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

# Georgia

# December 27, 2013

# Table of Contents

# 1    Survey Design Overview

The Georgia 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 Georgia 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.

**Exclusions**

- Residents of institutions (prisons, hospitals, etc);
- Residents of senior homes and hospices;
- Residents of other group dwellings such as college dormitories, halfway homes, workers' quarters, etc.;
- Persons living outside the country at the time of data;
- Persons living in the conflict regions Abkhazia and South Ossetia.

## 1.2   Sample Frame

The sample frame for the selection of first stage sample units is the file 'CRRC_SamplingFrame_09112012_CapitalUrban'. The sample frame is a list of 1553 first stage sampling units, i.e., Electoral Precincts, for the 2012 parliamentary elections in Georgia.

Table 1 provides the sample frame count of the number of PSUs & the estimated number of persons aged 15-64 by stratum.

**Table 1: Georgia Sample Frame Count of # of PSUs & # of Persons Aged 15-64 by Stratum**

| Stratum | # of PSUs (Electoral Precincts) | Estimated # of Persons Aged 15-64 |
|---|---|---|
| **Capital** | 718 | 792,672 |
| **Other Urban NE** | 146 | 149,879 |
| **Other Urban NW** | 373 | 373,680 |
| **Other Urban SE** | 169 | 171,117 |
| **Other Urban NW** | 147 | 150,326 |
| **Total** | **1,553** | **1,637,674** |

## 1.3   Sample Design

The Georgia sample design is a stratified 3 stage sample design. The sample was stratified by 5 geographic areas: Capital, Other Urban Northeast, Other Urban Northwest, Other Urban Southeast, and Other Urban Southwest.

First Stage Sample
The primary sample unit (PSU) is an Electoral Precinct; each PSU is uniquely defined by the sample frame variable 'PrecinctID'.

The first stage units were selected by the World Bank Survey Methodologist. The sampling objective was to select 225 PSUs, comprised of 200 Initial PSUs and 25 Reserve PSUs. Although 225 PSUs were selected, only 200 PSUs were activated; none of the 25 reserve PSUs was activated during data collection.

The PSUs were selected using a systematic probability proportional to size (PPS) sampling method, where the measure of size was the estimated number of persons aged 15 to 64 (i.e., final data file variable '*estimatedsize*') in a PSU.

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

For the 'Other Urban' strata, in order to provide sufficient sample to allow for a scenario of a 50% response rate the number of sampled cases was doubled in each selected 'Other  Urban' Electoral Precinct. In other words, a reserve sample of 15 households per PSU was selected in each 'Other Urban' stratum.

For the Capital stratum, the initial sample size was adjusted to allow for a possible expected response rate of 44%. In the case of the Capital stratum, a reserve sample of 19 households per PSU was selected. Also, during the survey fieldwork, there were 14 PSUs in the Capital stratum where, due to higher than expected non-response, the survey firm drew an additional 34 reserve households; the PrecinctIDs for the 14 PSUs are: 2015, 2025, 2033, 2060, 2076, 3003, 3012, 3022, 3029, 3071, 3083, 4014, 5070, and 6106.

Table 2 provides the overall sample allocation for the household sample for the activated PSUs.

| Table 2:   Georgia Overall Sample Allocation for 'Activated' PSUs by Stratum | | | | |
|---|---|---|---|---|
| Stratum | # of PSUs | Initial HH Sample | Reserve HH  Sample | Total |
| Capital | 97 | 1455 | 2319 | **3774** |
| Other Urban NE | 18 | 270 | 270 | **540** |
| Other Urban NW | 46 | 690 | 690 | **1380** |
| Other Urban SE | 21 | 315 | 315 | **630** |
| Other Urban NW | 18 | 270 | 270 | **540** |
| Total | **200** | **3000** | **3388** | **6864** |

Third 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.

## 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 3. A variable called '**final_status**' with these code values was created during the weighting process and a '**final_status**' was assigned to each sampled case.

| Table 3: *Full Literacy Assessment* 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 / selected address has no household |
| 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 not 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 3: *Full Literacy Assessment* 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 Georgia three-stage stratified cluster 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 sampling weight for each person that participated in the survey.

In general, the objectives of the 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.

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

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.

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

# 4 Files - Weighting Process

## 4.1 Input Files

The key file for the weighting process is the file 'STEP Georgia - Data for Weighting_S7.dta' provided by the WB core team. The file 'STEP Georgia - Data for Weighting_S7.dta' contains a data record for each of the 6,864 sampled households from the Georgia sample frame. An SPSS version of this file, i.e., 'STEP Georgia - Data for Weighting_S7.sav', was created for the weighting process.

### 4.1.1 Revisions to Input Data File

The following general revisions were made to the Georgia SPSS data file 'STEP Georgia - Data for Weighting_S7':

1) '*SmplID*', sample case sequential identification number from 1 to 6388 and 9001 to 9476
   a) For cases where the '*SmplID*' was missing in the final data file, the appropriate '*SmplID*' was obtained from the sample files.

2) '*m11_q07*', the Result of Interview Code (N.B. the survey firm assigned codes 21 to 88 only)
   a) added value labels; the value labels for '*m11_q07*' are as follows:

