
C. Computing the SIPP Sampling Weights

This appendix supplements the discussion in Chapter 8 (Using Sampling Weights on SIPP Files) with more detailed information about how the core wave file person-level weight FNLWGT and the full panel file person-level weights FNLWGT_x and PNLWGT are computed;¹ it is intended as a reference for users who require a comprehensive description of how the sampling weights are computed.

Sections 1 and 2 of this appendix discuss the algorithms that are used to compute the final core wave file person-level weights FNLWGT, with the first section discussing the Wave 1 weights and the second section discussing the Wave 2+ weights. The third section discusses the algorithm that computes the final full panel weights FNLWGT_x (the calendar year weight for year *x*) and PNLWGT (the panel weight).

Wave 1 Weights

For the 1996 Panel, the final weights used in deriving estimates consist of the product of four factors: the base weight, the duplication control factor, the household noninterview adjustment factor, and the second-stage adjustment factor. For panels prior to 1996, these four factors may have been multiplied by two other factors—the first-stage ratio estimate factor and the new construction noninterview adjustment factor—which are discussed later in this chapter.

Base Weight (BW)

The primary component of the sampling weight is the base weight. The base weight for any sampled person or sampled household is the reciprocal of the probability under the sample design of that person or household being selected. If there was full response and if there were no calibration adjustments, then the summation of base weights for a particular subgroup (e.g., Hispanics in the Southwest) is an unbiased estimator of the total U.S. population within that subgroup. In simplified terms, a base weight of 1,000 assigned to a sampled person means that the sampled person “represents” 1,000 people in the U.S. population. The base weight for a

¹ The remaining weights given in Table 12-2 (HWGT, FWGT, SWGT, P5WGT, H5WGT, and FINALWGT) are derived directly from the basic person-level weight FNLWGT. This derivation is discussed in the “How Weights Are Constructed” subsection of Chapter 8.

household and the base weight for a person within a household are the same, since every person within a sampled household is automatically selected (i.e., selected with a conditional probability of 1, given household selection).

Duplication Control Factor (DCF)

The duplication control factor, an integer value between 1 and 4 inclusive, is applied to the base weights of specified households to account for subsampling done in clusters of housing units selected at the last stage of sample selection. These clusters typically contain an unmanageable number of housing units. When this occurs, a sampling fraction, $1/N$, is determined by selecting a value of N such that the number of sample households in the cluster is reduced to a manageable size. After this is done, a duplication control factor of N or 4, whichever is smaller, is included as a weighting factor for sampled housing units in the cluster.

Household Noninterview Adjustment Factor (NAF)

The noninterview adjustment factor is intended to adjust for the presence of Type A noninterview households (households that are not interviewed because the occupants were temporarily absent, no one was home, the occupants refused participation, or the occupants could not be located). Noninterview adjustment factors are computed for each of a set of noninterview cells. These cells are based on 512 cells generated from all possible cross-classifications of the following household characteristics (256 cells for panels prior to 1996):

- Within-PSU oversampling strata: poverty stratum and nonpoverty stratum (only for 1996 and later panels);
- Census region;
- Race of reference person: black or nonblack;
- Tenure: owner or renter;
- Residence status: MSA urban, MSA nonurban, NonMSA Census place, or NonMSA not Census place; and
- Household size: one, two, three, or four or more persons.

Any cells with fewer than 30 interviewed households or with noninterview adjustment factors exceeding 2.0 are collapsed with a neighboring cell. To define cells as neighboring, the Census Bureau uses a sort order and scale values based on estimates of the 1979 poverty rate within the cell. The total number of noninterview cells is less than or equal to 512 for the 1996 Panel (256 or fewer for the earlier panels). In pre-1996 Panels, no cells were collapsed across the four cells defined by the cross-classification of race of reference person and tenure. For the 1996 Panel, no

cells are collapsed over the cross-cells defined by race of reference person, tenure, within-PSU oversampling strata, and Census region.

Within each final noninterview cell c , the formula for the noninterview adjustment factor (NAF $_c$) is

$$\text{NAF}_c = \frac{\text{sum of BW * DCF over all sampled households in cell } c}{\text{sum of BW * DCF over all interviewed households in cell } c} . \quad (\text{C-1})$$

This factor is applied to the weight of each interviewed household in the cell; with these noninterview-adjusted weights, the interviewed households in each cell can be seen to “represent” themselves and also the Type A noninterviewed households in the cell.²

Wave 1 Second-Stage Calibration Adjustment (SSCA)

For the second-stage calibration adjustments, the Census Bureau uses tallies of Current Population Survey (CPS) weights for independent population controls. The CPS weights are calibrated to match population controls provided by the population division of the Census Bureau and then a “March type” adjustment is done to equalize the weights of husbands and wives. Because the population division does not produce family-type controls, SIPP family-type controls are in fact CPS sample estimates. SIPP controls for age, sex, and race, on the other hand, should not differ appreciably from the original population division controls.

The primary steps in the calibration (or ratio estimation) process are the attaching of second-stage calibration adjustment factors to the pre-second-stage weights (BW*DCF*NAF) within particular cells (e.g., male Hispanic 14-year-olds) so that the resulting adjusted weights (BW*DCF*NAF*SSCA) aggregate to independent CPS-derived population estimates within the cell. The summation of the pre-second-stage weights within any cell are unbiased estimates (assuming the nonresponse adjustment successfully adjusts for all effects of nonresponse) of the population totals (e.g., the summation of BW*DCF*NAF over all male Hispanic 14-year-olds in the panel is an unbiased estimate of the total number of male Hispanic 14-years-olds in the U.S. population).

