

March 13, 2016

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

Philippines

March 13, 2016

Table of Contents

1	Survey Design Overview.....	1
1.1	Target Population.....	1
1.2	Sample Frame	1
1.2.1	Primary Sample Frame of Barangays	1
1.2.2	Sample Frame of PSUs, i.e., Barangay Segments.....	2
1.3	Sample Design.....	2
2	Data Collection Outcomes	4
3	Weighting Overview.....	6
4	Files - Weighting Process	7
4.1	Input Files.....	7
4.1.1	Revisions to Input Data File.....	7
4.1.1.1	Input File – Revised for Weighting.....	8
4.2	Final Output Files from Weighting Process.....	9
5	Weight Calculations based on World Bank (WB) Weight Requirement.....	12
5.1	First-stage Weight (PSU Weight).....	12
5.2	Second-stage Weight (Household Weight).....	13
5.2.1	Basic Household Weight	13
5.2.2	Second-stage Response Adjustment.....	14
5.2.2.1	Eligible Proportion of Households	17
5.2.2.2	Response Adjustment Factor for Non-participation.....	18
5.2.3	Final Household Weight.....	19
5.2.4	Weighted Household Counts	20
5.3	Third-stage Weight (Person Weight)	21
5.3.1	Basic Third-stage Person Weight	21
5.3.2	Third-stage Person-level Response Adjustment.....	22
5.3.3	Final Third-stage Person Weight.....	25
5.3.4	Weighted Person Counts	26
6	Response Rate.....	26
7	Jackknife Weights.....	28
8	Using the Jackknife Weights	28
8.1	Example: Philippines	29

March 13, 2016

1 Survey Design Overview

The Philippines survey firm implemented a Full Literacy Assessment design. The Full assessment required each selected person to attempt to complete a General Booklet comprising Reading Components and a set of Core Literacy Items. Subsequently, those respondents that correctly answered at least 3 Core Literacy items were assigned one of four assessment Exercise Booklets to complete. The full assessment sampling objective was to have a minimum of about 2400 completed Exercise Booklets, with a distribution of about 600 complete cases for each of the four Exercise Booklets.

1.1 Target Population

The target population for the Philippines STEP survey comprises all non-institutionalized persons 15 to 64 years of age (inclusive) living in private dwellings in urban areas of the country at the time of data collection. This includes all residents except foreign diplomats and non-nationals working for international organizations.

The following table was extracted from the Philippines National Sample Design & Planning Report (NSDPR); it shows the overall distribution of the urban households in the STEP target population by Region based on the 2010 Census of Population.

Table 1: Number of eligible urban Households, Census 2010

Regions	Urban
ARMM	70,820
CAR	96,823
CARAGA	134,327
N C R	2,704,838
REGION I - ILOCOS REGION	133,622
REGION II - CAGAYAN VALLEY	85,002
REGION III - CENTRAL LUZON	1,141,304
REGION IV-A - CALABARZON	1,708,795
REGION IV-B - MIMAROPA	126,634
REGION IX - ZAMBOANGA PENINS	244,284
REGION V - BICOL REGION	165,146
REGION VI - WESTERN VISAYAS	370,995
REGION VII - CENTRAL VISAYAS	33,578
REGION VIII - EASTERN VISAYAS	16,054
REGION X - NORTHERN MINDANAO	385,118
REGION XI - DAVAO REGION	600,971
REGION XII - SOCCSKSARGEN	398,603
Grand Total	8,416,914

1.2 Sample Frame

1.2.1 Primary Sample Frame of Barangays

The primary sample frame for the Philippines STEP Survey is the Excel file '[STEP Philippines - MOD Urban barangays for drawing_20150608](#)'. This sample frame, provided by the Philippines survey firm, is a list of 5028 urban barangay survey units, which are used to create the list of the first stage sampling units or PSUs, i.e., Barangay Segments.

Table 2 provides the primary sample frame count of the barangays & the households in the target population by stratum.

Stratum	# of Barangays	% of Total Barangays	# of Households	% of Total Households
NCR Urban	1,667	33.2%	2,704,838	32.1%
Other Urban	3,361	66.8%	5,712,076	67.9%
Total	5,028	100.0%	8,416,914	100.0%

1.2.2 Sample Frame of PSUs, i.e., Barangay Segments

Barangay Segments were created by dividing each barangay on the primary sample frame into approximately equal-sized segments based on the number of households in the barangay.

The SPSS file ‘PHL_F01 - BGY- SEGs Frame_08-Jun-2015’ is the Sample Frame of PSUs, i.e., Barangay Segments.

Each urban PSU in the sample frame of PSUs is uniquely identified by the variables ‘UniqueID’, i.e. the barangay ID#, and ‘SEG#_BGY’, i.e. Segment ID# within a barangay.

Stratum	# of Barangays	# of Barangay Segments	# of Households
NCR Urban	1,667	27,052	2,704,838
Other Urban	3,361	57,158	5,712,076
Total	5,028	84,210	8,416,914

1.3 Sample Design

The Philippines sample design is a 4 stage sample design. There was no explicit stratification but the sample is implicitly stratified by Urban Region. Implicit stratification was achieved by sorting the PSUs by Urban Region and selecting a systematic sample of PSUs.

First Stage Sample

The primary sample unit (PSU) is a Barangay Segment. The first stage units were selected by the World Bank Survey Methodologist.

Each PSU is uniquely defined by the sample frame variable ‘SEG#_BGY’, i.e., the Segment ID# within a Barangay. The sampling objective was to conduct interviews in 200 PSUs. In addition, 25 extra PSUs were selected for use in case it was impossible to conduct any interviews in one or more initially selected PSUs. (N.B. None of the 25 extra PSUs were required to be activated.)

Second Stage Sample

The second stage sample unit (SSU) is a dwelling. The sampling objective was to obtain interviews at 15 dwellings within each selected PSU. The dwellings were selected in each PSU using a systematic random method.

Third Stage Sample

The third stage sample unit is a household. The sampling objective is to select one household within each selected dwelling. The households are randomly selected with equal probability in each PSU.

