Attachment J

U.S. Census Bureau Report "Evaluation of the 2022 December Food Security Supplement Test Noninterview Adjustment Using Logistic Regression"

UNITED STATES DEPARTMENT OF COMMERCE
U.S. Census Bureau
Washington, DC 20233-0001

August 9, 2024

MEMORANDUM FOR   Michele Ver Ploeg
    Chief, Food Assistance Branch
    U.S. Department of Agriculture Economic Research Service

Through: Kyra M. Linse
    Survey Director for Current Population Survey & Time Use
    Associate Director for Demographic Programs – Survey Operations

From: Anthony G. Tersine, Jr. *Anthony G. Tersine, Jr.*
    Chief, Demographic Statistical Methods Division

Subject: Evaluation of the 2022 December Food Security Supplement Test
    Noninterview Adjustment Using Logistic Regression

This memorandum includes documentation of the evaluation of the 2022 Food Security Supplement test weighting procedure for the noninterview adjustment using logistic regression compared to the original weighting procedure.

The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data (Project No. P-7527681, Disclosure Review Board (DRB) approval number: CBDRB-FY24-POP001-0078).

If you should have any questions or need additional information, please contact Jana Hatch at 301-763-2230 or jana.s.hatch@census.gov or Dave Hornick at 301-763-4183 or david.v.hornick@census.gov.

Attachment: Evaluation of the December 2022 Food Security Supplement Test
    Noninterview Adjustment Using Logistic Regression

cc:
Michael Brennan   (ADDP)
Greg Weyland
Ibrahim Keita
James Farber   (DSMD)
David Hornick
Rebecca Hoop
Weimin Zhang
Jana Hatch
Kiauna Womack

census.gov

[This page intentionally left blank.]

Demographic Statistical Methods Division

Survey Statistics for Employment, Education, Crime, and Housing

Evaluation of the 2022 December Food Security Supplement Test Noninterview Adjustment Using Logistic Regression

August 9, 2024

Jana Hatch

James Farber, ADC

David Hornick, Lead Scientist

Evaluation of the 2022 December Food Security Supplement Test Noninterview Adjustment Using Logistic Regression

The Food Security supplement (FSS) to the Current Population Survey[1] (CPS), the original, or current, noninterview adjustment method uses basic CPS noninterview clusters to adjust for nonresponse. For purposes of this evaluation, it will be referred to as *orig_NIWGT*.

The test, or new, noninterview adjustment method for the FSS uses logistic regression to group records into response propensity cells and calculates an adjustment factor, total weight/response weight, for each cell. This ratio of total weight to response weight is the noninterview adjustment factor. The following variables are used in the logistic regression model: region (*GEREG*), state (*GESTFIPS*), age (*PEAGE*), sex (*PESEX*), race (*PRWTRACE*), Hispanic origin (*PEHSPNON*), family income (*HEFAMINC*), educational attainment (*PEEDUCA*), and labor force status (*PRCIVLF*). For purposes of this evaluation, this weight will be referred to as *test_NIWGT*.

In the current method, the noninterview clusters are defined by geography and metropolitan status (in principal city, not in principal city, not metropolitan), with no demographic variables. The new method incorporates ethnicity, race, age, and sex variables, as well as three socio-economic variables, to try to better capture and adjust for groups that may tend to not respond to the survey.

FSS data collected in December 2022 is used to evaluate the test method by comparing it to the original method through ratios and distributions. Attachment A contains SAS programs used for the evaluation.

Evaluation Part 1: Coverage Ratios By Demographic Group

Coverage ratios are calculated as the weighted population estimate before poststratification divided by the independent population control. Table 1 displays coverage ratios by ethnicity (*PEHSPNON*), race (*PRWTRACE*), age group (*PEAGE*), and sex (*PESEX*), where the noninterview adjustment weights are used in the numerator. Comparing the coverage ratios using *test_NIWGT* to the coverage ratios using *orig_NIWGT* can show how the new method improves coverage for these demographic groups.

---

[1] More information on confidentiality protection, methodology, sampling and nonsampling error, and definitions is available at <www2.census.gov/programs-surveys/cps/techdocs/cpsdec22.pdf>.