For SIPP, the monthly CPS estimates of the population totals in these cells are generally superior to the aggregations of nonresponse-adjusted SIPP weights (superior in the sense of having lower sampling and/or nonsampling error). The adjusted weights (BW*DCF*NAF*SSCA) give estimates then for these cells that are equal to the independent estimates. This adjustment generally improves the overall precision of all estimates of these cells or any other related survey characteristics that are prevalent in these cells.

² In pre-1996 Panels, group quarters housing units were not included in the nonresponse computations, and received nonresponse adjustments equal to 1. Group quarters housing units are treated as other households in the 1996 Panel.

The population cells for which adjustments are made to independent estimates are given in Figures C-1, C-2, and C-3 (see pages C-6–C-11). The cells include (as can be seen in the figures) age, race, sex, Spanish origin, family relationship, and household type. As noted earlier, the independently derived estimates for these cells are based on CPS March supplement-type estimates, except the estimates for family type. (The CPS estimates are not the usual CPS monthly estimates. [See U.S. Census Bureau (1998) for more details.] The estimates are specially computed for this purpose by summing the CPS weights within a given cell for all sample units in the relevant CPS sample [there are some extra steps also, such as the equalization of husbands' and wives' CPS weights, which are not generally part of the CPS estimation process]).

Outline of the Second-Stage Calibration Algorithm

The second-stage calibration algorithm uses as its inputs the pre-second-stage weights $BW*DCF*NAF$ computed for each sampled person represented on a completed questionnaire in a SIPP panel.³ These weights are run through a series of adjustments, which result in a final weight (FNLWGT).⁴ This final weight can be written as $FNLWGT = SSCA*BW*DCF*NAF$, with SSCA (the second-stage calibration adjustment) equal to the ratio of the pre-second-stage weight and the final weight after the calibration process is completed.

This algorithm can be segmented into five major steps⁵:

1. Calibration of Hispanic children weights;
2. Calibration of non-Hispanic children weights;
3. Initial calibration steps for all adults;
4. Calibration of Hispanic adults; and
5. Calibration of non-Hispanic adults.

Each of these steps consists of numerous substeps. The next two sections describe certain steps that are common to all of the steps in the algorithm (the ratio adjustment step, the raking step, the cell-collapsing step, and the computation of control totals), the third section discusses details of

³ Children do not answer any SIPP questionnaires, but any children who are indicated as dependents by a sampled household receive weights in this process.

⁴ In pre-1996 Panels, households with all adults categorized as military personnel were interviewed and assigned weights (except for households in barracks, which are ineligible for SIPP). These households were not included in the second-stage calibration process (as they are not eligible for CPS and are not included in the CPS-derived control totals), and they received final weights equal to their pre-second-stage weights. For the 1996 Panel, these households are assigned as ineligible households and are not included in the weighting at all.

⁵ Separate runs of the calibration algorithm are made for each reference month and each rotation group (a total of 16 calibration runs for each panel wave).

particular calibration steps, and the last section describes steps that were carried out only for pre-1996 Panels.

Ratio Adjustments, Raking, and Cell Collapsing

The most important steps in the algorithm are the ratio adjustment and raking steps. Each ratio adjustment step takes all of the person weights (as they are at that point in the algorithm) within particular second-stage cells and multiplies them by a common ratio adjustment factor. The common factor is chosen for the second-stage cell so that the summation of the adjusted person weights within the cell equals the control total for that second-stage cell. The common ratio adjustment factor for each cell is equal to the control total divided by the summation of the current person weights for all sample persons in the cell.

The raking step is similar to the ratio adjustment step except that there are two sets of second-stage cells, with separate control totals (one set of second-stage cells is called the “row dimension,” and the other set is called the “column dimension”). At the end of the raking process (also called iterative proportional fitting), each person weight (as it is at that point in the algorithm) has been adjusted so that all person weights aggregate to the appropriate control totals for both the row cells and the column cells. The adjusted person weights have the property of aggregating within the second-stage cells to each control total while remaining as “close as possible” (in terms of a particular algebraic distance function) to the person weight values at the beginning of the raking step. Thus, the new person weights are consistent with both sets of independent control totals and have been altered as little as possible from the person weights before the step.

Most of the ratio adjustment and raking steps are preceded by a cell-collapsing step. This step is designed to prevent extreme alterations in the person weights (which will increase variability of the estimators) in any of the ratio adjustment and raking steps. Each second-stage cell is checked in its sample size: if the sample size is less than 35, then the cell is collapsed with a neighboring cell. The second-stage cells are also checked by computing the ratio adjustment for that cell. If that adjustment is less than 0.67 or greater than 2.0, then the cell is collapsed with a neighboring cell.

Ratio adjustments are computed for each set of second-stage cells before the raking process is performed. Ratio adjustments are computed for the row cells and the column cells as if only a ratio adjustment were being done for the row cells alone or the column cells alone, rather than a full raking step. If the computed ratio adjustments for any of the row cells are less than 0.67 or greater than 2.0, or the sample size for any row cell is less than 35, then the row cell is collapsed with a neighboring row cell. The same process is carried out for the column cells. All collapsing of this kind is completed before the raking step is executed.