| Table 4: VALUE LABELS for the variable 'm11_q07' | |
|---|---|
| 00-No assigned code | 71-SP looked through GB-unable to read & write language of GB' |
| 21-Household Refusal to continue Household Module (i.e., begun but not complete)' | 72-SP refused to continue GB after starting' |
| 22-Household module terminated due to unusual circumstance' | 73-GB interruption too long' |
| 31-No eligible household member aged 15-64' | 74-GB terminated due to unusual circumstance' |
| 32-Refusal-Refusal by SP to begin modules 2-7' | 75-Attempted GB but no answers marked' |
| 33-Interview termination-SP unusual circumstance' | 76-Attempted part or all of GB; less than 3 correct Core answers' |
| 34-Refusal-Other Household Member refused to let SP participate' | 77-Marked some items & attempted all items in GB' |
| 35-Unable to contact SP after 3 attempts' | 78-Marked all items in GB' |
| 36-SP absent for entire survey period' | 81-No attempt at any exercise booklet items / refusal to start' |
| 37-SP deaf or gard of hearing' | 82-Interruption too long during exercise booklet administration' |
| 38-SP learning or mental disability' | 83-Unusual circumstance, Exercise Booklet administration terminated' |
| 39-SP speech impairment' | 84-Attempt at some Exercise Booklet items, no answers marked' |
| 40-SP language problem' | 85-Attempt at all Exercise Booklet items, no answers marked' |
| 51-Refusal by SP to continue modules 2-7' | 86-Marked some Exercise Booklet items, not all items attempted' |
| 52-Interview termination-SP unusual circumstance' | 87-Marked some Exercise Booklet items, all items attempted' |
| 61-SP unable to read language of General Booklet (GB)' | 88-Marked all Exercise Booklet items' |
| 62-SP refusal to begin GB' | |
| 63-SP unable to do GB due to unusual circumstance' | |
| 64-SP blindness or visual impairment-unable to do GB' | |
| 65-SP physical disability' | |

3) '*result*', the code for the Overall Outcome of the Interview (N.B., the survey firm assigned codes 1 to 88):
   a) added value labels; the value labels for '*result*' are as follows:

| Table 5: VALUE LABELS for the variable 'm11_q07' | |
|---|---|
| 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 / selected address has no household | 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' |
| 21-Household Refusal to continue Household Module (i.e., begun but not complete)' | 74-GB terminated due to unusual circumstance' |
| 22-Household module terminated due to unusual circumstance' | 75-Attempted GB but no answers marked' |
| 31-No eligible household member aged 15-64' | 76-Attempted part or all of GB; less than 3 correct Core answers' |
| 32-Refusal-Refusal by SP to begin modules 2-7' | 77-Marked some items & attempted all items in GB' |
| 33-Interview termination-SP unusual circumstance' | 78-Marked all items in GB' |
| 34-Refusal-Other Household Member refused to let SP participate' | 81-No attempt at any exercise booklet items / refusal to start' |
| 35-Unable to contact SP after 3 attempts' | 82-Interruption too long during exercise booklet administration' |
| 36-SP absent for entire survey period' | 83-Unusual circumstance, Exercise Booklet administration terminated' |
| 37-SP deaf or gard of hearing' | 84-Attempt at some Exercise Booklet items, no answers marked' |
| 38-SP learning or mental disability' | 85-Attempt at all Exercise Booklet items, no answers marked' |
| 39-SP speech impairment' | 86-Marked some Exercise Booklet items, not all items attempted' |
| 40-SP language problem' | 87-Marked some Exercise Booklet items, all items attempted' |
| | 88-Marked all Exercise Booklet items' |

4) New Variables needed for weighting calculations

   a) **final_status**: Final status code for sampled cases; used to create response categories for sampled households and sampled persons.

      i) Refer to Table 3 in section 2 for the list of **final_status** codes.

      ii) The data file variables '*result* - Overall Outcome of the Interview' and '*field_rst* - Fieldwork result code ' were used to create the variable 'final_status'.
         ◆ For '*result*', the description of the codes is included in Table 5 above.
         ◆ For , '*field_rst*' the description of the codes is as follows:

| Code | Code Description |
|---|---|
| 1 | Activated; for an 'initial sample' HH this means that it was contacted, for a 'reserve sample' HH it means that it was activated for an interview attempt. |
| 2 | Never activated; applicable to reserve sample HHs, the case was correctly never activated, i.e., the prescribed fieldwork procedures were followed. |

| | |
|---|---|
| 3 | Wrongly activated; applicable to reserve sample cases, the reserve case was activated, but according to the prescribed fieldwork procedures it should <u>not</u> have been activated (i.e., a reserve sample HH was incorrectly activated). |
| 4 | Wrongly never activated; applicable to reserve sample cases, the reserve case was never activated but according to the prescribed fieldwork procedures it should have been activated (i.e., a reserve HH should have been activated based on the provided activation order, but it wasn't). |
| 5 | Incorrect eligible HH member selected; during the random selection process, an incorrect household member was selected from the list of <u>eligible</u> members. |
| 6 | Ineligible HH member selected; during the random selection process, an <u>ineligible</u> household member was selected. |

iii) The specifications for creating the variable '**final_status**' are as follows;

- IF ((result ≠ 0) & (field_rst = 1 or field_rst = 3))then  final_status=result;
- IF (field_rst = 2) then final_status=95;
- IF (field_rst = 4) then final_status=97;
- IF (field_rst = 5) then final_status=98;
- IF (field_rst = 6) then final_status=99.

b) **EligPers_Cnt**: Number of <u>selected</u> Persons in the sampled Household; used in place of the variable '*m1c_a_clean*'; there was one case with a missing value for the variable '*m1c_a_clean*' (i.e., NUMBER A: Eligible Persons).

  i) **EligPers_Cnt**: Number of <u>selected</u> Persons in the sampled Household
  - for data records with <u>non-missing</u> variable '*m1c_a_clean*',
    - **EligPers_Cnt**= '*m1c_a_clean*'
  - for data records with <u>missing</u> variable '*m1c_a_clean*',
    - **EligPers_Cnt**= *m1c_a_median*, the PSU median value for non-missing cases of the variable '*m1c_a_clean*'

### 4.1.1.1 Input File – Revised for Weighting

The SPSS Revised Input File for weighting is 'W00-GEO Final Revised Data File(19-Dec-2013)'. This file is the primary file for the process of creating the survey weights; it includes all 6864 sampled cases; each case includes a final status code. Table 5 provides the list of key variables for weighting purposes that are included in the file.