Table 1. Coverage Ratios using *orig_NIWGT* and *test_NIWGT* by Demographic Group: December 2022

Characteristic		Coverage Ratios Using *orig_NIWGT0*	Coverage Ratios Using *test_NIWGT0*	Comparison of Coverage Ratios P-value
Ethnicity (*PEHSPNON*)	1 = Hisp	0.85	0.86	< 0.01*
	2 = Non-Hisp	0.90	0.91	< 0.01*
Race (*PRWTRACE*)	1 = White Alone	0.94	0.94	< 0.01**
	2 = Black Alone	0.72	0.75	< 0.01*
	3 = Asian Alone 4 = Other⁺	0.79	0.82	< 0.01*
Age (grouped) (*PEAGE*)	1 = <18	0.84	0.85	< 0.01*
	2 = 18-30	0.80	0.81	< 0.01*
	3 = 31-64	0.89	0.89	< 0.01*
	4 = 65+	1.06	1.06	0.14
Sex (*PESEX*)	1 = male	0.88	0.89	< 0.01*
	2 = female	0.91	0.91	< 0.01*

Source: U.S. Census Bureau, Current Population Survey, Food Security, December 2022

* Indicates the coverage ratio using *test_NIWGT* is significantly higher than the coverage ratio using *orig_NIWGT* at the 0.1 significance level.

** Indicates the coverage ratio using *test_NIWGT* is significantly lower than the coverage ratio using *orig_NIWGT* at the 0.1 significance level.

⁺ Asian alone and Other race categories are combined to match the population controls for residual race.

Note: Differences may not be apparent due to rounding.

The coverage ratios for most of the demographic groups show a significant increase in coverage when using the new noninterview adjustment weight. The White Alone category is the only group to have a significant decrease in coverage, showing that the adjustment is targeting the remaining race groups because they tend to have higher nonresponse. The other group that didn't have a significant increase in coverage consisted of those aged 65 and older. This age category did not have a significant difference between coverage ratios using *orig_NIWGT* and *test_NIWGT*, showing again that the new noninterview adjustment is not targeting this group.

Evaluation Part 2: Nonadjustment Weight Ratios By Demographic Group

Records are combined into four groups of cells based on region (*GEREG*), state (*GESTFIPS*), race (*PRWTRACE*), ethnicity (*PEHSPNON*), age (*PEAGE*), and sex (*PESEX*). Ages from *PEAGE* are grouped into <18, 18-30, 31-64, and 65+.

Group 1: *GESTFIPS, PEHSPNON, PEAGE, PESEX* – 816 cells

Group 2: *GESTFIPS, PWRTRACE, PEAGE, PESEX* – 1,632 cells

Group 3: *GEREG, PEHSPNON, PEAGE, PESEX* – 64 cells

Group 4: *GEREG, PWRTRACE, PEAGE, PESEX* – 128 cells

Significance Tests of Ratios

The following steps show the process used to evaluate the test noninterview weight (*test_NIWGT*) by demographic group using significance tests. The steps were completed for each of the four groups identified above.

Step 1. Sum *orig_NIWGTi* by cell group, where *i* = 0 to 160. This will result in 161 sums for each cell.

Step 2. Sum *test_NIWGTi* by cell group, where *i* = 0 to 160. This will result in 161 sums for each cell.

Step 3. For each replicate in each cell, calculate the ratio of *test_NIWGTi* to *orig_NIWGTi*.

$$\hat{\theta}_i = \frac{test\_NIWGTi}{orig\_NIWGTi}$$

where *i* = 0 is the full weight and *i* = 1 to 160 are the replicate weights.

Step 4. Calculate the standard error of the full weight ratio, $\hat{\theta}_0$, using the replicate ratios, $\hat{\theta}_i$, where *i* = 1 to 160.

$$Var(\hat{\theta}_0) = \frac{4}{160} \sum_{i=1}^{160} (\hat{\theta}_i - \hat{\theta}_0)^2$$

$$SE(\hat{\theta}_0) = \sqrt{Var(\hat{\theta}_0)}$$

Step 5. Perform a significance test comparing the full-weight ratio to 1.

$$Stat(\hat{\theta}_0) = \frac{\hat{\theta}_0 - 1}{SE(\hat{\theta}_0)}$$

If $Stat(\hat{\theta}_0)$ > 1.645, then the ratio is significantly higher than 1 at the 0.1 level: *test_NIWGT0* is adjusting the cells differently than *orig_NIWGT0*.

