Canadian Community Health Survey (CCHS)

Household weights documentation

TABLE OF CONTENTS

1.   Introduction

This document describes the weighting process used in the creation of the household weight for the Canadian Community Health Survey (CCHS). In using this weight, users should note that the survey is designed to represent individuals and not designed to represent households. Certain steps are taken in the design to ensure that the sample is representative of different demographic groups and this may affect how well the sample represents households of different compositions. Also, since the calibration for the household weight is done at the provincial level, it is possible to yield reliable estimates at the national and provincial level only. It is felt that individual responses to certain questions can be used to represent the household. This weight should only be used for variables where it can be assumed that the responses from the individual clearly represent the household and that the response would not be affected by who responded within the household. This is highlighted by the fact that throughout the document, we refer to responding persons. Users must remember that the responses from this selected individual are assumed to represent the household when using the household weight.

Those familiar with the CCHS will notice that these weight adjustments are very similar to those used for other weights that have been produced in the past. A CCHS interview can be seen as a two-part process. First, the interviewer gets the complete roster of the people living within the household (Household Response). Second, (s)he interviews the selected person within the household (Person Response). In the calculation of the household weight, the individual responses are used to represent the household. Nonresponse adjustments for both stages are still included since nonresponse can occur at either part of the interview process. Note that nonresponse adjustments for the household can be based on characteristics of the individual respondent. Since the survey is designed to collect information from the individual, the characteristics of the individual can have an effect on nonresponse of the household.

2.   Weighting Overview

In order for estimates produced from survey data to be representative of the covered population and not just the sample itself, users must incorporate the survey weights in their calculations. A survey weight is given to each person included in the final sample, that is, the sample of persons having answered the survey. This weight corresponds to the number of households in the entire population that are represented by the respondent.

The CCHS has recourse to three sampling frames for its sample selection: an area frame acting as the primary frame and two frames formed of telephone numbers complementing the area frame. Since only minor differences differentiate the two frames formed of telephone numbers in terms of weighting, they are treated together. They are referred to as being part of the telephone frame.

The weighting strategy treats both the area and telephone frames independently. Weights resulting from these two frames are afterwards combined into a single set of weights through a step called "*integration*". After some adjustments, this integrated weight becomes the final weight. Note that

depending on the need, one or two frames were used for the selection of the sample within a given health region (HR). The weighting strategy deals with this aspect at the integration step.
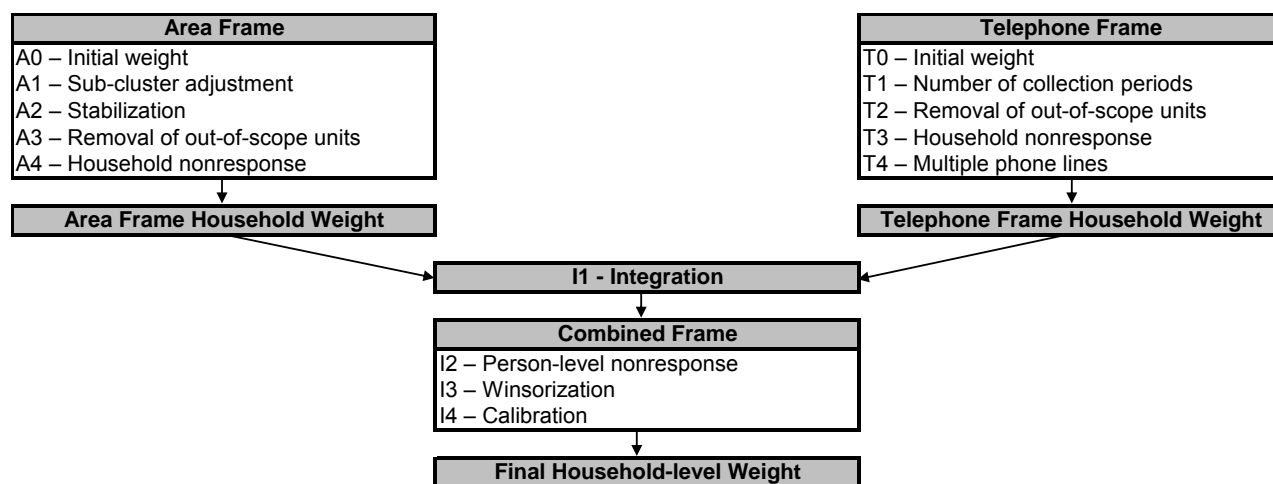
3. Sample weighting