| Table 5: Key Variables in Primary Input File for Weighting Process | | |
|---|---|---|
| **Variable Name** | **Source** | **Description** |
| **Country** | Sample File | Country abbreviation code |
| **SmplID** | Sample File | Sampled case ID# |
| **PSUsmpID** | Sample File | PSU Sample ID# |
| **PSUsmp_Type** | Sample File | PSU Sample Type (Initial, Reserve) |
| **HHsmpOrder** | Sample File | Sample activation order |
| **hhsmpordersf** | Final Data File | HHsmpOrder used by the survey firm |
| **HHsmpSeq** | Sample File | Household Sample Sequence #s 01 to 30 |
| **HHsmpType** | Sample File | HH Sample Indicator |
| **precinctid** | Sample Frame | Electoral Precinct ID# |
| **Region** | Sample Frame | Region |
| **macrostrata** | Sample Frame | Stratum |
| **estimatedsize** | Sample Frame | Estimated number of persons aged 15 to 64 |
| **HH_count_total** | Final Data File | Number of listed households in PSU |
| **result** | Final Data File | Overall outcome of the interview |
| **m1c_a_clean** | Final Data File | NUMBER A-# of eligible persons |
| **m11_q07** | Final Data File | Result code of interview |
| **field_rst** | Final Data File | Fieldwork result code |
| **final_status** | Created during weighting | Final Result Code |

Table 6 provides the breakdown of the 6864 sample cases by the final result of the interview attempt, i.e., by **'final_status'** code.

**Table 6: Georgia: Number & Percentage of Sampled Dwellings by Final Result Code of Interview 'final_status'**

| Final Result of Interview, i.e., *'final_status'* | # of Sample Cases | Percent of Sample Cases |
|---|---|---|
| 01-Refusal-Household | 1103 | 16.1% |
| 02-Unusual circumstance-Household | 17 | .2% |
| 03-No knowledgeable household member found after 3 visits | 73 | 1.1% |
| 04-Temporarily absent/unavailable for field period | 242 | 3.5% |
| 05-No competent household member (e.g., severe illness, mental disability, etc.) | 14 | .2% |
| 06-Language problem | 28 | .4% |
| 07-Dwelling could not be found / selected address has no household | 93 | 1.4% |
| 08-Vacant dwelling / selected address has no household | 255 | 3.7% |
| 09-Dwelling not habitable/destroyed/converted to commercial | 50 | .7% |
| 21-Household Refusal to continue Household Module (begun, not complete) | 5 | .1% |
| 22-Household module terminated - unusual circumstance | 1 | .0% |
| 31-No eligible household member aged 15-64 | 218 | 3.2% |
| 32-Refusal by SP to begin modules 2-7 | 12 | .2% |
| 33-Interview termination-SP unusual circumstance | 1 | .0% |
| 35-Unable to contact SP after 3 attempts | 4 | .1% |
| 36-SP absent for entire survey period | 13 | .2% |
| 38-SP learning or mental disability | 1 | .0% |
| 51-Refusal by SP to continue modules 2-7 | 4 | .1% |
| 52-Interview termination-SP unusual circumstance | 1 | .0% |
| 61-SP unable to read language of General Booklet (GB) | 44 | .6% |
| 62-SP refusal to begin GB | 250 | 3.6% |
| 63-SP unable to do GB due to unusual circumstance | 3 | .0% |
| 64-SP blindness or visual impairment-unable to do GB | 24 | .3% |
| 65-SP physical disability | 4 | .1% |
| 71-SP looked through GB-unable to read/write language of GB | 14 | .2% |
| 72-SP refused to continue GB after starting | 39 | .6% |
| 73-GB interruption too long | 3 | .0% |
| 74-GB terminated - unusual circumstance | 3 | .0% |
| 75-Attempted GB, no answers marked | 5 | .1% |
| 76-Attempted part or all of GB; <3 correct Core answers | 55 | .8% |
| 81-No attempt at any Exercise Booklet items / refusal to start | 104 | 1.5% |
| 82-Interruption too long during Exercise Booklet administration | 8 | .1% |
| 83-Unusual circumstance, Exercise Booklet administration terminated | 3 | .0% |
| 84-Attempt at some Exercise Booklet items, no answers marked | 28 | .4% |
| 85-Attempt at all Exercise Booklet items, no answers marked | 13 | .2% |
| 86-Marked some Exercise Booklet items, not all items attempted | 655 | 9.5% |
| 87-Marked some Exercise Booklet items, all items attempted | 779 | 11.3% |
| 88-Marked all Exercise Booklet items | 957 | 13.9% |
| 95-Household not activated | 1462 | 21.3% |
| 97-Household not interviewed - field procedure error | 271 | 3.9% |
| 98-Incorrect eligible household member selected | 5 | .1% |
| Total | 6864 | 100.0% |

## 4.2 Final Output Files from Weighting Process

The final output file from the weighting process is,

1) GEO-W01-STEP(HH) Weighted Data-Final Design Weights (26-Dec-2013)

   a) This file contains the original data variables from the Georgia data file 'STEP Georgia - Data for Weighting_S7.dta' and the applicable PSU weights, household weights, person weights, and other associated variables that were created during the weighting process.

   b) The file contains a data record for each of the 5402 cases that were activated or should have been activated during the data collection operation.
      i) 3032 participating households with a final household weight
         ◆ For participating households, the weighted sum of the household weights is 503,417.
      ii) 2996 participating selected persons with a final person weight
         ◆ For participating individuals, the weighted sum of the person weights is 1,346,787.

   c) In addition to the Georgia data file original variables, the weighted data file includes several 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 7 provides a list of additional variables from the weighting process that are included in the final weighted data file.

      ii) Table 8 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 9 provides the frequency count of final status codes for the 5402 sampled households that were activated or should have been activated in the final weighted data file 'GEO-W01-STEP(HH) Weighted Data-Final Design Weights (26-Dec-2013)'.