Table 2 shows the results of the significance tests[2] by showing the percentage of cells in each group that had ratios that were significantly different from 1. The groups using states (*GESTFIPS*) had smaller percentages of significant cells because the cells were much smaller, with some cells even empty, due to being split into 50 states and the District of Columbia, as opposed to four regions. The distributions of the significant cells are found by group in Tables 3.A – 6.A.

---

[2] All comparative statements in this report have undergone statistical testing, and, unless otherwise noted, all comparisons are statistically significant at the 0.1 significance level.

Table 2. Significance Test Results of Ratios by Demographic Groups: December 2022

Variables	Number of cells[+]	Percent significantly higher than one	Percent not significantly higher than one
Group 1: *GESTFIPS, PEHSPNON, PEAGE, PESEX*	794[+]	9.70	90.30
Group 2: *GESTFIPS, PWRTRACE, PEAGE, PESEX*	1,499[+]	16.34	83.66
Group 3: *GEREG, PEHSPNON, PEAGE, PESEX*	64	39.06	60.94
Group 4: *GEREG, PWRTRACE, PEAGE, PESEX*	128	42.19	57.81

Source: U.S. Census Bureau, Current Population Survey, Food Security, December 2022

[+] Some cells were empty due to the high number of cells; therefore this number does not match up to the number of possible cells listed above.

Compare Distributions

For each group, distributions were also explored for ratios in the top and bottom percentiles (10 percent for state groups, 25 percent for region groups) and ratios that were significantly higher than one. These were compared to the distributions of all unweighted cases for the demographic variables (designated by shading) to show how the new noninterview adjustment is affecting the weights for certain demographic groups. The new method hopes to target certain demographic groups that tend to have higher nonresponse and weight them up.

The distributions and comparisons can be found in Tables 3.A. – 6.A. Distributions of the geographic variables (state and region) were not evaluated.

Table 3.A. Distributions of Demographic Characteristics of Group 1 Cells State/Ethnicity/Age/Sex, Unweighted and With Certain Ratios of *test_NIWGT0* to *orig_NIWGT0*: December 2022

Characteristic		Total un-weighted	Cells with ratios significantly higher than 1		Cells with high ratios (above 90th percentile)		Cells with low ratios (below 10th percentile)	
		Percent	Percent	P-value	Percent	P-value	Percent	P-value
Ethnicity (*PEHSPNON*)	1 = Hisp	15.64	48.05	< 0.01*	79.75	< 0.01*	58.23	< 0.01*
	2 = Non-Hisp	84.36	51.95		20.25		41.77	
Age (*PEAGE* grouped)	1 = <18	20.86	20.78	< 0.01*	18.99	< 0.01*	22.78	0.30
	2 = 18-30	14.72	44.16		36.71		18.99	
	3 = 31-64	43.22	24.68		21.52		32.91	
	4 = 65+	21.20	10.39		22.78		25.32	
Sex (*PESEX*)	1 = male	48.68	59.74	0.05*	58.23	0.09*	46.84	0.74
	2 = female	51.32	40.26		41.77		53.16	

Source: U.S. Census Bureau, Current Population Survey, Food Security, December 2022

* Indicates distribution is significantly different from the total unweighted distribution at the 0.1 significance level.

Note: Within a category, percents may not sum to 100 due to rounding.

Table 3.B. Summary Statistics for Ratios of *test_NIWGT0* to *orig_NIWGT0* for State/Ethnicity/Age/Sex Cells: December 2022

Total cells:	794[+]
Number of cells in 10% tails:	79
Number of cells significantly higher than one:	77
Mean:	1.016
Median:	1.014
Maximum:	1.142
Minimum:	0.866

Source: U.S. Census Bureau, Current Population Survey, Food Security, December 2022
[+] Some cells were empty due to the high number of cells, therefore this number
does not match up to the number of possible cells listed above.