N.B. The Philippines firm indicated that all selected dwellings contained one household, i.e., there were no multiple household dwellings in the STEP sample.

Fourth Stage Sample

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

Sample Size

The Philippines firm's sampling objective is to obtain interviews from 3000 individuals in the urban areas of Philippines.

Table 4 provides the overall target sample allocation for the initial household sample and reserve household sample in the Initial PSUs and the Reserve PSUs.

Stratum	<i>Initial Sample PSU</i>				<i>Reserve Sample PSU</i>			
	# of PSUs	Initial HH Sample	Reserve HH Sample	Total	# of PSUs	Initial HH Sample	Reserve HH Sample	Total
NCR Urban	65	975	975	1950	8	120	120	240
Other Urban	135	2025	2025	4050	17	255	255	510
Total	200	3000	3000	6000	25	375	375	750

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 result code assigned to the sampled household and the result code assigned to the selected person. The result codes used in STEP are listed in Table 5. A variable called 'final_status' with these code values was assigned to each of the 6000 sampled cases.

Table 5: <i>Full Literacy Assessment</i> Result Codes for Sampled Households and Selected Persons	
I. Household-level Codes	
Code	Code Description
01	Refusal-Household
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 – insufficient address information in sample frame
08	Vacant dwelling / selected address has no household
09	Dwelling not habitable/dwelling destroyed/dwelling converted to commercial use
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	Household not interviewed due to sample selection error – wrong household selected in its place
II. Person-level Codes	
Code	Code Description
Individual modules 2-7 <u>not</u> begun or unacceptable interview	
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
65	Physical disability of Selected Person prevented completion of General Booklet

Table 5: <i>Full Literacy Assessment</i> Result Codes for Sampled Households and Selected Persons	
	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, General Booklet administration terminated
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
	Exercise Booklet (module 9)
81	No attempt at any of the exercise booklet items / refusal to start
82	Interruption too long during exercise booklet administration
83	Unusual circumstance, Exercise Booklet administration terminated
84	Attempt at some Exercise Booklet items but no answers marked
85	Attempt at all Exercise Booklet items but no answers marked
86	Marked some Exercise Booklet items, not all items attempted
87	Marked some Exercise Booklet items, all items attempted
88	Marked all Exercise Booklet items

3 Weighting Overview

While the Philippines four-stage sample design greatly enhanced the operational feasibility of data collection, it 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, STEP requires a basic sampling weight for each person that participated in the survey.

In general, the objectives of the STEP weighting process are

- 1) to construct a set of basic survey weights to,
 - a) compensate for unequal probabilities of selection;
 - b) compensate for household-level non-response and person-level non-response;
- 2) to create a set of replicate weights for use in calculating the standard error of the survey estimates.

The general weighting procedure for the Philippines STEP survey required the following:

- 1) Creation of a data file to input into 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;
 - 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.
 - d) Calculation of a set of household-level replicate weights;
 - e) Calculation of a set of replicate weights for selected persons.
- 3) The required output from the weighting process is a final Philippines data file with the survey design weights (i.e., for each sampled PSU, household, person) and the replicate weights (i.e., for each sampled household and sampled person) appended to each data record.

4 Files - Weighting Process

4.1 Input Files

The key file for the weighting process is the EXCEL file ‘STEP PHL - Data requirements for weighting_v4’ provided by the Philippines survey firm. This file contains a data record for each of the 6000 sampled households from the Philippines sample frame. An SPSS version of the file, i.e., ‘STEP STEP PHL - Data requirements for weighting_v4.sav’, was created for the weighting process.

4.1.1 Revisions to Input Data File

The following general revisions were made to the Philippines SPSS data file ‘STEP PHL - Data requirements for weighting_v4’:

- 1) ‘m11_q07’, the Result of Interview Code: added value labels; the value labels for ‘m11_q07’ are as follows:

00-No assigned code	51-Refusal by SP to continue modules 2-7'
01-Refusal-Household	52-Interview termination-SP unusual circumstance'
02-Unusual circumstance-Household	61-SP unable to read language of General Booklet (GB)'
03-No knowledgeable household member found after 3 visits	62-SP refusal to begin GB'
04-Temporarily absent/unavailable for field period	63-SP unable to do GB due to unusual circumstance'
05-No competent household member to interview (e.g., due to severe illness, mental disability, etc.)	64-SP blindness or visual impairment-unable to do GB'
06-Language problem	65-SP physical disability'
07-Dwelling could not be found - insufficient Sample Frame address info	71-SP looked through GB-unable to read & write language of GB'
08-Vacant dwelling / selected address has no household	72-SP refused to continue GB after starting'
09-Dwelling not habitable/dwelling destroyed/dwelling converted to commercial use	73-GB interruption too long'
10-household moved away	74-GB terminated due to unusual circumstance'
21-Household Refusal to continue Household Module (i.e., begun but not complete)'	75-Attempted GB but no answers marked'
22-Household module terminated due to unusual circumstance'	76-Attempted part or all of GB; less than 3 correct Core answers'
31-No eligible household member aged 15-64'	77-Marked some items & attempted all items in GB'
32-Refusal-Refusal by SP to begin modules 2-7'	78-Marked all items in GB'
33-Interview termination-SP unusual circumstance'	81-No attempt at any exercise booklet items / refusal to start'
34-Refusal-Other Household Member refused to let SP participate'	82-Interruption too long during exercise booklet administration'
35-Unable to contact SP after 3 attempts'	83-Unusual circumstance, Exercise Booklet administration terminated'
36-SP absent for entire survey period'	84-Attempt at some Exercise Booklet items, no answers marked'
37-SP deaf or hard of hearing'	85-Attempt at all Exercise Booklet items, no answers marked'
38-SP learning or mental disability'	86-Marked some Exercise Booklet items, not all items attempted'
39-SP speech impairment'	87-Marked some Exercise Booklet items, all items attempted'
40-SP language problem'	88-Marked all Exercise Booklet items'

- 2) **final_status**: Final status code for sampled cases; added value labels. Refer to Table 5 in section 2 for the list of **final_status** codes.