**Table 7: Variables Created During Weighting Process - *Added to Georgia Data File***

| Variable Name | Description | Values |
|---|---|---|
| m1c_a_clean_Med | Median # Eligible Persons in Stratum | n/a |
| EligPers_Cnt | Eligible Person Count in Participant HH | n/a |
| final_status | Final Result Code | Refer to Table 2 for code descriptions |
| RespCat_HH | Final Response Category-Households | 1-Participant HH<br>2-Ineligible HH<br>3-Non-participant Eligible HH<br>4-Non-participant Unknown Eligibility HH |
| RespCat_SP | Final Response Category-Selected Person | 1-Participant SP<br>2-Non-participant SP<br>3-Not applicable - non-participant HH |
| W1_PSU | Basic PSU Weight | n/a |
| W2_FinHHwt | Final Household Weight<br>Use for estimation of Household characteristics | n/a |
| W3_FinSPwt | Final Person Weight | n/a |

**Table 8: Variables Created During Weighting Process – Not *Added to Georgia Data File***

| Variable Name | Description | Values |
|---|---|---|
| St_estpopn_15to64 | Stratum estimated target population size | n/a |
| ST_PSUs | Stratum Number of PSUs | n/a |
| PSU_SmplSz | PSU Sample Size - HHs | n/a |
| 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 |
| EligPers_Cnt | Eligible Person Count in Sampled HH | n/a |
| W3_SP | Selected Person n weight within household | n/a |
| R_adj_SP | Response adjustment for selected person | n/a |
| EP_PSU_RCat | Total Eligible Persons-PSU Participant Households | n/a |
| EP_k | PSU - # of Participant SPs in participant HHs | n/a |
| EP_z | PSU - # of Non-participant SPs in participant HHs | n/a |
| EP_PSU | PSU - Total Eligible Persons | n/a |

**Table 9: Georgia, Frequency/Percent of Final Status Codes For Activated Sample Cases**

| Case Final Result Code | Frequency | Percent |
|---|---|---|
| 01-Refusal-Household | 1103 | 20.4% |
| 02-Unusual circumstance-Household | 17 | .3% |
| 03-No knowledgeable household member found after 3 visits | 73 | 1.4% |
| 04-Temporarily absent/unavailable for field period | 242 | 4.5% |
| 05-No competent household member (e.g., severe illness, mental disability, etc.) | 14 | .3% |
| 06-Language problem | 28 | .5% |
| 07-Dwelling could not be found - insufficient Sample Frame address info | 93 | 1.7% |
| 08-Vacant dwelling / selected address has no household | 255 | 4.7% |
| 09-Dwelling not habitable/destroyed/converted to commercial | 50 | .9% |
| 21-Household Refusal to continue Household Module (begun,not complete) | 5 | .1% |
| 22-Household module terminated - unusual circumstance | 1 | .0% |
| 31-No eligible household member aged 15-64 | 218 | 4.0% |
| 32-Refusal by SP to begin modules 2-7 | 12 | .2% |
| 33-Interview termination-SP unusual circumstance | 1 | .0% |
| 35-Unable to contact SP after 3 attempts | 4 | .1% |
| 36-SP absent for entire survey period | 13 | .2% |
| 38-SP learning or mental disability | 1 | .0% |
| 51-Refusal by SP to continue modules 2-7 | 4 | .1% |
| 52-Interview termination-SP unusual circumstance | 1 | .0% |
| 61-SP unable to read language of General Booklet (GB) | 44 | .8% |
| 62-SP refusal to begin GB | 250 | 4.6% |
| 63-SP unable to do GB due to unusual circumstance | 3 | .1% |
| 64-SP blindness or visual impairment-unable to do GB | 24 | .4% |
| 65-SP physical disability | 4 | .1% |
| 71-SP looked through GB-unable to read/write language of GB | 14 | .3% |
| 72-SP refused to continue GB after starting | 39 | .7% |
| 73-GB interruption too long | 3 | .1% |
| 74-GB terminated - unusual circumstance | 3 | .1% |
| 75-Attempted GB, no answers marked | 5 | .1% |
| 76-Attempted part or all of GB; <3 correct Core answers | 55 | 1.0% |
| 81-No attempt at any Exercise Booklet items / refusal to start | 104 | 1.9% |
| 82-Interruption too long during Exercise Booklet administration | 8 | .1% |
| 83-Unusual circumstance, Exercise Booklet administration terminated | 3 | .1% |
| 84-Attempt at some Exercise Booklet items, no answers marked | 28 | .5% |
| 85-Attempt at all Exercise Booklet items, no answers marked | 13 | .2% |
| 86-Marked some Exercise Booklet items, not all items attempted | 655 | 12.1% |
| 87-Marked some Exercise Booklet items, all items attempted | 779 | 14.4% |
| 88-Marked all Exercise Booklet items | 957 | 17.7% |
| 97-Household not interviewed - field procedure error | 271 | 5.0% |
| 98-Incorrect eligible household member selected | 5 | .1% |
| **Total** | 5402 | 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-GEO Final Revised Data File(19-Dec-2013)'.

## 5.1 First-stage Weight (PSU Weight)

The basic first stage weight, i.e., the PSU weight, is based on a sample 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 Georgia 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 Georgia sample of PSUs, the probability of selection of the $j^{th}$ PSU, say $p_j$, may be represented as,

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

where,

$y_g$ - the number of sampled PSUs in stratum g,

$m_{gj}$ - the measure of size for the $j^{th}$ PSU in stratum 'g', in the case of the Georgia STEP Survey the measure of size for the first stage selection of PSUs was the estimated number of persons aged 15 to 64 (i.e., final data file variable '*estimatedsize*') in a PSU,

$M_g$ - the total number of persons aged 15 to 64 in stratum 'g', based on the sample frame information, is represented by

$$2) \qquad M_g = \sum_1^{y_g} m_{gj}$$

Table 10 provides the Georgia stratum counts for $y_g$ and $M_g$.

| Table 10: Georgia Overall PSU Sample Allocation for 'Activated' PSUs by Stratum | | | |
|---|---|---|---|
| Stratum | | Sampled PSUs | Estimated Target Population |
| Area | g | $y_g$ | $M_g$ |
| Capital | 1 | 97 | 792,672 |
| Other Urban NE | 2 | 18 | 149,879 |
| Other Urban NW | 3 | 46 | 373,680 |
| Other Urban SE | 4 | 21 | 171,117 |
| Other Urban NW | 5 | 18 | 150,326 |
| Total | | 200 | 1,637,674 |

The basic first stage weight for the $j^{th}$ PSU, $W1_{gj}$ (i.e., weighted data file variable '**W1_PSU**'), is calculated as the inverse of its probability of selection:

$$3) \qquad W1_{gj} = \frac{1}{p_{gj}} = \frac{M_g}{y_g * m_{gj}}$$

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

14

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

## 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:

4) $\quad W2_{jk} = \dfrac{N_j}{n_j}$

where,

$N_j$ - the number of <u>listed households</u> (i.e., weighted data file variable '**HH_count_inhabited**') in the $j^{th}$ PSU,

$n_j$ - the number of households sampled in the $j^{th}$ PSU.