When a second-stage cell is designated as requiring collapsing during the cell-collapsing step, the neighboring cell is chosen through a predetermined mechanism. Hispanic second-stage cells (see Figure C-1) are collapsed by sex (e.g., Hispanic males 15–24 are collapsed with Hispanic

females 15–24). The same is true for the household status second-stage cells for non-Hispanic children (the column dimension for non-Hispanic children; see Figure C-2). For the household status second-stage cells for adults (the column dimension for adults; see Figure C-3, pp. C-8 through C-11), the following pairs are collapsed when collapsing is necessary (the numbers in parentheses are the column numbers in the Figure C-3 tables):⁶

- Spouse in primary family (1); spouse in subfamily (3).
- Householder, no spouse present, in household with family (2); householder in household without a family (5).
- Not a spouse in household with family (4); not a householder in household without family (6).

For the age status second stage for adults (the row dimension for adults: see Figure C-3), neighboring cells are found on the basis of the scale value (which is given for the 1996 Panel in Figure C-3). The cell with the scale value closest to that of the cell that requires collapsing becomes the neighboring cell used in collapsing.

Figure C-1. Second-Stage Cells for Hispanics

Second-stage cells for Hispanic children

Male	Female

Second-stage cells for Hispanic adults⁷

Male			Female		
15–24	25–44	45+	15–24	25–44	45+

Second-stage cells for unmarried Hispanic adults

Male	Female

⁶ Collapsing is never done across black and nonblack status, or across sex, but only within the four primary groups: black males and females, and nonblack males and females (see Figure C-3).

⁷ Hispanic adults in the military are not defined as Hispanics in the computation of control totals or in the calculation of second-stage adjustments.

Figure C-2. Second-Stage Cells for Non-Hispanic Children

Second-Stage Cells for Black Children (14 years of age and younger)

MALES Age (years)	Children in Family Households	Children Not in Family Households	SCALE
Under 2			15
2 to 3			17
4 to 5			25
6 to 7			27
8 to 9			45
10 to 11			47
12 to 13			55
14			57

FEMALES Age (years)	Children in Family Households	Children Not in Family Households	SCALE
Under 2			15
2 to 3			17
4 to 5			25
6 to 7			27
8 to 9			45
10 to 11			47
12 to 13			55
14			57

Second-Stage Cells for Nonblack Children (14 years of age and under)

MALES Age (years)	Children in Family Households	Children Not in Family Households	SCALE
Under 1			15
1			17
2			25
3			27
4			45
5			47
6			55
7			57
8			75
9			77
10 to 11			85
12 to 13			105
14			107

FEMALES Age (years)	Children in Family Households	Children Not in Family Households	SCALE
Under 1			15
1			17
2			25
3			27
4			45
5			47
6			55
7			57
8			75
9			77
10 to 11			85
12 to 13			105
14			107

Figure C-3. Second-Stage Cells for Non-Hispanic Adults

Second-Stage Cells for Black Males (15+ years of age)

Age (years)	Persons in Households That Contain a Primary Family or Subfamily				Persons Not in Households Containing a Primary Family or Subfamily		SCALE VALUE
	Husband of Primary Family	Male Householder, No Spouse Present	Other Household Members		Householder	Not a Householder or Person in Group Quarters	
			Husband of Subfamily	Not a Husband			
15							15
16-17							16
18-19							18
20-21							27
22-24							29
25-29							47
30-34							49
35-39							57
40-44							59
45-49							63
50-54							65
55-59							83
60-64							85
65-69							93
70+							95

(figure continues)

The cell-collapsing procedure in some cases requires more than one iteration if cells after collapsing to the nearest neighbor are still too small or show extreme ratio adjustments (this generally occurs only in row-dimension collapsing for adults). New scale values are computed for the collapsed cells and are used to designate neighboring cells for any further collapsing that is necessary.

Computation of Control Totals

The control totals are equal to the CPS March-type estimates within each second-stage cell for some of the earlier ratio adjustment and raking steps in the algorithm.⁸ For the remaining ratio adjustment and raking steps, the control totals are derived by taking the CPS March-type estimate within the second-stage cell and subtracting from this the adjusted weights of any

⁸ For the 1984 and 1985 Panels, the control totals excluded people illegally residing in the United States. For the 1986 Panel and all panels following, the people are included in the control totals.

Figure C-3. Second-Stage Cells for Non-Hispanic Adults (*continued*)

Second-Stage Cells for Black Females (15+ years of age)

Age (years)	Persons in Households That Contain a Primary Family or Subfamily				Persons Not in Households Containing a Primary Family or Subfamily		SCALE VALUE
	Wife of Primary Family	Female House- holder, No Spouse Present	Other Household Members		House- holder	Not a Householder or Person in Group Quarters	
			Wife of Subfamily	Not a Wife			
15							15
16-17							16
18-19							18
20-21							27
22-24							29
25-29							47
30-34							49
35-39							57
40-44							59
45-49							63
50-54							65
55-59							83
60-64							85
65-69							93
70-74							94
75+							96

(figure continues)

subgroups whose weights have been completed. For example, control totals are derived for non-Hispanic children by taking the CPS March-type estimates for all children in each row cell and column cell (see Figure C-2) and subtracting the adjusted weights of all SIPP panel-rotation-group Hispanic children within that cell.

Details of the Calibration Steps

The first step (for Hispanic children) is a direct ratio adjustment to CPS control totals (using only two cells defined by sex). The second step (for non-Hispanic children) is a raking adjustment to derived controls; for row cells and column cells, the second-stage cells given in Figure C-2 are used. The derived control totals for each second-stage cell are equal to CPS control totals for all children in the cell minus the adjusted weights of all sampled Hispanic children in the cell.