4.1.1.1 Input File – Revised for Weighting

The SPSS Revised Input File for weighting is ‘W00-PHL Final Revised Data File(6-Mar-2016)’. This file is the primary file for the process of creating the survey weights; it includes all 6000 sampled cases; each case includes a final status code. Table 7 provides the list of key variables for weighting purposes that are included in the input file.

Table 7: Key Variables in Primary Input File for Weighting Process		
Variable Name	Source	Description
Country	Sample File	Country abbreviation code
Cluster	Philippines Final Data File	Cluster
SMPLID_HH	Sample File	Sampled case ID#
PSUsmpID	Sample File	PSU Sample ID#
STEP_Stratum	Sample File	Stratum
FrmeHHs	Sample File	HHs in Sample Frame (est.)
HHSmpid	Sample File	HH Sample Sequence within PSU, #s 01-30
HHsmpType	Sample File	Type of Sampled HH - Initial or Reserve
HHactOrd	Sample File	Sample activation order (HHsmpOrder in Sample File)
Resp_Gender	Philippines Final Data File	Gender of Selected Person
Resp_Age	Philippines Final Data File	Age of Selected Person
m1c_a	Philippines Final Data File	NUMBER A-# of eligible persons
m11_q07	Philippines Final Data File	Result code of interview
final_status	Philippines Final Data File	Final Result Code
PSUhhs_LSTd	Philippines Final Data File	# of Listed HHs in PSU

Table 8 provides the breakdown of the 6000 sample cases by the final result of the interview attempt, i.e., by *'final_status'* code.

Table 8: Philippines: Number of Sampled Dwellings by Final Status Code of Interview *'final_status'*

FINAL STATUS OF INTERVIEW	# of Dwellings
01-Refusal-Household	105
02-Unusual circumstance-Household	57
03-No knowledgeable household member found after 3 visits	3
61-SP unable to read language of GB	24
62-SP refusal to begin GB	23
63-SP unable to do GB due to unusual circumstance	66
64-SP blindness or visual impairment-unable to do GB	25
65-SP physical disability	36
71-SP looked through GB-unable to read/write language of GB	16
72-SP refused to continue GB after starting	5
73-GB interruption too long	87
74-GB terminated - unusual circumstance	4
75-Attempted General Booklet, no answers marked	3
76-Attempted part or all of General Booklet; <3 correct Core answers	83
81-No attempt at any exercise booklet items / refusal to start	178
82-Interruption too long during exercise booklet administration	228
83-Unusual circumstance, Exercise Booklet administration terminated	14
84-Attempt at some Exercise Booklet items, no answers marked	34
85-Attempt at all Exercise Booklet items, no answers marked	14
86-Marked some Exercise Booklet items, not all items attempted	21
87-Marked some Exercise Booklet items, all items attempted	142
88-Marked all Exercise Booklet items	368
95- Sample case justifiably not activated, e.g., reserve sample HH not needed	2835
TOTAL	6000

4.2 Final Output Files from Weighting Process

The final output file from the weighting process is,

1) *PHL-W01-STEP(HH) Weighted Data-Final Design Wts & Repl Wts(10-Mar-2016)*

- a) This file contains the original data variables from the Philippines data file *'STEP PHL - Data requirements for weighting_v4'* and the applicable PSU weights, household weights, person weights, replicate household weights, replicate person weights, and other associated variables that were created during the weighting process.
- b) The file contains a data record for each of the 3165 cases that were activated or should have been activated during the data collection operation.
 - i) 3000 participating households with a final non-zero household weight
 - ♦ For participating households, the weighted sum of the household weights is 8,704,367.
 - ii) 3000 participating selected persons with a final non-zero person weight
 - ♦ For participating individuals, the weighted sum of the person weights is 22,902,060.
 - iii) Set of 30 household replicate weights
 - iv) Set of 30 person weights.

- c) In addition to the Philippines data file original variables, the 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 9 provides a list of the variables that are included in the final weighted data file '*PHL-W01-STEP(HH) Weighted Data-Final Design Wts & Repl Wts(10-Mar-2016)*'.
 - ii) Table 11 provides a list of variables from the weighting process that are not included in the final weighted data file. These variables are available from other interim files that were created during the weighting process.
 - iii) Table 10 provides the frequency count of result codes for sampled households that were activated or should have been activated in the final weighted data file '*PHL-W01-STEP(HH) Weighted Data-Final Design Wts & Repl Wts(10-Mar-2016)*'.

Table 9: Variables in Final Weighted File '*PHL-W01-STEP(HH) Weighted Data-Final Design Wts & Repl Wts(10-Mar-2016)*'

Variable Name	Source	Description
Country	Sample File	Country abbreviation code
Cluster	Philippines Final Data File	Cluster
SMPLID_HH	Sample File	Sampled case ID#
PSUmpID	Sample File	PSU Sample ID#
STEP_Stratum	Sample File	Stratum
FrmeHHs	Sample File	HHs in Sample Frame (est.)
HHSmpid	Sample File	HH Sample Sequence within PSU, #s 01-30
HHsmpType	Sample File	Type of Sampled HH - Initial or Reserve
HHactOrd	Sample File	Sample activation order (HHsmpOrder in Sample File)
Resp_Gender	Philippines Final Data File	Gender of Selected Person
Resp_Age	Philippines Final Data File	Age of Selected Person
m1c_a	Philippines Final Data File	NUMBER A-# of eligible persons
m11_q07	Philippines Final Data File	Result code of interview
final_status	Philippines Final Data File	Final Result Code
PSUhhs_LSTd	Philippines Final Data File	# of Listed HHs in PSU
W1_PSUwt	Weighting Process	Basic PSU Weight
RespCat_HH	Weighting Process	Final Response Category-Households
W2_FinHHwt	Weighting Process	Final Household Weight
RespCat_SP	Weighting Process	Final Response Category-Selected Person
W3_FinSPwt	Weighting Process	Final Person Weight
jkgrp	Weighting Process	JackKnife Group #
hh_rpl01 to hh_rpl30	Weighting Process	30 Household Replicate Weights
sp_rpl01 to sp_rpl30	Weighting Process	30 Selected Person Replicate Weights