As mentioned previously, units from both the area and telephone frames are treated separately up to the integration step. These weighting steps for the household weight, up to and including the integration of the frames, are the same as the steps from the main weight. Please refer to the CCHS User Guide for more information about these steps. The final three weighting steps, person nonresponse, winsorization and calibration, are explained in sub-sections 3.1-3.3.

Although these two frames were used to cover the three territories, some modifications had to be done relative to their use. These modifications substantially affected the weighting of these three regions and they are reported in sub-section 3.4.

Diagram A presents an overview of the different adjustments that are part of the weighting strategy, in the order in which they are applied. A numbering system is used to identify each adjustment applied to the weight and will be used throughout the section. Letters A and T are used as prefixes to refer to adjustments applied to the units on the Area and Telephone frames respectively, while prefix I identifies adjustments applied from the Integration step onwards.

Diagram A: Weighting strategy overview (Household weight)

Area Frame	**Telephone Frame**
A0 – Initial weight A1 – Sub-cluster adjustment A2 – Stabilization A3 – Removal of out-of-scope units A4 – Household nonresponse	T0 – Initial weight T1 – Number of collection periods T2 – Removal of out-of-scope units T3 – Household nonresponse T4 – Multiple phone lines
Area Frame Household Weight	**Telephone Frame Household Weight**

I1 - Integration

Combined Frame
I2 – Person-level nonresponse I3 – Winsorization I4 – Calibration

Final Household-level Weight

3.1 Person-level nonresponse (I2)

A CCHS interview can be seen as a two-part process. First, the interviewer gets the complete roster of the people within the household. Second, (s)he interviews the selected person. In some cases, interviewers can only get through the first part, either because they cannot get in touch with the selected person or because that selected person refuses to be interviewed. Such individuals are

defined as person nonrespondents and an adjustment factor must be applied to the weights of the respondents to account for this nonresponse. Using the same methodology that was used in the treatment of household nonresponse (see User Guide, Section 8.2 - A4), the adjustment is applied within response homogeneity groups based on characteristics available for both respondents and non-respondents. All characteristics collected when creating the roster of household members are available for the creation of the groups as well as geographic information and some paradata. The scoring method is used to define the classes. In the end, the following adjustment factor is calculated within each group:

$$\frac{\text{Sum of weight I1 for all selected persons}}{\text{Sum of weight I1 for all selected persons who respond}}$$

Weight I1 is multiplied by the above adjustment factor to produce weight I2. Nonresponding persons are dropped from the weighting process from this point onward.

3.2   Winsorization (I3)

Following the series of adjustments applied to the respondents, some units may come out with extreme weights compared with other units in the same domain of interest. For the household weights, the domains include province by household size, where household size is defined by: 1-person household, 2-person household and at least 3-person household. Some responding households could represent a large proportion of their province by household size domain or have a large impact on the variance. In order to prevent this, the weights of the outlier units that represent a large proportion of their domain are adjusted downward using a "Winsorization" trimming approach.

3.3   Calibration (I4)

The last step in obtaining the final CCHS household weight is calibration. Calibration is done using the program CALMAR to ensure that the sum of the final weights corresponds to the household estimates defined at the province by household size level. These groups of interest are defined by the sizes: 1-person household, 2-person household and at least 3-person household. At the same time, the weights are seasonally adjusted to ensure that the each two month collection period is equally represented within the sample. In terms of geography, all calibration is at the provincial level only.