---

**Example 2**: Second Stage Basic Weight

The basic second stage weight for each of the 34 sampled households in Georgia PSU # 101 which contains 458 listed households is calculated as follows:

$$W2_{jk} = \frac{m'_j}{n_j} = \frac{458}{34} \approx 13.4706$$

---

The basic second stage weight is similarly calculated for each of the 6864 sampled households.

### 5.2.2　Second-stage Response-adjusted Household Weight

#### 5.2.2.1　Second-stage Household Response Categories

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 <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, 51-88 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 |

| I. Household-level Codes – Participant, Eligible Household | |
|---|---|
| **Code** | **Code Description** |
| 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 |
| 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 |
| 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.

$$5) \quad n_j = n_{jk} + n_{jx} + n_{jz} + n_{ju}$$

where,

$n_j$ - the number of households sampled in the jth PSU,

$n_{jk}$ - the number of households 'k' that participate in the jth PSU,

$n_{jx}$ - the number of ineligible households 'x' in the jth PSU,

$n_{jz}$ - the number of non-participating households 'z' with known eligibility in the jth PSU,

$n_{ju}$ - the number of non-participating households 'u' with unknown eligibility in the jth PSU.

Table 11 provides the STEP counts of <u>activated</u> households from the Georgia data file by the outcome categories described above.

| Table 11: Sample Count of 'Activated' Households by Final Response Category | | | | | | |
|---|---|---|---|---|---|---|
| | Final Response Category-Households | | | | # of HHs | % of Total |
| Case Final Status Code | Participant | Ineligible | Non-participant: Eligible | Non-participant: Unknown Eligibility | | |
| 01-Refusal-Household | | | | 1103 | 1103 | 20.4% |
| 02-Unusual circumstance-HH | | | | 17 | 17 | .3% |
| 03-No knowledgeable HH member found | | | | 73 | 73 | 1.4% |
| 04-Temporarily absent | | | | 242 | 242 | 4.5% |
| 05-No competent HH member | | | | 14 | 14 | .3% |
| 06-Language problem | | | | 28 | 28 | .5% |
| 07-Dwelling not found - no address info | | 93 | | | 93 | 1.7% |
| 08-Vacant dwelling | | 255 | | | 255 | 4.7% |
| 09-Dwelling not habitable | | 50 | | | 50 | .9% |
| 21-HHRefusal to continue HH Module | | | 5 | | 5 | .1% |
| 22-HH module terminated - unusual circumstance | | | 1 | | 1 | .0% |
| 31-No eligible household member aged 15-64 | | 218 | | | 218 | 4.0% |
| 32-Refusal by SP to begin modules 2-7 | 12 | | | | 12 | .2% |
| 33-Interview termination-SP unusual circumstance | 1 | | | | 1 | .0% |
| 35-Unable to contact SP after 3 attempts | 4 | | | | 4 | .1% |
| 36-SP absent for entire survey period | 13 | | | | 13 | .2% |
| 38-SP learning or mental disability | 1 | | | | 1 | .0% |
| 51-Refusal by SP to continue modules 2-7 | 4 | | | | 4 | .1% |
| 52-Interview termination-SP unusual circumstance | 1 | | | | 1 | .0% |
| 61-SP unable to read language of GB | 44 | | | | 44 | .8% |
| 62-SP refusal to begin GB | 250 | | | | 250 | 4.6% |
| 63-SP unable to do GB - unusual circumstance | 3 | | | | 3 | .1% |
| 64-SP blindness or visual impairment | 24 | | | | 24 | .4% |
| 65-SP physical disability | 4 | | | | 4 | .1% |
| 71-SP unable to read/write language of GB | 14 | | | | 14 | .3% |
| 72-SP refused to continue GB after starting | 39 | | | | 39 | .7% |
| 73-GB interruption too long | 3 | | | | 3 | .1% |
| 74-GB terminated - unusual circumstance | 3 | | | | 3 | .1% |
| 75-Attempted GB, no answers marked | 5 | | | | 5 | .1% |
| 76- <3 correct Core answers | 55 | | | | 55 | 1.0% |
| 81-No attempt at any EB items / refusal to start | 104 | | | | 104 | 1.9% |
| 82-Interruption too long during EB administration | 8 | | | | 8 | .1% |
| 83-Unusual circumstance, EB terminated | 3 | | | | 3 | .1% |
| 84-Attempt at some EB items, no answers marked | 28 | | | | 28 | .5% |
| 85-Attempt at all EB items, no answers marked | 13 | | | | 13 | .2% |
| 86-Marked some EB items, not all items attempted | 655 | | | | 655 | 12.1% |
| 87-Marked some EB items, all items attempted | 779 | | | | 779 | 14.4% |
| 88-Marked all EB items | 957 | | | | 957 | 17.7% |
| 97-HH not interviewed - field procedure error | | | 271 | | 271 | 5.0% |
| 98-Incorrect eligible HH member selected | 5 | | | | 5 | .1% |
| Total | 3032 | 616 | 277 | 1477 | 5402 | 100.0% |

## 5.2.2.2 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 $\varepsilon_j$, 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

$$6) \quad \varepsilon_j = \frac{\textit{Total \# of known eligible cases}}{\textit{Total \# of known eligible \& known ineligible cases}} = \frac{n_{jk} + n_{jz}}{n_{jk} + n_{jx} + n_{jz}}$$