Figure C-3. Second-Stage Cells for Non-Hispanic Adults (*continued*)

Second-Stage Cells for Nonblack Males (15+ years of age)

Age (years)	Persons in Households That Contain a Primary Family or Subfamily				Persons Not in Households Containing a Primary Family or Subfamily		SCALE VALUE
	Husband of Primary Family	Male Householder, No Spouse Present	Other Household Members		Householder	Not a Householder or Person in Group Quarters	
			Husband of Subfamily	Not a Husband			
15							15
16-17							16
18-19							18
20-21							27
22-24							29
25-29							47
30-34							49
35-39							57
40-44							59
45-49							63
50-54							65
55-59							83
60-64							85
65-69							93
70-74							95
75-79							103
80-84							104
85+							106

(figure continues)

Following the steps for children (which complete all second-stage adjustments for the children's weights) are the initial calibration steps for adults. Those steps are as follows:

1. A raking adjustment to CPS control totals that uses the Figure C-3 second-stage cells (the input weights are the pre-second-stage weights of all sampled adults);
2. A direct ratio adjustment to CPS control totals for sampled Hispanic adults; the input weights are the adjusted weights from step 1, and the second-stage cells are the cells given in Figure C-3 (for adults);
3. An equalization of all husbands' weights to their wives' weights (so that spouses in one family have equal weights);
4. A second raking adjustment identical to step 1 except that the input weights are the adjusted weights after steps 1 through 3 are completed;
5. A second Hispanic adult ratio adjustment identical to step 2 except that the input weights are the Hispanic adult adjusted weights from step 4.

Figure C-3. Second-Stage Cells for Non-Hispanic Adults (continued)

Second-Stage Cells for Nonblack Females (15+ years of age)

Age (years)	Persons in Households That Contain a Primary Family or Subfamily				Persons Not in Households Containing a Primary Family or Subfamily		SCALE VALUE
	Wife of Primary Family	Female House- holder, No Spouse Present	Other Household Members		House- holder	Not a Householder or Person in Group Quarters	
			Wife of Subfamily	Not a Wife			
15							15
16-17							16
18-19							18
20-21							27
22-24							29
25-29							47
30-34							49
35-39							57
40-44							59
45-49							63
50-54							65
55-59							83
60-64							85
65-69							93
70-74							95
75-79							103
80-84							104
85+							106

The next two steps complete the weights for Hispanic adults. The first step is an equalization of all husbands' weights in married couples, including at least one Hispanic, to their wives' weights. The exception to this is when the wife is not Hispanic, in which case the wife's weight is set equal to the husband's weight. At this point, all married couples including at least one Hispanic have their final weights. The second step is a ratio adjustment for sampled unmarried Hispanics (only males and females are used as second-stage cells) to derived control totals, which are CPS control totals for all Hispanic adults minus the adjusted weights of the sampled married Hispanics.

The last steps complete the calibration process for sampled non-Hispanic adult weights. Those steps are as follows:

6. An equalization of wives' weights to their husbands' weights.
7. A raking adjustment to derived control totals that uses the Figure C-3 second-stage cells (the input weights are the current adjusted weights of all non-Hispanic adults). The control totals are the CPS control totals for all adults for the second-stage cells minus the adjusted weights of Hispanic adults within those cells.
8. An equalization of husbands' weights to their wives' weights. This step finalizes the weights for all non-Hispanic females and all non-Hispanic husbands.
9. A raking adjustment to derived control totals; the Figure C-3 second-stage cells for adult males (with the two husband columns deleted) are used, and the current adjusted weights of all non-Hispanic nonhusband males are used. The derived control totals are the CPS control totals minus the adjusted weights of all groups who have had their weights completed. This step produces the final weights for all non-Hispanic nonhusband male adults (the last group without completed weights).

Weighting Factors Used in Panels Prior to 1996

In all panels prior to the 1996 Panel, a first-stage ratio estimate factor (FSF) was applied to the base weight of each person in non-self-representing PSUs (i.e., PSUs not sampled with certainty). This first-stage factor was a ratio adjustment step that used as cells Census region, residence status, and race; it was designed to reduce the variance resulting from sampling of PSUs. Although this factor is no longer computed in the 1996 Panel, the cells are now used in the computation of noninterview adjustment factors.

Also, beginning with the 1985 Panel, a new construction noninterview adjustment factor (NCF) was applied to the base weight of new households in new construction housing-unit clusters. This factor was used to account for newly constructed housing units that were selected for the sample but were unavailable for interviewing. It was set equal to 1 in the 1986–1993 Panels (it was not used in the 1984 Panel), and eventually it was discontinued.

Thus, in the 1984 Panel, FNLWGT was equal to $BW*DCF*HNF*FSF*SSCA$ (excludes NCF). FNLWGT was equal to $BW*DCF*NCF*HNF*FSF*SSCA$ in the 1985–1993 Panels.

Wave 2+ Weights

The later wave cross-sectional weight is computed separately for each reference month of each wave. This Wave 2+ FNLWGT has the following factors for people in households whose residents have not changed from Wave 1: an initial weight (IW), a later wave noninterview

adjustment (LWNIA), and a second-stage calibration adjustment (SSCA). The initial weight is generally equal to the pre-second-stage weight for the Wave 1 household weight (with some exceptions). For households that have had people move into or out of the household after Wave 1, there is an adjustment to the initial weight called the mover's weight (MW). For these people, the cross-sectional weight has as factors the mover's weight, the later wave noninterview adjustment, and the second-stage calibration adjustment. In summary, people in households that do not need mover's adjustments receive the cross-sectional weight $FNLWGT = IW * LWNIA * SSCA$, and persons in households that do require a mover's adjustment receive the Wave 2+ final weight $FNLWGT = MW * LWNIA * SSCA$.