As shown in Table 3.A., the distributions of the cells that had ratios of *test_NIWGT0* to *orig_NIWGT0* that were significantly higher than one was significantly different than the unweighted distributions for the three demographic variables in Group 1 (ethnicity, age group, sex). Also, the distributions of the cells with ratios of *test_NIWGT0* to *orig_NIWGT0* in the top 10th percentile were significantly differently than the unweighted distributions for ethnicity, age, and sex. For the age variable, this shows that the test noninterview adjustment shifts cells from the older age groups to the 18 – 30-year-old age group. Also, many more cells are falling in the Hispanic category and male category due to the new noninterview adjustment, compared to the unweighted distribution of records.

Looking at the other side, the cells in the bottom 10th percentile had distributions that were not significantly different from the unweighted variable distributions for age and sex. However, there was a significant difference in distributions for ethnicity.

Table 4.A. Distributions of Demographic Characteristics of State/Race/Age/Sex Cells, Unweighted and With Certain Ratios of *test_NIWGT0* to *orig_NIWGT0*: December 2022

Characteristic		Total un-weighted	Cells with ratios significantly higher than 1		Cells with high ratios (above 90th percentile)		Cells with low ratios (below 10th percentile)	
		Percent	Percent	P-value	Percent	P-value	Percent	P-value
Race (*PRWTRACE*)	1 = White Alone	79.64	10.20	< 0.01*	0.00	< 0.01*	22.15	< 0.01*
	2 = Black Alone	10.36	45.71		41.61		18.79	
	3 = Asian Alone	5.71	37.96		46.31		18.79	
	4 = Other	4.28	6.12		12.08		40.27	
Age (*PEAGE* grouped)	1 = <18 yrs	20.86	26.12	< 0.01*	28.19	< 0.01*	18.79	0.01*
	2 = 18-30 yrs	14.72	32.65		34.90		23.49	
	3 = 31-64 yrs	43.22	22.86		12.75		32.89	
	4 = 65+ yrs	21.20	18.37		24.16		24.83	
Sex (*PESEX*)	1 = male	48.68	56.33	0.02*	58.39	0.02*	40.94	0.06*
	2 = female	51.32	43.67		41.61		59.06	

Source: U.S. Census Bureau, Current Population Survey, Food Security, December 2022
* Indicates distribution is significantly different from the total unweighted distribution at the 0.1 significance level.
Note: Within a category, percents may not sum to 100 due to rounding.

Table 4.B. Summary Statistics for Ratios of *test_NIWGT0* to *orig_NIWGT0* for State/Race/Age/Sex Cells: December 2022

Total cells:	1,499[+]
Number of cells in 10% tails:	149
Number of cells significantly higher than one:	245
Mean:	1.028
Median:	1.021
Maximum:	1.255
Minimum:	0.802

[+] Some cells were empty due to the high number of cells, therefore this number does not match up to the number of possible cells listed above.
Source: U.S. Census Bureau, Current Population Survey, Food Security, December 2022

Looking at Table 4.A., the distributions of the cells that had ratios of *test_NIWGT0* to *orig_NIWGT0* that were significantly higher than one was significantly different than the unweighted distributions for the three demographic variables in Group 2 (race, age, sex). Also, the distributions of the cells with ratios of *test_NIWGT0* to *orig_NIWGT0* in the top 10th percentile were significantly different than the unweighted distributions for race, age, and sex. For the race variable, this shows that the test noninterview adjustment shifts cells from the White Alone race group largely to the Black Alone and Asian Alone race groups. Also, within the age groups, 31 – 64-year-old age group cells shift to the 18 – 30-year-old age group and more cells are falling in the male category due to the new noninterview adjustment, compared to the unweighted distribution of records.

Looking at the other side, the cells in the bottom 10th percentile had distributions that were also significantly different from the unweighted variable distributions for the three demographic variables in Group 2: race, age, and sex. In particular, the race variable cells shifted from the White Alone race group largely to the Other race group.

