligible Selected Household

Code	Code Description
	Individual modules 2-7 begun but not completed
51	Refusal by Selected Person to continue modules 2-7
52	Unusual circumstance - Selected Person
	General Booklet (module 9) not begun
61	Unable to read or write in language of General Booklet – Selected Person refused to begin
62	Refusal by Selected Person to begin General Booklet
63	Unusual circumstance - Selected Person
64	Blindness or visual impairment of Selected Person prevented completion of General Booklet
65	Physical disability of Selected Person prevented completion of General Booklet
	General booklet (module 9) looked through or attempted but not completed
71	Selected Person looked at Booklet but unable to read & write the language of Booklet - did not begin
72	Refusal by Selected Person to continue the General Booklet
73	Interruption – unacceptably long
74	Unusual circumstance - Selected Person
75	Attempted General Booklet but no answers marked
76	Marked some items but did not attempt all items in General Booklet
77	Marked some items & attempted all items in General Booklet
78	Marked all items in General Booklet

Non-participant SP - A non-participating person '**q**' results when an eligible person does not participate due to refusal, unusual circumstance, non-contact, language problem, physical disability (hearing, vision, other) or the person was selected in error. The result codes for non-participating persons '**q**' are as follows:

II. Person-level Codes – Non-participant in Eligible Selected Household

Code	Code Description
	Individual modules 2-7 not begun
32	Refusal by Selected Person to begin modules 2-7
33	Unusual circumstance - Selected Person (e.g., death in family, illness, fire in dwelling, etc.)
34	Refusal - Another household member refused to let selected individual do individual modules
35	Unable to contact Selected Person after three visits during field period
36	Temporarily absence/unavailability of Selected Person for field period (info from household member)
37	Hearing Problem - Selected Person is deaf or so hard-of-hearing, no translator available
38	Learning or mental disability (includes emotional conditions like severe depression) of Selected Person
39	Speech impairment of Selected Person prevented participation, no translator available
40	Language problem – Selected Person's language not understood by survey team, no translator available
98	Sample selection error – an incorrect eligible person was selected by the interviewer

Ineligible SP - A sampled person '**f**' is categorized as <u>ineligible</u> if the person is not a member of the target population, for example, the person's age is outside the eligible 15-64 age range. Ineligible sampled persons '**f**' are as assigned the result code 99.

III. Person-level Codes – Non-participant in Eligible Selected Household	
Code	**Code Description**
99	Sample selection error – an ineligible person was selected by the interviewer

The third stage person weight is only relevant for sampled households in the final response category FinRespCat_HH=1, i.e., participating households. Table 14 provides for the activated sample cases the frequency counts of participant and non-participant households by stratum of the Final Status Codes.

Table 14: Sri Lanka – Activated Sample Cases: Frequency of Final Status Codes by Stratum

		Urban	Rural	Total	
FINRESPCAT_HH	**Final Status Code**	**n**	**n**	**n**	**%**
4-Non-participant	01-Refusal-Household	95	71	166	3.4%
	02-Unusual circumstance-Household	18	9	27	0.6%
	03-No contact after 3 visits	8	33	41	0.9%
	04-Temporarily absent/unavailable for field period	104	87	191	4.0%
	05-No competent household member to interview	5	10	15	0.3%
	06-Language problem	1		1	0.0%
3-Non-participant	21-Household Refusal to continue Household Module	4		4	0.1%
	22-Household module terminated-unusual circumstance	1		1	0.0%
4-Non-participant	97-Household not interviewed - field procedure error	530	805	1335	27.7%
Sub-Total: Non-participant Households		**766**	**805**	**1781**	**50.8%**
1-Participant HH	32-Refusal-Refusal by SP to begin modules 2-7	7	3	10	0.2%
	33-Interview termination-SP unusual circumstance	1	1	2	0.0%
	35-Unable to contact SP after 3 attempts	13	13	26	0.5%
	36-SP absent for entire survey period	2	2	4	0.1%
	51-Refusal by SP to continue modules 2-7	5	2	7	0.1%
	52-Interview termination-SP unusual circumstance	3	0	3	0.1%
	61-SP unable to read language of GB	1	0	1	0.0%
	62-SP refusal to begin GB	1	3	4	0.1%
	71-SP unable to read/write language of GB	1	6	7	0.1%
	72-SP refused to continue GB after starting	3	20	23	0.5%
	73-GB interruption	0	1	1	0.0%
	74-GB terminated due to unusual circumstance	0	1	1	0.0%
	75-SP attempted GB-did not mark any answers	0	18	18	0.4%
	76-SP marked some GB items-did not attempt all items	31	100	131	2.7%
	77-SP marked some GB items & attempted all GB items	409	696	1105	22.9%
	78-SP marked all GB items	742	946	1688	35.0%
	98-Sample selection error – incorrect eligible person	4	7	11	0.2%
Sub-Total: Participant Households		**1223**	**1819**	**3042**	**49.2%**
Total: Activated Sample Cases		**1989**	**2834**	**4823**	**100.0%**