The household count estimates are based on the most recent census. The average of these monthly estimates is used to calibrate for each of the province by household size post-strata within a collection period. The weight I3 is therefore adjusted to obtain the final weight I4 with the help of the adjustment factor I4 defined as follows:

$$\frac{\text{Population estimate for the Province by Household size of the respondents}}{\text{Sum of weights I3 for the Province by Household size of the respondents}}$$

Consequently, the weight I4 corresponds to the ***final CCHS household weight*** that can be found on the household weight file with the variable name WTS_MHH for the master weight and WTS_SHH for the share weight.

As of 2010, differences in the counts of households of size 3+ may be noted in some areas. This is due to a change in the methods used to estimate the current number of households and an update to use information from the 2006 census. The counts starting in 2010 represent an improved estimate of the number of households.

3.4 Particular aspects of the weighting in the three territories

The sampling frame used in the three territories is somewhat different from the one used in the provinces. Therefore the weighting strategy is adapted to comply with these differences. This section summarises the changes applied to the weighting steps in the territories.

For the area frame, an additional stage of selection is added in the territories where each territory is initially stratified into groups of communities and one community is selected within each group. Note that the capital of each territory forms a stratum on its own and is consequently selected automatically at this first stage. This has an effect in the computation of the probability of selection and therefore in the value of the initial weight (A0). Once the initial weight is calculated, the same series of adjustments (A1 to A4) is applied to the area frame units. The household-level adjustment classes are built in the same way as for the provinces, using the same set of variables available.

For the weighting of the telephone frame units, it should first be noted that only the RDD frame is used for the territories, and exclusively in the Yukon and Northwest Territories capitals. All of the standard telephone adjustments are applied.

The two sets of weights (area and telephone) are subsequently integrated and then adjusted for person level nonresponse, winsorization and finally calibrated in a similar way to what is done for the provinces, with the exception of four details. First, the integration is applied only to units located in the Yukon and Northwest Territories capitals since the other communities are covered only by the area frame. Second, for Nunavut, the household counts used for calibration only represent the 10 largest communities (73% of the households) because of the under-coverage of the area frame (for more details, see User Guide, section 5.4.1). Third, in the Yukon and Northwest Territories, starting with the 2008 and 2007/2008 reference periods, calibration is used to control for the proportion of households located inside the capital cities versus the proportion of households located outside of the capital cities. The same approach has been adapted for Nunavut starting in 2009. Finally, due to the differences in collection strategies, the number of collection periods used in calibration for the seasonal effect in the territories is different from the provinces. Since 2009, two 6-month periods are used in the three Territories.

3.5   Creation of the Share weight (WTS_SHH)

Along with the master file and PUMF, which contain all CCHS responding persons, a share file is created which contains only a portion (>90%) of the original CCHS responding persons. The individuals on this share file have agreed to share their data with certain partners. To compensate for the loss of some respondents from the file, the weights of these "sharers" must be adjusted by the factor:

$$\frac{\text{Sum of weight I4 for all respondents}}{\text{Sum of weight I4 for all respondents agreeing to share their data}}$$

Similar to the nonresponse adjustments, this factor is calculated within homogeneity groups, where in this case, individuals with similar estimated propensity to share will be grouped together. The final weight after this adjustment is called WTS_SHH.

3.6   The Food Security Module in 2010

The Food Security Module (FSC) is one of the few modules that are part of the CCHS where it can be appropriate to use the household weights. For 2007-2008, the FSC module was part of the common content, and therefore was asked to respondents in all provinces and territories. For the 2009-2010 collection period, the FSC module is part of optional content and was not selected for some of the provinces. Given this change, it is no longer appropriate to calculate national estimates with variables in this module since the results would only represent those provinces and territories where the module was asked. For more information on optional content selection, please refer to the appendices of the User Guide.