Wave 2+ Initial Weights

The initial weight is essentially the pre-second-stage Wave 1 weight, that is, $IW = BW * DCF * NAF$.⁹ The second-stage calibration adjustment for the Wave 1 reference months is not included as a factor: the second-stage calibration adjustment is redone using control totals current for the later wave reference months. The initial weight allows the original sample person to represent unsampled persons in the population and persons in households who were not successfully interviewed in Wave 1. The initial weight does not generally change from wave to wave after Wave 1, unless special circumstances arise that cause an alteration in the panel sample (such as a cut in the sample for budgetary or other reasons).

Movers' Weights

People in any households that an original sample person enters during later waves, or any people who become part of a Wave 1 sample household during later waves, also become part of the sample for those waves. If the original sample person moves away from the household containing those people, the additional people immediately drop from the sample (their in-sample status in any given wave is entirely dependent on the presence of original sample persons in the household). Any of the additional people who were part of the SIPP population in Wave 1 (and therefore could have been sampled) and who become members of households with original sample persons are called associated sample persons. If any of these additional persons were not part of the SIPP population in Wave 1 (because they were out of the country, institutionalized, etc.), then they are called additional sample persons.

⁹ The 1985 Panel had an initial weight that was computed differently. The initial weight for this panel included a new-construction noninterview adjustment factor and a first-stage ratio estimate factor. The Wave 1 noninterview adjustment factor was also recomputed in the 1985 Panel to account for sampled households mistakenly left off the sample roster during Wave 1, and sampled households that were noncooperative in Wave 1 but were converted during Wave 2. There was also an added "sample cut" factor, adjusting for sampled households that were deselected because of a reduction in the 1985 Panel sample. Pre-1996 Panels following 1985 had only one difference from the 1996 Panel initial weight described in the text: the presence of the first-stage ratio estimate factor.

Any household that consists of people who were in the SIPP universe who lived in separate households during the Wave 1 reference period (with at least one of the households sampled in Wave 1) is called an enhanced household. In most cases, an enhanced household consists of original sample persons from a Wave 1 sample household and associated sample persons from a household (or households) not sampled in Wave 1. In a few rare cases, an enhanced household will contain original sample persons from more than one Wave 1 sample household. Those households are rare because the probability of selection of any given household in SIPP is quite small, making the joint probability of a later wave merged household having two or more of its Wave 1 predecessor households selected in Wave 1 quite small (but the situation does occur in the SIPP panels).

Enhanced households require an adjustment of the Wave 1 base weight for each person in the household. These people in effect had multiple chances of being in the selected enhanced household: they could have been selected as original sample persons in the household they were in during Wave 1 (which then became an enhanced household), or they could become an associated sample person if their Wave 1 household was not selected but merged later with a sampled Wave 1 household. Their true probability of being included in the enhanced household is higher than their nominal Wave 1 probability of selection, and their assigned base weight should be the reciprocal of this true sample inclusion probability.

This true inclusion probability is not computed directly, for it requires the computation of joint probabilities of selection of multiple households, some of which were not in the original Wave 1 household sample. Instead, a “mover’s weight” is assigned to each original and associated sample person in the enhanced household, which has as its expectation the inverse of the true sample inclusion probability. In other words, the movers’ weights are unbiased weights, taking into account the complex realized sample design for enhanced households.

In the case in which an enhanced household is formed from only one Wave 1 sample household (with associated persons added to it), the mover’s weight for each person in the household (original, associated, or additional) is computed as follows for reference month t , enhanced household i :

$$W_{ti} = \frac{W_{1i} S_{1ti}}{S_{ti} - S_{tai}}, \quad (\text{C-2})$$

where W_{1i} is the initial weight that is common to all original sample persons in the i th enhanced household, S_{1ti} is the number of original sample persons in the i th enhanced household in month t , S_{ti} is the size of the i th enhanced household in month t (all persons), and S_{tai} is the number of additional sample persons in the i th enhanced household in month t . The numerator of this expression is the sum of the initial weights over all original sample persons in the household during month t , and the denominator of this expression is the number of original and associated sample persons in the i th enhanced household in month t . For a discussion of why these are unbiased weights, see, for example, Kalton and Brick (1994).

When two Wave 1 sample households merge, the mover's weight for each sample person (original, associated, or additional) in the household is computed as follows:

$$W_{ti} = \frac{W_{1i}S_{1ti} + W'_{1i}S'_{1ti}}{S_{ti} - S_{tai}}. \quad (\text{C-3})$$

The two terms in the numerator are for the first and second Wave 1 sample households. The movers' weights for more than two merged Wave 1 sample households are computed analogously.

Wave 2+ Later Wave Noninterview Adjustments

The initial weights have an adjustment for noncooperation in Wave 1; that is, the sample households with nonzero initial weights represent households for which an interview was not completed in Wave 1. There are, however, further losses of sample households in later waves for several reasons:

- The household refuses to cooperate in some or all of the later waves.
- The people in the household have moved and cannot be found.
- The household has moved, and has been found, but is too far away for a personal interview and cannot be reached by telephone.¹⁰

The weights of households for which later wave interviews are completed are adjusted to “represent” sample households (who cooperated in Wave 1) whose interviews are not completed for any of the above reasons. Those adjustments are computed by assigning each sample household with a nonzero initial weight to one of 109 later wave noninterview cells.¹¹ The noninterview cells are based on the following household characteristics:

1. Reference person is a non-Hispanic white person, or other (two categories).
2. Reference person is a female householder without a spouse and with her own children, a householder 65 years of age or older, or other (three categories).
3. Household income includes welfare payments (AFDC, WIC, Food Stamps, Medicaid, or other welfare), or not (two categories).
4. Household size is 1, 2, 3, or 4 or more persons (four categories).
5. Household has some bond-type financial assets, or not (two categories).