Table 10: Variables Created During Weighting Process – Not Added to Philippines Data File

Variable Name	Description	Values
W2_HH	Household Weight within PSU	n/a
W2_HHwt_basic	Basic Household Weight	n/a
PSU_RespFreq	Count of 'RespCat_HH' in PSU	n/a
HH_k	# of participant HHs in PSU	n/a
HH_x	# of Ineligible HHs in PSU	n/a
HH_z	# of Non-participant Eligible HHs in PSU	n/a
HH_u	# of Non-participant Unknown Eligible HHs in PSU	n/a
Elig_Prop_HH	Proportion of Eligible HHs - known eligibility	n/a
R_adj_HH	Response adjustment ratio for HHs	n/a
W3_SP	Selected Person weight within household	n/a

Table 11: Philippines, Frequency/Percent of Final Status Codes For Activated Sample Cases

Case Final Result Code	Frequency	Percent
01-Refusal-Household	105	3.3%
02-Unusual circumstance-Household	57	1.8%
03-No knowledgeable household member found after 3 visits	3	.1%
61-SP unable to read language of General Booklet (GB)	23	.7%
62-SP refusal to begin GB	66	2.1%
63-SP unable to do GB due to unusual circumstance	25	.8%
64-SP blindness or visual impairment-unable to do GB	36	1.1%
65-SP physical disability	16	.5%
71-SP looked through GB-unable to read/write language of GB	5	.2%
72-SP refused to continue GB after starting	87	2.7%
73-GB interruption too long	4	.1%
74-GB terminated - unusual circumstance	3	.1%
76-Attempted part or all of GB; <3 correct Core answers	83	2.6%
81-No attempt at any Exercise Booklet items / refusal to start	178	5.6%
82-Interruption too long during Exercise Booklet administration	228	7.2%
83-Unusual circumstance, Exercise Booklet administration terminated	14	.4%
84-Attempt at some Exercise Booklet items, no answers marked	34	1.1%
85-Attempt at all Exercise Booklet items, no answers marked	14	.4%
86-Marked some Exercise Booklet items, not all items attempted	21	.7%
87-Marked some Exercise Booklet items, all items attempted	142	4.5%
88-Marked all Exercise Booklet items	368	11.6%
Total	3165	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 '[W00-PHL Final Revised Data File\(7-Mar-2016\)](#)'.

5.1 First-stage Weight (PSU Weight)

The basic first stage weight, i.e., the PSU weight, is based on a sample size of 200 PSUs. Although 225 PSUs were initially selected to allow for 25 reserve PSUs for non-responding PSUs, only the 200 'initial' PSU sample was actually activated during data collection; none of the 25 'reserve' PSUs was used for the Philippines STEP data collection. Therefore, the STEP survey weight is based on the probability of selection for each of the 200 'initial' PSU sample.

For the Philippines systematic sample of PSUs, the probability of selection of the j^{th} PSU, say p_j , may be represented as,

$$1) \quad p_j = \frac{m}{M} = \frac{200}{84210} = 0.002375$$

where,

m - the number of sampled PSUs (i.e., m=200 PSUs in the case of the Philippines STEP Survey),

M - the total number of PSUs in the target population

The basic first stage weight for the j^{th} PSU, $W1_j$ (i.e., weighted data file variable '**W1_PSUwt**'), is calculated as the inverse of its probability of selection:

$$2) \quad W1_j = \frac{1}{p_j} = \frac{M}{m} = \frac{1}{0.002375} = 421.05$$

- **The basic first stage PSU weight is the same for each of the 200 sampled PSUs; it is appended to the STEP weighted data file.**

5.2 Second-stage Weight (Household Weight)

5.2.1 Basic Household Weight

The basic second stage weight for the k^{th} household in PSU 'j', $W2_{jk}$, is calculated as follows:

$$3) \quad W2_{jk} = \frac{N_j}{n_j}$$

where,

- N_j - the number of listed households (i.e., weighted data file variable 'PSUhhhs_LSTd') in the j^{th} PSU,
- n_j - the number of households sampled in the j^{th} PSU, (i.e., $n_j = 30$ households in each Philippines sampled PSU).

Example 1: Second Stage Basic Weight

The basic second stage weight for each of the 30 sampled households in the Philippines PSU # 1 which contains 115 listed households is calculated as follows:

$$W2_{jk} = \frac{N_j}{n_j} = \frac{115}{30} \approx 3.83$$

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

5.2.2 Second-stage Response Adjustment

The second-stage response adjustment was carried out for each sampled PSU, i.e., at the PSU level.

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 eligible 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, 51-88 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
81	No attempt at any of the exercise booklet items / refusal to start
82	Interruption too long during exercise booklet administration
83	Unusual circumstance, Exercise Booklet administration terminated
84	Attempt at some Exercise Booklet items but no answers marked
85	Attempt at all Exercise Booklet items but no answers marked
86	Marked some Exercise Booklet items, not all items attempted
87	Marked some Exercise Booklet items, all items attempted
88	Marked all Exercise Booklet items
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 ineligible 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
08	Vacant dwelling
09	Dwelling not habitable/dwelling destroyed/dwelling converted to commercial use
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
97	Sample selection error – a wrong household was selected in the field

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
07	Dwelling could not be found due to insufficient Sample Frame address info
95	Sample case not activated
96	General non-response; Reason unspecified by survey firm

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

$$4) \quad n_j = n_{jk} + n_{jx} + n_{jz} + n_{ju}$$

where,

n_j - the number of households sampled in the j^{th} PSU, (i.e., $n_j = 30$ households in each Philippines sampled PSU),

n_{jk} - the number of households ‘k’ that participate in the j^{th} PSU,

n_{jx} - the number of ineligible households ‘x’ in the j^{th} PSU,

n_{jz} - the number of non-participating households ‘z’ with known eligibility in the j^{th} PSU,

n_{ju} - the number of non-participating households ‘u’ with unknown eligibility in the j^{th} PSU.

Table 12 provides the STEP counts of activated households from the Philippines data file by the outcome categories described above. Note that the Philippines Step had no cases of ineligible dwellings/households and no cases of non-participant dwellings/households with known eligibility.