➢ **Note that the eligible proportion is the same for all eligible participating households.**

---

**Example 3**: Eligible Proportion of Households

For PSU #101, the following household results were obtained:

| $n_j$ | $n_{jk}$ | $n_{jx}$ | $n_{jz}$ | $n_{ju}$ |
|-------|----------|----------|----------|----------|
| 34    | 15       | 6        | 0        | 13       |

For PSU # 101, the proportion of households with known eligibility is calculated as,

$$\varepsilon_{101} = \frac{\textit{Total \# of known eligible cases}}{\textit{Total \# of known eligible \& known ineligible cases}}$$

$$= \frac{n_{jk} + n_{jz}}{n_{jk} + n_{jx} + n_{jz}}$$

$$= \frac{15 + 0}{15 + 6 + 0} = \frac{15}{21} \approx 0.7142857$$

### 5.2.2.3 Response Adjustment Factor for Household 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:

$$7) \quad R_{HH_j} = \frac{\textit{Total eligible households in PSU 'j'}}{\textit{Total particpating households in PSU 'j'}} = \frac{n_{jk} + n_{jz} + (\varepsilon_j * n_{ju})}{n_{jk}}$$

➤ The response adjustment ratio, $R_{HHj}$, incorporates the adjustment for unknown eligibility, $\varepsilon_j$.
➤ The households that are categorized as ineligible, $n_{jx}$, are excluded from the calculation of the household-level response adjustment.

---

**Example 4**: Second Stage Response Adjustment Factor for sampled Households

**For example**, the following results were obtained for PSU #101

| $n_j$ | $n_{jk}$ | $n_{jx}$ | $n_{jz}$ | $n_{ju}$ |
|-------|----------|----------|----------|----------|
| 34 | 15 | 6 | 0 | 13 |

Therefore,

$$R_{102} = \frac{\textit{Total eligible households in PSU  \#101}}{\textit{Total particpating households in PSU  \#101}}$$

$$= \frac{n_{jk} + n_{jz} + (\varepsilon_{102} * n_{ju})}{n_{jk}}$$

$$= \frac{15 + 0 + (.7142857 * 13)}{15} \approx 1.61905$$

---

The response adjustment ratio, $R_{HHj}$, is similarly calculated for all 200 sampled PSUs.

### 5.2.2.4   *Response-adjusted Household Weight*

The second stage response-adjusted household weight is calculated as the product of 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, $W2_{Radj}$, is calculated as follows:

$$8) \quad W2_{Radj} = W2_{jk} * R_{HH_j}$$

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

---

**Example 5**: Second Stage Response-adjusted Household Weight

For PSU #101,

$$W2_{Radj} = W2_{jk} * R_{HH_j}$$
$$= 13.4706 \times 1.61905$$
$$\approx 21.8095$$

---

The response-adjusted household weight is similarly calculated for all sampled households.

## 5.2.3   Second-stage 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 second stage household weight is calculated as product of the response-adjusted household weight and the final first-stage PSU weight. Therefore, for household 'k' in PSU 'j', the final second stage household weight, $W2_{HHk}$, is calculated as follows:

$$9) \quad W2_{HH_k} = W2_{Radj} * W1_{gj}$$

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

---

**Example 6**: Final Second Stage Household Weight

For PSU #101,

$$W2_{HH_k} = W2_{Radj} * W1_{gj}$$
$$= 21.8095 \times 5.6749$$
$$\approx 123.7672$$

---

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

## 5.2.4 Weighted Household Counts

The STEP data collection found that some of the sampled households are ineligible for reasons such as vacant, not habitable, no eligible household member, etc.

Table 12 provides,
 i.  the weighted count of households based on the listed households in the 200 sampled PSUs;
 ii. the weighted count of households based on the final status of the cases in the Georgia data file.

| Table 12:  Georgia – Weighted Number of Households – Sample Frame vs. STEP Data File | | |
|---|---|---|
| Source of PSU Household Count | # of PSUs | Weighted Sum of Households |
| 1.  Based on # of listed PSU Hhlds in STEP Data File 'STEP Georgia - Data for Weighting_S7' [**all participants**, i.e., participants, non-participants, ineligibles] | 200 | 591,961 |
| 2.  Based on # of PSU Hhlds in STEP Data File 'STEP Georgia - Data for Weighting_S7' [**eligible households only**] | 200 | 503,417 |

As shown in the second row of Table 12, when ineligible households are excluded, the weighted number of eligible households is **503,417**.

For participating selected households, Table 13 provides a summary of the central tendency for the participating households.

| Table 13:  Summary - Household Weights | |
|---|---|
| Statistic | Final Household Weight |
| No. of Cases | 3032 |
| Mean | 166.03 |
| Std. Error of Mean | 1.01 |
| Median | 171.32 |
| Minimum | 26.61 |
| Maximum | 296.85 |
| Sum of Weights | 503,417 |

## 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, $W3_{jki}$ , is calculated as follows:

$$10) \quad W3_{jki} = \frac{a_{jk}}{b_{jk}} = \frac{a_{jk}}{1} = a_{jk}$$

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 Georgia sampled household).