Table 5.A. Distributions of Demographic Characteristics of Region/Ethnicity/Age/Sex Cells, Unweighted and With Certain Ratios of *test_NIWGT0* to *orig_NIWGT0*: December 2022

Characteristic		Total un-weighted	Cells with ratios significantly higher than 1		Cells with high ratios (above 75th percentile)		Cells with low ratios (below 25th percentile)	
		Percent	Percent	P-value	Percent	P-value	Percent	P-value
Ethnicity (*PEHSPNON*)	1 = Hisp	15.64	60.00	< 0.01*	87.50	< 0.01*	12.50	0.73
	2 = Non-Hisp	84.36	40.00		12.50		87.50	
Age (*PEAGE* grouped)	1 = <18	20.86	20.00	< 0.01*	25.00	0.02*	18.75	0.66
	2 = 18-30	14.72	48.00		37.50		6.25	
	3 = 31-64	43.22	16.00		12.50		43.75	
	4 = 65+	21.20	16.00		25.00		31.25	
Sex (*PESEX*)	1 = male	48.68	68.00	0.05*	68.75	0.11	37.50	0.37
	2 = female	51.32	32.00		31.25		62.50	

Source: U.S. Census Bureau, Current Population Survey, Food Security, December 2022
* Indicates distribution is significantly different from the total unweighted distribution at the 0.1 significance level.
Note: Within a category, percents may not sum to 100 due to rounding.

Table 5.B. Summary Statistics for Ratios of *test_NIWGT0* to *orig_NIWGT0* for Region/Ethnicity/Age/Sex Cells: December 2022

Total cells:	64
Number of cells in 25% tails:	16
Number of cells significantly higher than one:	25
Mean:	1.015
Median:	1.016
Maximum:	1.044
Minimum:	0.990

Source: U.S. Census Bureau, Current Population Survey, Food Security, December 2022

For Group 3 cells in Table 5.A. (region/ethnicity/age/sex), the distributions of the cells that had ratios of *test_NIWGT0* to *orig_NIWGT0* that were significantly higher than one was significantly different than the unweighted distributions for the three demographic variables. The distributions of the cells with ratios of *test_NIWGT0* to *orig_NIWGT0* in the top quarter were significantly different than the unweighted distributions for ethnicity and age, but not significantly different for sex. For the age variable, this shows that the test noninterview adjustment shifts cells from the 31 – 64-year-old age group to the 18 – 30-year-old age group. Also, many more cells are falling in the Hispanic category due to the new noninterview adjustment, compared to the unweighted distribution of records.

Looking at the other side, the cells in the bottom quarter had distributions that were not significantly different from the unweighted variable distributions for all three demographic categories (ethnicity, age, and sex).

Table 6.A. Distributions of Demographic Characteristics of Region/Race/Age/Sex Cells, Unweighted and With Certain Ratios of *test_NIWGT0* to *orig_NIWGT0*: December 2022

Characteristic		Total un-weighted	Cells with ratios significantly higher than 1		Cells with high ratios (above 75th percentile)		Cells with low ratios (below 25th percentile)	
		Percent	Percent	P-value	Percent	P-value	Percent	P-value
Race (*PRWTRACE*)	1 = White Alone	79.64	3.70		0.00		68.75	
	2 = Black Alone	10.36	46.30	< 0.01*	40.63	< 0.01*	0.00	< 0.01*
	3 = Asian Alone	5.71	48.15		59.38		0.00	
	4 = Other	4.28	1.85		0.00		31.25	
Age (*PEAGE* grouped)	1 = <18 yrs	20.86	24.07		31.25		18.75	
	2 = 18-30 yrs	14.72	31.48	< 0.01*	34.38	< 0.01*	12.50	0.34
	3 = 31-64 yrs	43.22	24.07		12.50		34.38	
	4 = 65+ yrs	21.20	20.37		21.88		34.38	
Sex (*PESEX*)	1 = male	48.68	55.56	0.31	59.38	0.23	37.50	0.21
	2 = female	51.32	44.44		40.63		62.50	

Source: U.S. Census Bureau, Current Population Survey, Food Security, December 2022
* Indicates distribution is significantly different from the total unweighted distribution at the 0.1 significance level.
Note: Within a category, percents may not sum to 100 due to rounding.