¹⁰ The SIPP sample is designed so that most of the field work takes place within the SIPP PSUs, to reduce traveling costs. If a household moves too far away from the field areas, a telephone interview is attempted.

¹¹ In pre-1996 Panels, 53 noninterview cells were used, based on the first 7 of the 10 listed household characteristics.

6. Reference person's education level is less than 8 years, 8 to 11 years, 12 to 15 years, or 16 or more years (four categories).
7. Household owns housing unit, is renter, or is living in a public housing project or receiving a rent subsidy from the government (three categories).
8. Census division (nine categories).
9. Number of imputations in household Wave 1 questionnaire is none, 1, or more than 1 (three categories).
10. Household income as a percentage of the household poverty threshold (with both averaged over 4 reference months): less than or equal to 175 percent, 176 through 450 percent, and more than 450 percent (three categories).

These categories have been found in empirical research to be consistently heterogeneous in later wave noninterview rates (i.e., the categories have divergent noninterview rates). The later wave noninterview adjustment for each noninterview cell is equal to the sum of the initial or mover's weights of all households that have had the later wave interview completed, divided by the sum of the initial or mover's weights of all Wave 1 sample households.¹² (The mover's weight is used whenever a mover's weight is computed for the household.) These adjustments are made separately for each reference month of each later wave of the panel.

Before the final noninterview adjustment is computed for each wave, each noninterview cell is checked. Any noninterview cell with fewer than 30 interviewed households, or with a noninterview adjustment greater than 2, is collapsed with a neighboring cell. Cells are defined as neighboring on the basis of a set of scale values assigned to each noninterview cell. This procedure prevents extreme noninterview adjustments from being made (which will increase sampling variability). The final noninterview adjustment (LWNIA) for the cell, or collapsed cell, is assigned to each household within the cell.

Table C-1 presents the major groupings of noninterview cells (the noninterview cells within these major groupings have similar scale values and would be collapsed together within these groupings before any collapsing was done across groupings).

Wave 2+ Second-Stage Calibration Adjustment (SSCA)

A second-stage calibration adjustment is carried out for each reference month in each later wave, for each rotation group of the panel separately. This adjustment uses the same algorithm as described for Wave 1 weights, with new CPS or CPS-derived control totals computed for each

¹² In pre-1996 Panels, general quarters households were not included in these calculations and receive noninterview adjustments equal to 1. In the 1996 Panel, these households are treated in the same way as family households in noninterview calculations, but households with only military adults were included.

**Table C-1. Major Groupings of Later Wave
Noninterview Cells**

Household Characteristics	Number of Nonresponse Cells
Hispanic or nonwhite	
Minimal assets	15
Assets include bonds	9
White Non-Hispanic	
Single female householder	1
Householder 65 and older	14
Other householder	
No welfare income	
One person in household	20
Two people in household	14
Three people in household	7
Four or more in household	19
Has welfare income	10
Total	109

new reference month. The pre-second-stage weights in this case are $IW * LWNIA$, or $MW * LWNIA$ if a mover's weight was computed for the household. The second-stage calibration adjustments reduce sampling variability by calibrating the final weights to agree with independent control totals. With the later wave cross-sectional weights, the second-stage calibration adjustments also have the effect of reducing biases from population undercoverage (arising from eligible people entering the U.S. population after the Wave 1 reference months).

Calendar Year and Panel Weights

The algorithm for generating the calendar year and panel weights is very similar to that used for computing Wave 2+ weights, with some differences. The most important differences are the following:

- A control date is associated with each calendar year and panel weight (rather than the weight being associated with a month, as with the Wave 1 and Wave 2+ weights).
- For a sample person to have a nonzero weight, data must be present for the sequence of months defined for the weight (12 months for the calendar year weights and all months of the panel for the panel weights). Months for which the sample person is ineligible are excluded from this check.

Calendar Year and Panel Initial Weights

The initial weight computed for each sample person for all calendar year and panel weights is $IW = BW * DCF * NAF$, that is, the same quantity that is used as the initial weight for all Wave 2+ weights. This initial weight allows each original sample person who has interviews for the months for which they are eligible in the calendar year (or panel) to represent unsampled people in the population and people in households that were not successfully interviewed in Wave 1.

Calendar Year and Panel Noninterview Adjustments

The noninterview adjustments for each calendar year and panel weight are computed by first assigning each sampled person with a nonzero initial weight to one of 149 noninterview cells.¹³ These noninterview cells are based on the following person-level characteristics:

1. Person is a non-Hispanic white person, or other (two categories).
2. Person was self-employed, or not (two categories).
3. Family income was a percentage of the family poverty threshold (with both averaged over 4 reference months): less than or equal to 175 percent, 176 through 450 percent, and more than 450 percent (three categories).¹⁴
4. Person in household whose income includes welfare payments (SSI, AFDC, WIC, Food Stamps, Medicaid, or other welfare), person receiving unemployment compensation but not in household with welfare payments, or neither (three categories).
5. Person in household with some bond-type financial assets, or not (two categories).
6. Person's education level is less than 12 years, 12 to 15 years inclusive, or 16 or more years (three categories).
7. Person was in labor force at least 1 month of wave, or not (two categories).
8. Census division of household (nine categories).
9. Number of imputations in household Wave 1 questionnaire is none, 1, or more than 1 (three categories).
10. Within PSU, stratum code of household is poverty stratum or nonpoverty stratum (two categories).