Table 12: Sample Count of 'Activated' Households by Final Household Response Category						
Case Final Status Code	Final Response Category-Households					% of Total
	Participant (n_{jk})	Ineligible (n_{ix})	Non- participant Eligible (n_{jz})	Non- participant Eligibility(??) (n_{ju})	Total (n_j)	
01-Refusal-Household				105	105	3.3%
02-Unusual circumstance-HH				57	57	1.8%
03-No knowledgeable HH member found				3	3	0.1%
61-SP unable to read General Booklet (GB) language	23				23	0.7%
62-SP refusal to begin GB	66				66	2.1%
63-SP unable to do GB due to unusual circumstance	25				25	0.8%
64-SP visual impairment-unable to do GB	36				36	1.1%
65-SP physical disability	16				16	0.5%
71-SP looked through GB-unable to read/write language of GB	5				5	0.2%
72-SP refused to continue GB	87				87	2.7%
73-GB interruption too long	4				4	0.1%
74-GB terminated - unusual circumstance	3				3	0.1%
75-Attempted GB, no answers marked	83				83	2.6%
76-Attempted GB; <3 correct Core answers	178				178	5.6%
81-No attempt at Exercise Booklet(EB) items	228				228	7.2%
82-Interruption too long-EB termination	14				14	0.4%
83-Unusual circumstance, EB terminated	34				34	1.1%
84-Attempt at some EB items, no answers marked	14				14	0.4%
85-Attempt at all EB items, no answers marked	21				21	0.7%
86-Marked some EB items, some items attempted	142				142	4.5%
87-Marked some EB items, all items attempted	368				368	11.6%
88-Marked all EB items	1653				1653	52.2%
Total	3000	0	0	165	3165	100.0%

5.2.2.1 Eligible Proportion of Households

The STEP assumes that the households 'u' with unknown eligibility in the j^{th} PSU are comprised of some proportion, say ε_j , that are eligible households and the complementary proportion, $1 - \varepsilon_j$, that are ineligible households.

Amongst the n_{ju} non-participating households 'u' with unknown eligibility in PSU 'j', the proportion of households with known eligibility is estimated to be

$$5) \quad \varepsilon_j = \frac{\text{Total \# of known eligible cases}}{\text{Total \# of known eligible \& known ineligible cases}} = \frac{n_{jk} + n_{jz}}{n_{jk} + n_{jx} + n_{jz}}$$

- Note that the eligible proportion is the same for all eligible participating households in a PSU. In the case of the Philippines Step, as illustrated in Example 3 below, $\varepsilon_j = 1$ for all cases.
- For the Philippines STEP, the final sample did not include any ineligible households and there were also no non-participating households with known eligibility.

Example 3: Eligible Proportion of Households

For all PSUs in the Philippines STEP, the following household results were obtained:

n_j	n_{jk}	n_{jx}	n_{jz}	n_{ju}
30	15	0	0	$0 \leq n_{ju} \leq 105$

For all PSUs in the Philippines STEP, the proportion of households with known eligibility is calculated as,

$$\begin{aligned} \varepsilon_j &= \frac{\text{Total \# of known eligible cases}}{\text{Total \# of known eligible \& known ineligible cases}} \\ &= \frac{n_{jk} + n_{jz}}{n_{jk} + n_{jx} + n_{jz}} \\ &= \frac{15 + 0}{15 + 0 + 0} = \frac{15}{15} \approx 1.0 \end{aligned}$$

5.2.2.2 Response Adjustment Factor 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 factor for the j^{th} PSU, R_j , is calculated as follows:

$$6) \quad R_{HHj} = \frac{\text{Total eligible households in PSU 'j'}}{\text{Total participating households in PSU 'j'}} = \frac{n_{jk} + n_{jz} + (\epsilon_j * n_{ju})}{n_{jk}}$$

- As illustrated in Example 4, the response adjustment ratio, R_{HHj} , is the same for all 200 sampled PSUs in the Philippines, i.e., $R_{HHj} = 2.0$.

Example 4: Second Stage Response Adjustment Factor for sampled Households

For example, the following results were obtained for all PSUs in the Philippines STEP

n_j	n_{jk}	n_{jx}	n_{jz}	n_{ju}
30	15	0	0	15

Therefore,

$$\begin{aligned} R_{HHj} &= \frac{\text{Total eligible households in PSU 'j'}}{\text{Total participating households in PSU 'j'}} \\ &= \frac{n_{jk} + n_{jz} + (\epsilon_j * n_{ju})}{n_{jk}} \\ &= \frac{15 + 0 + (1.0 * 15)}{15} \approx 2.0 \end{aligned}$$

5.2.3 Final Household Weight

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 final PSU weight, the basic household weight, and the response adjustment factor for non-participant households. Therefore, for household 'k' in PSU 'j', the final second stage household weight, $W2_{HHk}$, is calculated as follows:

$$7) \quad W2_{HHk} = W1_j * W2_{jk} * R_{HHj}$$

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

Example 5: Second Stage Response-adjusted Household Weight

For PSU #1,

$$\begin{aligned} W2_{HHj} &= W1_j * W2_{jk} * R_{HHj} \\ &= 421.05 \times 3.83 \times 2.0 \\ &\approx 3228.05 \end{aligned}$$

The household weight is similarly calculated for all sampled households.

5.2.4 Weighted Household Counts

The original sample frame 'PHL_F01 - BGY- SEGs Frame_08-Jun-2015' showed there is a total of 8,416,914 households in the target population.

Table 13 provides,

- i. the count of households from the sample frame of households for the 200 sampled PSUs;
- ii. the weighted count of households based on the final status of the sampled cases in the Philippines data file.

Source of PSU Household Count	# of PSUs	Weighted Sum of Households
1. Based on # of PSU Hhlds in Sample Frame 'PHL_F01 - BGY- SEGs Frame_08-Jun-2015'	84,210	8,416,914
2. Based on # of PSU Hhlds in STEP Data File 'STEP PHL - Data requirements for weighting_v4' [all activated sample cases, i.e., participants, non-participants]	200	8,704,367

For participating selected households, Table 14 provides a summary of the central tendency for the participating households.