Table 6.B. Summary Statistics for Ratios of *test_NIWGT0* to *orig_NIWGT0* for Region/Race/Age/Sex Cells: December 2022

Total cells:	128
Number of cells in 25% tails:	32
Number of cells significantly higher than one:	54
Mean:	1.029
Median:	1.024
Maximum:	1.111
Minimum:	0.982

Source: U.S. Census Bureau, Current Population Survey, Food Security, December 2022

For the last group, Table 6.A. shows that the distributions of the cells that had ratios of *test_NIWGT0* to *orig_NIWGT0* that were significantly higher than one was significantly different than the unweighted distributions for two of the demographic variables, race and age; whereas, it was not significantly different from the unweighted distribution for sex. Also, the distributions of the cells with ratios of *test_NIWGT0* to *orig_NIWGT0* in the top quarter were significantly differently than the unweighted distributions for ethnicity and age, but not significantly different for sex. For the race variable, this shows that the test noninterview adjustment shifts cells from the White Alone race group to the Black Alone and Asian Alone race groups. Also, within the age groups, 31 – 64-year-old age group cells shift to the younger age groups due to the new noninterview adjustment, compared to the unweighted distribution of records.

Looking at the other side, the cells in the bottom quarter had distributions that were not significantly different from the unweighted variable distributions for age and sex. However, there was a significant difference in distributions for race, where cells shifted out of Black Alone and Asian Alone and into Other.

Evaluation Part 3: Nonadjustment Weight Ratios By Individual Record

Significance Tests of Ratios

When looking at all cases individually (without being grouped), the steps for testing the ratios for significance are as follows:

Step 1. Calculate the ratio of *test_NIWGTi* to *orig_NIWGTi* for each record.

$$\hat{\theta}_i = \frac{test\_NIWGTi}{orig\_NIWGTi}$$

where *i* = 0 is the full weight and *i* = 1 to 160 are the replicate weights.

Step 2. Calculate the standard error of the full weight ratio, $\hat{\theta}_0$, using the replicate ratios, $\hat{\theta}_i$, where *i* = 1 to 160.

$$Var(\hat{\theta}_0) = \frac{4}{160}\sum_{i=1}^{160}\left(\hat{\theta}_i - \hat{\theta}_0\right)^2$$

$$SE(\hat{\theta}_0) = \sqrt{Var(\hat{\theta}_0)}$$

Step 3. Perform a significance test comparing the full-weight ratio to 1.

$$Stat(\hat{\theta}_0) = \frac{\hat{\theta}_0 - 1}{SE(\hat{\theta}_0)}$$

If $Stat(\hat{\theta}_0) > 1.645$, then the ratio is significantly different from 1 at the 0.1 level: *test_NIWGT0* is adjusting the cells differently than *orig_NIWGT0*.

Results: Of the 75,000 cases, 11.46 percent of the cases (8,500) had ratios that were significantly different from 1 at the 0.1 significance level.

<u>Compare Distributions</u>

Next, various distributions are explored to see how different demographic characteristics were affected by the new weighting. The distribution of the full weight ratio, *test_NIWGT0 / orig_NIWGT0*, for all cases is displayed through summary statistics in Table 7. Several distributions are compared for different demographic variables (race, ethnicity, age, sex) in Tables 8 and 9.

Table 7. Summary statistics for ratios of *test_NIWGT0* to *orig_NIWGT0*

Total records:	75,000
Mean:	1.011
Median:	1.008
Standard Deviation:	0.08054
Minimum:	0.704
Maximum:	1.432

Source: U.S. Census Bureau, Current Population Survey, Food Security, December 2022

Table 8 compares the distributions of the ratios of *orig_NIWGT0* to *test_NIWGT0* that were significantly higher than one and that were in the top and bottom 5th percentiles. These were compared to the distributions of all unweighted cases for the demographic variables of ethnicity, race, age, and sex (designated by shading).