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

where,

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

---

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

For PSU #102, selected household # 04 has 3 eligible persons. Therefore,

$$W3_{jki} = \frac{a_{jk}}{b_{jk}} = \frac{3}{1} = 3$$

---

## 5.3.2 Third-stage Response-adjusted Person Weight

### 5.3.2.1 Third-stage Person-level Response Categories

**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 <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 |
| 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 |

Note: there were no ineligible selected persons in the Georgia 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 14 provides the pertinent frequency counts of the Final Status Codes for the participant households.

| Table 14: Final Status Code for 'Participating' Households (i.e., Households with RespCat_HH=1) | | |
|---|---|---|
| **Case Final Status Code** | **Frequency** | **Percent** |
| 32-Refusal by SP to begin modules 2-7 | 12 | .4% |
| 33-Interview termination-SP unusual circumstance | 1 | .0% |
| 35-Unable to contact SP after 3 attempts | 4 | .1% |
| 36-SP absent for entire survey period | 13 | .4% |
| 38-SP learning or mental disability | 1 | .0% |
| 51-Refusal by SP to continue modules 2-7 | 4 | .1% |
| 52-Interview termination-SP unusual circumstance | 1 | .0% |
| 61-SP unable to read language of General Booklet (GB) | 44 | 1.5% |
| 62-SP refusal to begin GB | 250 | 8.2% |
| 63-SP unable to do GB due to unusual circumstance | 3 | .1% |
| 64-SP blindness or visual impairment-unable to do GB | 24 | .8% |
| 65-SP physical disability | 4 | .1% |
| 71-SP looked through GB-unable to read/write language of GB | 14 | .5% |
| 72-SP refused to continue GB after starting | 39 | 1.3% |
| 73-GB interruption too long | 3 | .1% |
| 74-GB terminated - unusual circumstance | 3 | .1% |
| 75-Attempted GB, no answers marked | 5 | .2% |
| 76-Attempted part or all of GB; <3 correct Core answers | 55 | 1.8% |
| 81-No attempt at any Exercise Booklet items / refusal to start | 104 | 3.4% |
| 82-Interruption too long during Exercise Booklet administration | 8 | .3% |
| 83-Unusual circumstance, Exercise Booklet administration terminated | 3 | .1% |
| 84-Attempt at some Exercise Booklet items, no answers marked | 28 | .9% |
| 85-Attempt at all Exercise Booklet items, no answers marked | 13 | .4% |
| 86-Marked some Exercise Booklet items, not all items attempted | 655 | 21.6% |
| 87-Marked some Exercise Booklet items, all items attempted | 779 | 25.7% |
| 88-Marked all Exercise Booklet items | 957 | 31.6% |
| 98-Incorrect eligible household member selected | 5 | .2% |
| **Total** | **3032** | **100.0%** |

## 5.3.2.2 Response Adjustment Factor for Selected Person Non-participation

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, $W3_{jki}$, would also be the final person weight within a sampled PSU if all persons in the PSU participated. 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 PSU may be represented by the following equation:

$$11) \quad A_{jk} = A_{jki} + A_{jkq}$$

where,

$A_{jk} = \sum_{k=1}^{n_{jk}} a_{jk}$ - the total number of eligible persons for all $n_{jk}$ participating households in the $j^{th}$ PSU,

$A_{jki} = \sum_{k=1}^{n_{jk}} a_{jki}$ - the total number of eligible persons 'i' that participate for all $n_{jk}$ participating households in the $j^{th}$ PSU,

$A_{jkq} = \sum_{k=1}^{n_{jk}} a_{jkq}$ - the total number of non-participating eligible persons 'q' for all $n_{jk}$ participating households in the $j^{th}$ PSU,

Therefore, the person-level response adjustment for the $j^{th}$ PSU, $R_{jk}$, is calculated as follows:

$$12) \quad R_{SP_j} = \frac{\textit{Total eligible persons in all participating sampled households in PSU 'j'}}{\textit{Total particpating persons in all participating sampled households in PSU 'j'}} = \frac{A_{jk}}{A_{jki}}$$

---

**Example 8**: Third Stage – Basic Person Weight Response Adjustment

i) For PSU #101, there are 39 eligible persons across all participating households; also, all selected persons participated. Therefore,

$$A_{jk} = \sum_{k=1}^{n_{jk}} a_{jk} = 39 \; ; \qquad A_{jki} = \sum_{k=1}^{n_{jk}} a_{jki} = 39$$

$$R_{SP_j} = \frac{A_{jk}}{A_{jki}} = \frac{39}{39} = 1.00$$

ii) For PSU # 102, the total number of eligible persons across all participating households is 41. For PSU # 102, there are six non-participating selected persons. Therefore,

$$A_{jk} = \sum_{k=1}^{n_{jk}} a_{jk} = 41 \; ; \qquad A_{jki} = \sum_{k=1}^{n_{jk}} a_{jki} = 35$$

$$R_{SP_j} = \frac{A_{jk}}{A_{jki}} = \frac{41}{35} \approx 1.17143$$

---

Table 14 provides the sample counts of the Final Response Category for selected persons.

| Table 15: Sample Counts by Selected Person Final Response Category for Participating HHs | | |
|---|---|---|
| **For Participating Households** | | |
| **Final Response Category-SP** | **Frequency** | **Percent** |
| 1-Participant SP | 2996 | 98.8 |
| 2-Non-participant, Eligible | 36 | 1.2 |
| **Total** | **3032** | **100.0** |

The above table shoes that there are 2996 participating selected persons for which a final non-zero Person-weight was calculated. The sampled persons in category 2 have a zero-value final person weight.

### 5.3.2.3   *Response-adjusted Person Weight*

The response-adjusted person weight is calculated as the product of 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 response-adjusted person weight, $W_{Radj}$, is calculated as follows:

$$13) \quad W3_{Radj} = W3_{jki} * R_{SP_j}$$

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

---

**Example 9**: Third Stage – Response-adjusted Person Weight

i)   For PSU #101, selected household # 01 with 3 eligible persons, the selected person was a participant. All selected persons participated in PSU #101, therefore, $R_{SP_j} = 1$

$$W3_{jki} = 3 * 1 = 3$$

$$W3_{Radj} = W3_{jki} * R_{SP_j}$$
$$= 3 * 1 = 3$$

ii)  For PSU #102, in participant household # 14 with 4 eligible persons, the selected person was a participant. For this household,

$$W_{3jki} = 4 ; \qquad R_{SP_j} \approx 1.17143$$

$$W3_{Radj} = W3_{jki} * R_{SP_j}$$
$$= 4 * 1.17143 \approx 4.685714$$

---

The response-adjusted person weight is similarly calculated for all sampled persons.