Statistic	Final Household Weight
No. of Cases	3000
Mean	2901.46
Std. Error of Mean	5.786
Median	2863.14
Minimum	2021.04
Maximum	3564.89
Sum of Weights	8,704,367

5.3 Third-stage Weight (Person Weight)

5.3.1 Basic Third-stage Person Weight

The basic third stage weight for the i^{th} person, W_{3jki} , is calculated as follows:

$$8) \quad W_{3jki} = \frac{a_{jk}}{b_{jk}} = \frac{a_{jk}}{1} = a_{jk}$$

where,

a_{jk} - the number of eligible persons in the k^{th} household in the j^{th} PSU,

b_{jk} - the number of persons sampled in the k^{th} household in the j^{th} PSU, (i.e., $n_{jk} = 1$ person in each Philippines sampled household).

N.B. $A_{jk} = \sum_{k=1}^{n_{jk}} a_{jk} = \text{total \# of selected persons in the } j^{\text{th}} \text{ PSU}$
where,

n_{jk} - the number of households 'k' that participate in the j^{th} PSU

Example 6: Third Stage – Basic Person Weight within sampled Household

For PSU #1, selected household # 03 has 5 eligible persons. Therefore,

$$W_{3jki} = \frac{a_{jk}}{b_{jk}} = \frac{5}{1} = 5$$

5.3.2 Third-stage Person-level 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
81	No attempt at any of the exercise booklet items / refusal to start
82	Interruption too long during exercise booklet administration
83	Unusual circumstance, Exercise Booklet administration terminated
84	Attempt at some Exercise Booklet items but no answers marked
85	Attempt at all Exercise Booklet items but no answers marked
86	Marked some Exercise Booklet items, not all items attempted
87	Marked some Exercise Booklet items, all items attempted
88	Marked all Exercise Booklet items

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 ineligible 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

Note: there were no ineligible selected persons in the Philippines STEP Survey.

The third stage person weight is only relevant for sampled households in the final response category RespCat_HH=1, i.e., participating households. Table 15 provides the pertinent frequency counts of the Final Status Codes for the participant households.

Case Final Status Code	Frequency	Percent
61-SP unable to read language of General Booklet (GB)	23	.8%
62-SP refusal to begin GB	66	2.2%
63-SP unable to do GB due to unusual circumstance	25	.8%
64-SP blindness or visual impairment-unable to do GB	36	1.2%
65-SP physical disability	16	.5%
71-SP looked through GB-unable to read/write language of GB	5	.2%
72-SP refused to continue GB after starting	87	2.9%
73-GB interruption too long	4	.1%
74-GB terminated - unusual circumstance	3	.1%
75-Attempted GB, no answers marked	83	2.8%
76-Attempted part or all of GB; <3 correct Core answers	178	5.9%
81-No attempt at any Exercise Booklet items / refusal to start	228	7.6%
82-Interruption too long during Exercise Booklet administration	14	.5%
83-Unusual circumstance, Exercise Booklet administration terminated	34	1.1%
84-Attempt at some Exercise Booklet items, no answers marked	14	.5%
85-Attempt at all Exercise Booklet items, no answers marked	21	.7%
86-Marked some Exercise Booklet items, not all items attempted	142	4.7%
87-Marked some Exercise Booklet items, all items attempted	368	12.3%
88-Marked all Exercise Booklet items	1653	55.1%
Total	3000	100.0%

Note: For the Philippines STEP Survey, there were no non-participant selected persons; all 3000 selected persons were categorized as participants.

Table 16 provides the sample counts of the Final Response Category for selected persons.

For Participating Households		
Final Response Category-SP	Frequency	Percent
1-Participant SP	3000	100.0
2-Non-participant SP, Eligible	0	0.0
Total	3000	100.0

Based on the above table, there are 3000 participating selected persons for which a final non-zero Person-weight was calculated.

5.3.3 Final Third-stage Person Weight

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_{3SP} (i.e., weighted data file variable 'W_FinSPwt'), is calculated as follows:

$$9) \quad W_{3SP} = W_{2HH_k} * W_{3jki} * R_{SP_j}$$

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

Example 8: Third Stage – Final Person Weight

- i) For PSU #1, selected household # 03 with 5 eligible persons, the selected person was a participant. All selected persons participated in PSU #1, therefore, $R_{SP_j} = 1$

$$W_{2H} = 3228.05 * 5 * 1 \approx 16140.25$$

The final person weight is similarly calculated for all sampled persons.

5.3.4 Weighted Person Counts

The weighted count of the non-institutionalized persons 15 to 64 years of age (inclusive) living in private dwellings in urban areas of **Philippines** at the time of data collection is provided in Table 17. Based on the 3000 participating selected persons, the sum of the final person weights from the STEP survey is 1,185,359; the sum of the benchmark-adjusted final person weights from the STEP survey is 1,355,817

Table 17: Sample Frequency & Weighted Persons Totals by Final Selected Person Response Category For Participating Households		
Response Category (SP)	Sampled Cases	Total # of Eligible Persons in Target Population
		(based on Final Person weight)
1-Participant SP	3,000	22,902,060
2-Non-participant, Eligible	0	0
Total	3,000	22,902,060

For participating selected persons, Table 18 provides a summary of the central tendency of the number of eligible persons in the participating households. Based on the average of 2.77 eligible persons per participating household and an estimated 426,484 households, there are approximately 1,185,500 persons aged 15 to 64 in the target population.

Table 18: Summary - Person Weights	
Statistic	Final Person Weight
No. of Cases	3000
Mean	7634.02
Std. Error of Mean	73.42
Median	6399.96
Minimum	2021.04
Maximum	31438.40
Sum of Weights	22,902,060

6 Response Rate

The STEP response rate, R, is calculated as follows:

$$14) \quad R = \frac{P}{n - X} * 100\%$$

where,

P is the number of participating persons aged 15 to 64,

n is the number of activated sample cases,

X is the number of ineligible sample cases.

An overall response rate of 94.8% was achieved in the Philippines STEP Survey. Table 19 provides the detailed percentage distribution by final status code.