Table 8. Distribution of Demographic Characteristics Across all Records, Unweighted and with Certain Ratios of *test_NIWGT0* to *orig_NIWGT0*: December 2022

Characteristic		Total Un-weighted	Cases with ratios significantly higher than 1		Cases with high ratios (above 95th percentile)		Case with low ratios (below 5th percentile)	
		Percent	Percent	P-value	Percent	P-value	Percent	P-value
Ethnicity (*PEHSPNON*)	1 = Hisp	15.64	22.63	< 0.01*	25.81	< 0.01*	16.84	0.04*
	2 = Non-Hisp	84.36	77.37		74.19		83.16	
Race (*PRWTRACE*)	1 = White Alone	79.64	69.71	< 0.01*	59.01	< 0.01*	88.25	< 0.01*
	2 = Black Alone	10.36	15.40		23.18		7.10	
	3 = Asian Alone	5.71	12.01		15.05		2.09	
	4 = Other	4.28	2.89		2.76		2.56	
Age (grouped) (*PEAGE*)	1 = <18	20.86	19.53	< 0.01*	22.10	< 0.01*	15.42	< 0.01*
	2 = 18-30	14.72	19.41		20.93		12.99	
	3 = 31-64	43.22	42.42		39.64		47.11	
	4 = 65+	21.20	18.64		17.33		24.48	
Sex (*PESEX*)	1 = male	48.68	52.08	< 0.01*	52.84	< 0.01*	45.48	< 0.01*
	2 = female	51.32	47.92		47.16		54.52	

Source: U.S. Census Bureau, Current Population Survey, Food Security, December 2022
* Indicates distribution is significantly different from the total unweighted distribution at the 0.1 significance level.
Note: Within a category, percents may not sum to 100 due to rounding.

Table 8 shows that the distributions of the cases that had ratios of *test_NIWGT0* to *orig_NIWGT0* that were significantly higher than one was significantly different than the unweighted distributions for all four of the measured demographic variables, ethnicity, race, age, and sex. Also, the distributions of the cases with ratios of *test_NIWGT0* to *orig_NIWGT0* in the top 5th percentile were significantly different than the unweighted distributions for these four characteristics. For ethnicity, the test noninterview adjustment shifted cases from Non-Hispanic to Hispanic, and for race, White Alone to Black Alone and Asian Alone. Also, there was a greater share of 18 – 30-year-old cases, and a greater share of male cases due to the test noninterview adjustment, compared to the unweighted distribution.

Looking at the other side, the cases in the bottom 5th percentile had distributions that were also significantly different from the unweighted variable distributions for all demographic characteristics but shifted differently than the other two types of ratios, with less Black Alone and Asian Alone, less younger age groups, and less male cases.

The remaining distributions used for evaluation of the test noninterview weight are in Tables 9 and 10 and include all cases weighted by both the original and test noninterview weights, as well as the independent distribution from the December 2022 pop controls for comparison (designated by shading).

Table 9. Weighted and Independent Distributions of Demographic Characteristics Across all Records: December 2022

Characteristic		Dec 2022 Pop Control	Weighted orig_NIWGT0		Weighted test_NIWGT0		orig_NIWGT0 compared to test_NIWGT0
		Percent	Percent	P-value	Percent	P-value	P-value
Ethnicity (PEHSPNON)	1 = Hisp	19.22	18.24	< 0.01*	18.42	0.02*	0.15
	2 = Non-Hisp	80.78	81.76		81.58		
Race (PRWTRACE)	1 = White Alone	75.65	79.45	N	78.84	N	< 0.01**
	2 = Black Alone	13.44	10.88		11.21		
	3 = Asian Alone	10.91	6.20		6.46		
	4 = Other		3.47		3.49		
Age (grouped) (PEAGE)	1 = <18	22.04	20.71	< 0.01*	20.86	< 0.01*	< 0.01**
	2 = 18-30	16.79	14.96		15.13		
	3 = 31-64	43.65	43.56		43.35		
	4 = 65+	17.53	20.77		20.66		
Sex (PESEX)	1 = male	49.20	48.37	< 0.01*	48.53	< 0.01*	0.03**
	2 = female	50.80	51.63		51.47		

Source: U.S. Census Bureau, Current Population Survey, Food Security, December 2022

* Indicates distribution is significantly different from the Dec 2022 Pop Control distribution at the 0.1 significance level.