¹³ In pre-1996 Panels, 126 noninterview cells were used, based on the first 7 of the 10 listed person characteristics.

¹⁴ In pre-1996 Panels, household income (averaged over 4 reference months) was used instead: less than \$1,200 a month, between \$1,200 and \$4,000 a month, and greater than or equal to \$4,000 a month.

These categories have been found in empirical research to be consistently heterogeneous in later wave noninterview rates. The noninterview adjustment for the noninterview cell (for the particular calendar year [panel] weight) is equal to the sum of the initial weights of all sampled persons whose households were interviewed in Wave 1,¹⁵ divided by the sum of the initial weights of all sampled persons who have interviews for every month of the calendar year (panel) in which they are eligible.¹⁶

As with other noninterview adjustments discussed in this appendix, each noninterview cell is checked for small sample sizes and extreme noninterview adjustments. Any noninterview cell with fewer than 30 sampled persons with complete interview strings, or with a calendar year (panel) noninterview adjustment greater than 2, is collapsed with a neighboring cell for that calendar year and panel weight. If necessary, this process can be iterative: a cell may be collapsed into another cell, and then the combined cell may be collapsed further with other cells. A set of scale values determines how cells are collapsed when collapsing is necessary. Table C-2 presents the major groupings of noninterview cells (i.e., the noninterview cells with similar scale values). The noninterview cells within these groupings would be collapsed together among themselves before any collapsing would be done outside of these groupings.

Table C-2. Major Groupings of Calendar Year (Panel) Noninterview Cells

Person Characteristics	Number of Nonresponse Cells
Hispanic or nonwhite	50
White Non-Hispanic	
Less than 12 years of education	25
12 to 15 years of education	
In labor force	32
Not in labor force	18
16 or more years of education	24
Total	149

¹⁵ People who entered the sample during or after the calendar year (panel) period (by entering a sampled household) are excluded from these calculations (and receive calendar year [panel] weights of zero). Children who move without their parents (into nonsampled households) during the period are also excluded from these computations and receive calendar year (panel) weights of zero.

¹⁶ In pre-1996 Panels, sample persons living in group quarters are not included in these noninterview adjustments, and those people are given noninterview adjustments equal to 1 (when their calendar year and panel weights are nonzero). In the 1996 Panel, sample persons living in group quarters are treated in the same way as other sample persons.

Calendar Year and Panel Second-Stage Adjustments

The calendar year and panel weights that have been computed up to this point (called the pre-second-stage weights) for each sampled person (with a complete set of interviews for their eligible months) are equal to $BW*DCF*NAF*LWNIA$. The formula for the final calendar year weights (FNLWGT) is $BW*DCF*NAF*LWNIA*SSCA$, where SSSA is the second-stage calibration adjustment. The final panel weight follows the same formula: $PNLWGT = BW*DCF*NAF*LWNIA*SSCA$, though LWNIA and SSSA are computed differently here. The final weight is computed in both cases from the pre-second-stage weights $BW*DCF*NAF*LWNIA$ in accordance with the algorithm described below. As with the Wave 1 and Wave 2+ weights, the algorithm for second-stage adjustment for calendar year and panel weights can be segmented into the following five major steps:

1. Calibration of Hispanic children weights;
2. Calibration of non-Hispanic children weights;
3. Initial calibration steps for all adults;
4. Calibration of Hispanic adults; and
5. Calibration of non-Hispanic adults.

However, the actual steps within these five major steps are different in their details for calendar year (panel) weights. The primary difference between the calendar year (panel) weights second-stage calibration algorithm and the Wave 2+ weights second-stage calibration algorithm is that a married couple weighting equalization is not done for the calendar year (panel) weights, and married and unmarried persons are not separated out for separate calibration steps in the calendar year (panel) weights algorithm.

The independent estimates for the control month are the same CPS March supplement-type estimates that were used for the Wave 2+ weights, except they are computed for different second-stage cells when used for calendar year (panel) weights. The second-stage cells for calendar year (panel) weights are given in Figures C-4, C-5, and C-6. The second-stage calibration algorithm is run separately for each rotation group, with the control totals for each rotation group equal to one-quarter of the CPS control totals.

Figure C-4. Calendar Year and Panel Weight Second-Stage Cells for Hispanics

Second-Stage Cells for Hispanics (14 years and younger)

Male	Female

Second-Stage Cells for Hispanics (15+ years of age)¹⁷

Male			Female		
15-24	25-44	45+	15-24	25-44	45+

Figure C-5. Calendar Year and Panel Weight Second-Stage Cells for Non-Hispanic Children

Cells for Children (14 years and younger)

Age	Nonblack Males	Nonblack Females	Black Males	Black Females	SCALE
Under 2					15
2 to 3					17
4 to 5					25
6 to 7					27
8 to 9					45
10 to 11					47
12 to 13					55
14					57

¹⁷ Hispanic adults in the military are not defined as Hispanics in the computation of control totals or in the calculation of second-stage adjustments.