Table 19: Response Rate: Philippines, Activated HHs by Final Status Code

Case Final Status Code	Frequency	Percent
08-Vacant dwelling / selected address has no household	0	0.0%
09-Dwelling not habitable/destroyed/converted to commercial	0	0.0%
31-No eligible household member aged 15-64	0	0.0%
Ineligible HHs (N.B. % is based on total # of activated HHs, i.e., 3165 HHs)	0	0.0%
01-Refusal-Household	105	3.3%
02-Unusual circumstance-Household	57	1.8%
03-No knowledgeable household member found after 3 visits	3	0.1%
04-Temporarily absent/unavailable for field period	0	0.0%
05-No competent household member (e.g., severe illness, mental disability, etc.)	0	0.0%
06-Language problem	0	0.0%
07-Dwelling could not be found - insufficient Sample Frame address info	0	0.0%
21-Household Refusal to continue Household Module (begun, not complete)	0	0.0%
32-Refusal by SP to begin modules 2-7	0	0.0%
34-Refusal-Other Household Member refused to let SP participate	0	0.0%
35-Unable to contact SP after 3 attempts	0	0.0%
36-SP absent for entire survey period	0	0.0%
38-SP learning or mental disability	0	0.0%
97-Household not interviewed - field procedure error	0	0.0%
98-Incorrect eligible household member selected	0	0.0%
Non-respondents (Eligible) (N.B. % is based on total # of Eligible HHs, i.e., 3165 HHs)	165	5.2%
51-Refusal by SP to continue modules 2-7	23	0.7%
52-Interview termination-SP unusual circumstance	66	2.1%
61-SP unable to read language of General Booklet (GB)	25	0.8%
62-SP refusal to begin GB	36	1.1%
63-SP unable to do GB due to unusual circumstance	16	0.5%
64-SP blindness or visual impairment-unable to do GB	5	0.2%
65-SP physical disability	87	2.7%
72-SP refused to continue GB after starting	4	0.1%
74-GB terminated - unusual circumstance	3	0.1%
76-Attempted part or all of GB; <3 correct Core answers	83	2.6%
81-No attempt at any Exercise Booklet items / refusal to start	178	5.6%
82-Interruption too long during Exercise Booklet administration	228	7.2%
83-Unusual circumstance, Exercise Booklet administration terminated	14	0.4%
84-Attempt at some Exercise Booklet items, no answers marked	34	1.1%
85-Attempt at all Exercise Booklet items, no answers marked	14	0.4%
86-Marked some Exercise Booklet items, not all items attempted	21	0.7%
87-Marked some Exercise Booklet items, all items attempted	142	4.5%
88-Marked all Exercise Booklet items	368	11.6%
Respondents (N.B. % is based on total # of Eligible HHs, i.e., 5952 HHs)	3000	94.8%
Total <u>Eligible</u> Households	3165	100.0%
Total <u>Activated</u> Households	3165	100.0%

7 Jackknife Weights

The Jackknife Replication Method *JK1* was used to generate a set of 30 replicate weights for the Philippines STEP Survey. *Wesvar* software was used to create the initial replicates, and to adjust the initial replicate weights for non-response.

Table 20: STEP # of PSUs by JKgrp for Philippines

<i>JKgrp</i>	# of PSUs
1	6
2	7
3	7
4	6
5	7
6	7
7	6
8	7
9	7
10	6
11	7
12	7
13	7
14	6
15	7
16	7
17	6
18	7
19	7
20	6
21	7
22	7
23	7
24	6
25	7
26	7
27	6
28	7
29	7
30	6
TOTAL	200

8 Using the Jackknife Weights

The jackknife replicate weights are used to calculate the standard errors of the full-sample survey estimates. Regardless of the sample design used by a country, the jackknife replication method uses one common formula for the calculation of the standard error (s.e.) of a survey estimate,

$$s.e.(\hat{\theta}) = \sqrt{v(\hat{\theta})} = \sqrt{c \sum_{g=1}^G f_g h_g (\hat{\theta}_g - \hat{\theta})^2} \dots\dots\dots 4.0$$

where,

$\hat{\theta}$ is the full-sample estimate of a population parameter θ , e.g., θ could be an average, a total, a proportion, etc.

$v(\hat{\theta})$ is the variance of the full-sample estimate of the population parameter θ ,

$\hat{\theta}_g$ is the estimate of θ based on the observations included in the g^{th} replicate,

G is the total number of replicates formed,

$$c = \begin{cases} \frac{G-1}{G} & \text{for JK1 method} \\ 1 & \text{for JKn method} \end{cases}$$

f_g is the finite population correction (fpc) factor. N.B., the fpc can be ignored if less than 20% of the population is sampled in a stratum.

$$h_g = \begin{cases} 1 & \text{for unstratified sample designs} \\ \frac{n_{h'}-1}{n_{h'}} & \text{for stratified sample designs} \end{cases}$$

STEP Equation for Calculation of Standard Errors for All Participating Wave 1-2 Countries

Notes:

- The software *Wesvar* automatically incorporates the stratum factors, h_g , in the replicate weights.
- The fpc is not applicable for the STEP country datasets.

Therefore, equation 4.0 may be reduced to the following equation 4.1 for the calculation of standard errors of the STEP survey estimates for the STEP country datasets:

$$s.e.(\hat{\theta}) = \sqrt{c \sum_{g=1}^G (\hat{\theta}_g - \hat{\theta})^2} \dots\dots\dots 4.1$$

where,

$$c = \begin{cases} \frac{G-1}{G} & \text{for JK1 method} \\ 1 & \text{for JKn method} \end{cases}$$

8.1 Example: Philippines

The Philippines sample design did not incorporate a stratification strategy. For the Jackknife weighting process, the 200 PSUs in the Philippines sample were allocated to 30 replicate groups.