### 5.3.3  Third-stage Final 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) and the response adjusted basic person weight. Therefore, for the selected person 'i' in household 'k' in PSU 'j', the final person weight, $W3_{SP}$ (i.e., weighted data file variable '**W_FinSPwt**'), is calculated as follows:

$$14)\quad W3_{SP} = W2_{HH_k} * W3_{Radj}$$

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

---

**Example 8**: Third Stage – Final Person Weight

iii)    For PSU #101, selected household # 03 with 3 eligible persons, the selected person was a participant.

$$W2_{HH_k} = 123.7672 \,; W3_{Radj} = 3$$

$$W3_{SP} = W2_{HH_k} * W3_{Radj}$$

$$W3_{SP} = 123.7672 * 3 \approx 371.3015$$

iv)    For PSU #102, in participant household # 14 with 4 eligible persons, the selected person was a participant. For this household,

$$W2_{HH_k} \approx 130.4654 \,; \ W3_{Radj} = 4.6857$$

$$W3_{SP} = W2_{HH_k} * W3_{Radj}$$

$$W3_{SP} = 130.4654 * 4.6857 \approx 611.3244$$

---

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.

### 5.3.4 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 **Georgia** at the time of data collection is provided in Table 16. Based on the 2996 participating selected persons, the sum of the final person weights from the STEP survey is 1,346,787.

| **Table 16:** | Sampled Persons & Weighted # of Persons by Response Category | |
|---|---|---|
| **Response Category (SP)** | **Sampled Cases** | **Total # of Eligible Persons in Target Population** |
| | | **(based on Final Person weight)** |
| 1-Participant SP | 2,996 | 1,346,787 |
| 2-Non-participant, Eligible | 36 | 0 |
| **Total** | **3,032** | **1,346,787** |

For participating selected persons, Table 17 provides a summary of the central tendency of the number of eligible persons in the participating households.

| **Table 17: Summary - Eligible Persons, Person Weights** | | |
|---|---|---|
| **Statistic** | **Eligible Persons** | **Final Person Weight** |
| **No. of Cases** | 3032 | 3032 |
| **Mean** | 2.67 | 444.19 |
| **Std. Error of Mean** | 0.022 | 4.85 |
| **Median** | 2.00 | 390.16 |
| **Minimum** | 1 | 0.00 |
| **Maximum** | 10 | 1854.36 |
| **Sum** | **8109** | **1,346,787** |

# 6    Response Rate

An overall response rate of 62.6% was achieved in the Georgia STEP Survey. Table 18 provides the detailed percentage distribution by final status code.

**Response Rate: Georgia, Activated HHs by Final Status Code**

| Case Final Status Code | Frequency | Percent |
|---|---|---|
| 07-Dwelling could not be found / selected address has no household | 93 | |
| 08-Vacant dwelling / selected address has no household | 255 | 9.4% |
| 09-Dwelling not habitable/destroyed/converted to commercial | 50 | 0.9% |
| 31-No eligible household member aged 15-64 | 218 | 2.4% |
| **Ineligible HHs** *(N.B. % is based on total # of activated HHs, i.e., 5402 HHs)* | **616** | **11.4%** |
| 01-Refusal-Household | 1103 | 20.4% |
| 02-Unusual circumstance-Household | 17 | .3% |
| 03-No knowledgeable household member found after 3 visits | 73 | 1.4% |
| 04-Temporarily absent/unavailable for field period | 242 | 4.5% |
| 05-No competent household member (e.g., severe illness, mental disability, etc.) | 14 | .3% |
| 06-Language problem | 28 | .5% |
| 21-Household Refusal to continue Household Module (begun, not complete) | 5 | .1% |
| 22-Household module terminated - unusual circumstance | 1 | .0% |
| 32-Refusal by SP to begin modules 2-7 | 12 | .2% |
| 33-Interview termination-SP unusual circumstance | 1 | .0% |
| 35-Unable to contact SP after 3 attempts | 4 | .1% |
| 36-SP absent for entire survey period | 13 | .2% |
| 38-SP learning or mental disability | 1 | .0% |
| 97-Household not interviewed - field procedure error | 271 | 5.0% |
| 98-Incorrect eligible household member selected | 5 | .1% |
| **Non-respondents (Eligible)** *(N.B. % is based on total # of Eligible HHs, i.e., 4786 HHs)* | 1790 | 37.4% |
| 51-Refusal by SP to continue modules 2-7 | 4 | .1% |
| 52-Interview termination-SP unusual circumstance | 1 | .0% |
| 61-SP unable to read language of General Booklet (GB) | 44 | .8% |
| 62-SP refusal to begin GB | 250 | 4.6% |
| 63-SP unable to do GB due to unusual circumstance | 3 | .1% |
| 64-SP blindness or visual impairment-unable to do GB | 24 | .4% |
| 65-SP physical disability | 4 | .1% |
| 71-SP looked through GB-unable to read/write language of GB | 14 | .3% |
| 72-SP refused to continue GB after starting | 39 | .7% |
| 73-GB interruption too long | 3 | .1% |
| 74-GB terminated - unusual circumstance | 3 | .1% |
| 75-Attempted GB, no answers marked | 5 | .1% |
| 76-Attempted part or all of GB; <3 correct Core answers | 55 | 1.0% |
| 81-No attempt at any Exercise Booklet items / refusal to start | 104 | 1.9% |
| 82-Interruption too long during Exercise Booklet administration | 8 | .1% |
| 83-Unusual circumstance, Exercise Booklet administration terminated | 3 | .1% |
| 84-Attempt at some Exercise Booklet items, no answers marked | 28 | .5% |
| 85-Attempt at all Exercise Booklet items, no answers marked | 13 | .2% |
| 86-Marked some Exercise Booklet items, not all items attempted | 655 | 12.1% |
| 87-Marked some Exercise Booklet items, all items attempted | 779 | 14.4% |
| 88-Marked all Exercise Booklet items | 957 | 17.7% |
| **Respondents** *(N.B. % is based on total # of Eligible HHs, i.e., 4786 HHs)* | **2996** | **62.6%** |
| **Total _Eligible_ Households** | **4786** | **100.0%** |
| **Total <u>Activated</u> Households** | **5402** | **100.0%** |