Figure C-6. Calendar Year and Panel Weight Second-Stage Cells for Non-Hispanic Adults

1996 Panel Second-Stage Cells for Nonblack Females (15+ years of age)

Age (years)	Householder				Not Householder			SCALE VALUE
	1. Female Householder No Spouse Present with Own Children	2. Other Female Householder No Spouse Present	3. Other Female Householder Living with Relative	4. Female Householder Not Living with Relative	6. Spouse of Householder or Spouse of Related Subfamily	7. Other Female Related to Householder	9. Other Female Not Related to Householder	
15								15
16-17								16
18-19								18
20-21								27
22-24								29
25-29								47
30-34								49
35-39								57
40-44								59
45-49								63
50-54								65
55-59								73
60-61								74
62-64								76
65-69								93
70-74								95
75-79								103
80-84								104
85+								106

(figure continues)

Details of the Calendar Year and Panel Second-Stage Calibration Steps

The individual steps in the calendar year (panel) second-stage calibration algorithm are generally the same as the corresponding steps in the Wave 1 and Wave 2+ second-stage calibration

Figure C-6. Calendar Year and Panel Weight Second-Stage Cells for Non-Hispanic Adults (*continued*)

1996 Panel Second-Stage Cells for Black Females (15+ years of age)

Age (years)	Householder			Not Householder			SCALE VALUE
	2. Female Householder No Spouse Present	3. Other Female Householder Living with Relative	4. Female Householder Not Living with Relative	6. Spouse of Householder or Spouse of Related Subfamily	7. Other Female Related to Householder	9. Other Female Not Related to Householder	
15							15
16-17							16
18-19							18
20-21							27
22-24							29
25-29							47
30-34							49
35-39							57
40-44							59
45-49							63
50-54							65
55-59							73
60-61							74
62-64							76
65-69							93
70-74							94
75+							96

(figure continues)

algorithm.¹⁸ The differences in the two calibration algorithms are primarily the second-stage cells, with some other minor differences, as described in this section.

The first step (for Hispanic children) is a ratio adjustment to CPS control totals that uses only the two cells defined by sex (this step is identical to the Wave 1 and Wave 2+ algorithm step for Hispanic children). The second step (for non-Hispanic children) is a ratio adjustment step to derived controls that uses as cells the second-stage cells given in Figure C-5.

¹⁸ The cell-collapsing procedures described for the Wave 1 and Wave 2+ weights are also used as stated in that section for the calendar year and panel weights, except for the column dimension collapsing for non-Hispanic adults. For calendar year and panel weights, and for any of the four race/sex groups given in Figure C-6, columns 1 and 2 (see Figure C-6 for the numbering of the columns) are collapsed if either does not meet the criterion (which is the same as described in the earlier section on ratio adjustment, raking, and cell collapsing), column 4 is collapsed with column 2 if it does not meet the criterion, column 7 is collapsed with column 9 if either does not meet the criterion, and column 8 is collapsed with column 10. Collapsing of columns 3, 5, and 6 and further collapsing of the other columns should never be necessary in practice.

Figure C-6. Calendar Year and Panel Weight Second-Stage Cells for Non-Hispanic Adults (*continued*)

1996 Panel Second-Stage Cells for Nonblack Males (15+ years of age)

Age (years)	Householder		Not Householder			SCALE VALUE
	3. Male Householder Living with Relative	5. Male Householder Not Living with Relative	6. Spouse of Householder or Spouse of Related Subfamily	8. Other Male Related to Householder	10. Other Male Not Related to Householder	
15						215
16-17						216
18-19						218
20-21						227
22-24						229
25-29						247
30-34						249
35-39						257
40-44						259
45-49						263
50-54						265
55-59						273
60-61						274
62-64						276
65-69						293
70-74						295
75-79						303
80-84						304
85+						306

(figure continues)

Following these steps for children (which complete all second-stage adjustments for the children's weights) are the initial calibration steps for adults. Those steps are as follows:

1. A raking adjustment to CPS control totals that uses the Figure C-6 second-stage cells; the input weights are the pre-second-stage weights of all sampled adults.
2. A direct ratio adjustment to CPS control totals for sampled Hispanic adults; the input weights are the adjusted weights from step 1, and the second-stage cells are the cells given in Figure C-4 (for adults).
3. A second raking adjustment identical to step 1 except that the input weights are the adjusted weights after steps 1 and 2 are completed.

Figure C-6. Calendar Year and Panel Weight Second-Stage Cells for Non-Hispanic Adults (*continued*)

1996 Panel Second-Stage Cells for Black Males (15+ years of age)

Age (years)	Householder		Not Householder			SCALE VALUE
	3. Male Householder Living with Relative	5. Male Householder Not Living with Relative	6. Spouse of Householder or Spouse of Related Subfamily	8. Other Male Related to Householder	10. Other Male Not Related to Householder	
15						215
16-17						216
18-19						218
20-21						227
22-24						229
25-29						247
30-34						249
35-39						257
40-44						259
45-49						263
50-54						265
55-59						273
60-61						274
62-64						276
65-69						293
70+						295

4. A second Hispanic adult ratio adjustment identical to step 2 except that the input weights are the Hispanic adult adjusted weights from step 3.

At this point, the weights are completed for Hispanic adults. The final step is a raking adjustment to derived control totals that uses the Figure C-6 second-stage cells. The derived control totals are the CPS control totals for all adults for the second-stage cells minus the adjusted weights of Hispanic adults within those cells. The input weights are the current adjusted weights for non-Hispanic adults.