With regard to the calculation of the standard errors of STEP survey estimates for the Philippines, the following factor value is applicable,

$$1) \quad c = \frac{G-1}{G} = \frac{29}{30}$$

Therefore, for the Philippines, the calculation of the standard errors of STEP survey estimates can be carried out using the following equation:

$$s.e.(\hat{\theta}) = \sqrt{\frac{29}{30} \sum_{g=1}^{30} (\hat{\theta}_g - \hat{\theta})^2}$$

To illustrate the procedure for calculating the standard error of an estimate using the replicate information, suppose it is desired to obtain $\hat{\theta}$, the survey estimate of the average age of persons in the the Philippines STEP target population. The standard error calculation may be summarized in 4 steps as follows.

Step 1: Calculate $\hat{\theta}$ from full sample of respondents.

$$\hat{\theta} = \frac{\sum_1^n w_i x_i}{\sum_1^n w_i} = 37.78, \quad \text{i.e., calculated from the STEP the Philippines full sample of respondents,}$$

where,

- x_i is the age for responding individual 'i' in the full sample of STEP respondents,
- w_i is the final person weight (i.e., W_FinSPwt) for responding individual 'i' in the full sample of the Philippines STEP respondents,
- n is the number of responding individuals

Step 2: Calculate $\hat{\theta}_g$ for each of the $g=30$ replicates.

$$\hat{\theta}_g = \frac{\sum_1^{n_g} w_{i,g} x_{i,g}}{\sum_1^{n_g} w_{i,g}}, \quad \text{for the respondents in each replicate 'g' of the the Philippines STEP sample of respondents,}$$

where,

- $x_{i,g}$ is the age for responding individual 'i' in the replicate 'g' sample of STEP respondents,
- $w_{i,g}$ is the replicate 'g' person weight for responding individual 'i' in the replicate 'g' sample of STEP respondents
- n_g is the number of responding individuals in replicate 'g'

Column 'c' in Table 4 includes the calculated values for $\hat{\theta}_g$.

The following table summarizes the Philippines inputs for the standard error equation.

Table 10: Philippines Inputs for Standard Error Calculation				
Weight	Average Age Estimate		$\hat{\theta}_g - \hat{\theta}$	$(\hat{\theta}_g - \hat{\theta})^2$
(a)	(b)	(c)	(d)	(e)
Full Sample	$\hat{\theta}$	37.78	-	-
Replicate 01	$\hat{\theta}_1$	37.77	-0.0030140067988285	0.0000090842369834
Replicate 02	$\hat{\theta}_2$	37.76	-0.0222562447085721	0.0004953404285278
Replicate 03	$\hat{\theta}_3$	37.75	-0.0242301694356541	0.0005871011108805
Replicate 04	$\hat{\theta}_4$	37.72	-0.0602618600483709	0.0036314917764894
Replicate 05	$\hat{\theta}_5$	37.81	0.0355805480847664	0.0012659754020124
Replicate 06	$\hat{\theta}_6$	37.77	-0.0091768495528015	0.0000842145677148
Replicate 07	$\hat{\theta}_7$	37.71	-0.0690943088317013	0.0047740235129305
Replicate 08	$\hat{\theta}_8$	37.81	0.0329274700403488	0.0010842182832581
Replicate 09	$\hat{\theta}_9$	37.78	0.0063086103729901	0.0000397985648382
Replicate 10	$\hat{\theta}_{10}$	37.67	-0.1035507053479140	0.0107227485780504
Replicate 11	$\hat{\theta}_{11}$	37.71	-0.0672229757207958	0.0045189284647587
Replicate 12	$\hat{\theta}_{12}$	37.68	-0.1005328773205820	0.0101068594223552
Replicate 13	$\hat{\theta}_{13}$	37.85	0.0748483091391918	0.0056022693809960
Replicate 14	$\hat{\theta}_{14}$	37.83	0.0552647942392781	0.0030541974823098
Replicate 15	$\hat{\theta}_{15}$	37.70	-0.0771053206080197	0.0059452304660655
Replicate 16	$\hat{\theta}_{16}$	37.83	0.0494891148353105	0.0024491724871826
Replicate 17	$\hat{\theta}_{17}$	37.82	0.0395346625096167	0.0015629895397493
Replicate 18	$\hat{\theta}_{18}$	37.76	-0.0195538086587703	0.0003823514330638
Replicate 19	$\hat{\theta}_{19}$	37.84	0.0605793734126081	0.0036698604830642
Replicate 20	$\hat{\theta}_{20}$	37.81	0.0302257111469615	0.0009135936143396
Replicate 21	$\hat{\theta}_{21}$	37.84	0.0672920861766428	0.0045282248620047
Replicate 22	$\hat{\theta}_{22}$	37.81	0.0284995417237539	0.0008122238784640
Replicate 23	$\hat{\theta}_{23}$	37.85	0.0768105240083159	0.0058998565984321
Replicate 24	$\hat{\theta}_{24}$	37.82	0.0431676012407323	0.0018634417968789
Replicate 25	$\hat{\theta}_{25}$	37.71	-0.0665673988141720	0.0044312185848850
Replicate 26	$\hat{\theta}_{26}$	37.77	-0.0091796491686935	0.0000842659588603
Replicate 27	$\hat{\theta}_{27}$	37.82	0.0470401822091091	0.0022127787422662
Replicate 28	$\hat{\theta}_{28}$	37.76	-0.0199249784145294	0.0003970047648195
Replicate 29	$\hat{\theta}_{29}$	37.76	-0.0174759359494345	0.0003054083373087
Replicate 30	$\hat{\theta}_{30}$	37.80	0.0212429262518938	0.0004512619157434
TOTAL				0.0818851346752328

Step 3: Calculate $(\hat{\theta}_g - \hat{\theta})^2$, the squared differences between $\hat{\theta}_g$ and $\hat{\theta}$.

Column 'e' in Table 4 includes the calculated values for $(\hat{\theta}_g - \hat{\theta})^2$.

Step 4: Calculate the standard error of the average age estimate, $\hat{\theta}$, using the equation,

$$s.e.(\hat{\theta}) = \sqrt{\frac{29}{30} \sum_{g=1}^{30} (\hat{\theta}_g - \hat{\theta})^2}$$

Using the squared sum of differences total of column 'e' from Table 4,

$$s.e.(\hat{\theta}) = \sqrt{\frac{29}{30} (0.0818851346752328)} = 0.2813461$$