In each participating household in the STEP Survey, one eligible person was selected to be interviewed. Therefore, for each participating household the basic person weight, W_{5gjski} , would also be the final person weight within a sampled PSU <u>if all persons in the PSU participated</u>. However, when there is one or more non-participating selected persons within a sampled PSU, the basic person-weights for the PSU require a response adjustment so that the final person weight for each participating person in the PSU is also representative of the non-participating person(s) in the sampled households in the PSU. In other words, the total number of participating persons in a sampled PSU must represent the total number of eligible persons in the sampled households in the PSU.

Therefore, within each sampled PSU, a person-level response adjustment to the basic person weight is necessary to compensate for persons that did not participate in the survey.

The total number of eligible persons in each SSU may be represented by the following equation:

$$12) \; A_{gjsk} = A_{gjski} + A_{gjskq}$$

where,

$$A_{gjsk} = \sum_{k=1}^{n_{gjsk}} a_{gjsk}$$ - the total number of eligible persons for all n_{gjsk} participating households in the jth PSU and s^{th} SSU,

$$A_{gjski} = \sum_{k=1}^{n_{gjsk}} a_{gjski}$$ - the total number of eligible persons 'i' that participate for all n_{gjsk} participating households in the jth PSU and s^{th} SSU,

$$A_{gjskq} = \sum_{k=1}^{n_{gjsk}} a_{gjskq}$$ - the total number of non-participating eligible persons 'q' for all n_{gjk} participating households in the jth PSU and s^{th} SSU,

Therefore, the person-level response adjustment for the js^{th} PSU_SSU, R_{jsk}, is calculated as follows:

$$13) \quad R_{gjsk} = \frac{\textit{Total eligible persons in all participating sampled households in PSU 'j'}}{\textit{Total particpating persons in all participating sampled households in PSU 'j'}} = \frac{A_{gjsk}}{A_{gjski}}$$

Note that ineligible persons are excluded from the above ratio adjustment for person-level non-response.

Example 7: Third Stage – Basic Person Weight Response Adjustment

i) For PSU #165, there are 42 eligible selected persons across all participating households. Also, all selected persons participated in the survey and thus the selected persons represent the 53 eligible persons in the participating households. Therefore,

$$A_{gjsk} = \sum_{k=1}^{n_{gjsk}} a_{gjk} = 42 \; ; \; A_{gjski} = \sum_{k=1}^{n_{gjsk}} a_{gjski} = 42$$

$$\boldsymbol{R_{gjsk}} = \frac{A_{gjsk}}{A_{gjski}} = \frac{\mathbf{42}}{\mathbf{42}} = \mathbf{1.00}$$

ii) For PSU # 101, the total number of eligible selected persons across all participating households is 49. For PSU # 101, there is one non-participating selected person from sampled household #23 with 4 eligible persons. Hence, the total number of eligible persons across the participating households where a selected person participates is 45 (i.e., 49 minus 4). Therefore,

$$A_{gjsk} = \sum_{k=1}^{n_{gjsk}} a_{gjsk} = 49 \; ; A_{gjski} = \sum_{k=1}^{n_{gjsk}} a_{gjski} = 45$$

$$\boldsymbol{R_{gjsk}} = \frac{A_{gjsk}}{A_{gjski}} = \frac{\mathbf{49}}{\mathbf{45}} \approx \mathbf{1.09}$$

Table 15 provides the activated households sample counts of the Final Response Category for selected persons.

Table 15: Activated Household Sample Counts (n) by Selected Person (SP) Final Response Category						
Final Response Category-SP (For Participating Households)	**Urban**		**Rural**		**Total**	
	n	**% of Total n**	**n**	**% of Total n**	**n**	**% of Total n**
1-Participant SP	1196	36.3%	1793	14.5%	**2989**	**50.8%**
2-Non-participant SP, Eligible	27	0.0%	26	0.0%	**53**	**0.1%**
3-Non-participant SP, Ineligible	0	0.0%	0	0.0%	0	0.0%
9-Not applicable - non-participating HH	766	35.1%	1015	14.0%	**1781**	**49.1%**
Total	**1989**	**71.4%**	**2834**	**28.6%**	**4823**	**100.0%**

Based on the above table, there are 2989 participating selected persons (i.e., 1196 urban participants and 1793 rural participants) for which a final non-zero person weight was calculated. The sampled persons in categories 2, 3, and 9 will have a zero value final person weight.

5.5.2 Person Weight - Final

The final person weight is the weight used for producing person-level estimates, such as the proportion of persons with a specific attribute. This weight is appended to the STEP data file.

The final person weight is calculated as the product of the final household weight (which incorporates the PSU weight and the response adjustment for the basic household weight), the basic third stage person weight, and the response adjustment for the basic person weight. Therefore, for the selected person 'i' in household 'k' in PSU 'j', the final person weight, W_{SP} is calculated as follows:

$$14)\ W_{SP} = W_H * W_{3gjski} * R_{gjsk}$$

➢ **The final third stage person weight is appended to the STEP data file.**

Example 8: Third Stage – Final Person Weight

i) For PSU #101, selected household # 7 with 6 eligible persons, the selected person was a participant. There was on non-participating selected person in PSU #101, therefore, $R_{gjsk} \approx 1.09$

$$W_{SP} = 438.79 * 6 * 1.09 \approx 2866.74$$

ii) For PSU #165, in participant household # 28 with 6 eligible persons, the selected person was a participant. For this household,

$$W_H \approx 503.36; \qquad W_{5gjki} = 6; \qquad R_{jk} \approx 1.00$$

$$W_{SP} = 503.36 * 6 * 1.00 \approx 3020.16$$

The final person weight is similarly calculated for all sampled persons. Note that the final person weight will be zero for all non-participating selected persons.

5.5.3 Weighted Person Counts

The weighted count of the non-institutionalized persons 15 to 64 years of age (inclusive) living in eligible private dwellings in urban and rural areas of **Sri Lanka** at the time of data collection is provided in Table 16. The 'Weighted # of Eligible Persons' in this table is the estimated number of persons aged 15 to 64 in the Sri Lanka target population based on the STEP sample.

Table 16: Eligible Activated Households - Sample Frequency & Weighted Persons Totals by Final Selected Person Response Category						
	Urban		Rural		Total	
Final Response Category-SP **(For Participating Households only)**	**n**	**Weighted # of Eligible Persons**	**n**	**Weighted # of Eligible Persons**	**n**	**Weighted # of Eligible Persons**
1-Participant SP	1196	2,216,805	1793	10,997,707	**2989**	**13,214,513**
2-Non-participant SP, Eligible	27	0	26	0	**53**	0
7-Not applicable - Ineligible HH	53	0	33	0	**86**	0
8-Not applicable - non-activated HH	358	0	733	0	**1091**	0
9-Non-participant SP - non-participating activated HH	766	0	1015	0	**1781**	0
Total	**2400**	**2,216,805**	**3600**	**10,997,707**	**6000**	**13,214,513**

For participating selected persons, Table 17 provides a summary of the central tendency of the number of eligible persons in the participating households. Based on the average of 2.73 eligible persons per participating household and an estimated 4,960,680 households, there are approximately 13,500,000 persons aged 15 to 64 in the target population.

Table 17: Summary of Average # of Eligible Persons, Average Person Weights & Average Household Weights			
Statistic	**NUMBER A- # of eligible persons**	**Final Person Weight**	**Final Household Weight**
	3042	3042	3042
Mean	2.7295	4,412.9787	1,630.7297
Std. Error of Mean	.02180	83.7451	25.7161
Median	2.0000	3,233.0840	1,517.2076
Mode	2.00	0.0000	591.0865
Minimum	1	0.0000	151.7709
Maximum	11	65,835.2138	13,167.0428
Sum	8303	13,424,281	4,960,680