** Indicates the distributions weighted by the original noninterview weight (orig_NIWGT) are significantly different than the distributions weighted by the test noninterview weight (test_NIGT) at the 0.1 significance level.

N The comparison distribution of Dec 2022 pop controls is not comparable to the race noninterview weight (orig_NIWGT and test_NIWGT) distributions because of differing race definitions.

Note: Within a category, percents may not sum to 100 due to rounding.

Table 8 shows that the distributions of the cases that were weighted by the original noninterview adjustment weight, orig_NIWGT0, were significantly different than the independent distributions of December 2022 population controls for ethnicity, age, and sex. Also, the distributions of the cases that were weighted by the test noninterview adjustment weight, test_NIWGT, were significantly different than the December 2022 population controls for ethnicity, age, and sex. Note that this comparison was not possible for the race characteristic because the definitions did not match up.

When comparing the two weighted sets of distributions against each other, they were found significantly different for race, age, and sex, with the distributions of the test weights shifting towards the distribution of the pop controls. The two distributions were not significantly different for ethnicity.

Table 10. Weighted Percents of Demographic Characteristics Across all Records: December 2022

Characteristic		Weighted *orig_NIWGT0*	Weighted *test_NIWGT0*	*orig_NIWGT0* compared to *test_NIWGT0*
		Percent	Percent	P-value
Ethnicity (*PEHSPNON*)	1 = Hisp	18.24	18.42	< 0.01*
	2 = Non-Hisp	81.76	81.58	< 0.01*
Race (*PRWTRACE*)	1 = White Alone	79.45	78.84	< 0.01*
	2 = Black Alone	10.88	11.21	< 0.01*
	3 = Asian Alone	6.20	6.46	< 0.01*
	4 = Other	3.47	3.49	0.14
Age (grouped) (*PEAGE*)	1 = <18	20.71	20.86	< 0.01*
	2 = 18-30	14.96	15.13	< 0.01*
	3 = 31-64	43.56	43.35	< 0.01*
	4 = 65+	20.77	20.66	< 0.01*
Sex (*PESEX*)	1 = male	48.37	48.53	< 0.01*
	2 = female	51.63	51.47	< 0.01*

Source: U.S. Census Bureau, Current Population Survey, Food Security, December 2022
* Indicates the percent weighted by the original noninterview weight (*orig_NIWGT0*) is significantly different from the percent weighted by the test noninterview weight (*test_NIWGT0*) at the 0.1 significance level.
Note: Within a category, percents may not sum to 100 due to rounding.

The final evaluation table, Table 10, shows that the weighted percent of cases using the original noninterview adjustment weight, *orig_NIWGT0*, were significantly different than the weighted percent of cases using the test noninterview adjustment weight, *test_NIWGT0*, for all ethnicity, race, age, and sex sub-groups except for Other race.

Conclusion

In conclusion, comparing the distributions of the highest ratios (of test noninterview adjustment to original noninterview adjustment) to the unweighted distributions seemed to show a general shift in the weights from non-Hispanic to Hispanic, from White-alone to Black alone and Asian alone, from the older age groups to the younger age groups, and from female to male.[3] The two differently weighted distributions were compared at the record level to an independent population distribution, and significant differences were found for all characteristics compared (ethnicity, age, sex). However, these two weighted distributions were also significantly different from each other for all characteristics except for ethnicity. The new noninterview adjustment does seem to make a difference in the weighting for key demographic characteristics.

---

[3] In group 4 where the ratios are grouped into cells based on region, race, age, and sex, the distribution of the highest ratios were not significantly different from the total unweighted distribution for sex.

SAS Code for Evaluation of Test Noninterview Weights   **Appendix A**

Program 1: evaluation.sas

```
*data with orig niwgt;
data fst.orig_niwgt (keep = nonresp gereg gestfips pesex1 hefaminc1 peeduca1 prwtrace1 pehspnon1
prcivlf1 peage1 niwgt0-niwgt160 qstnum occurnum sswgt0);
        merge    dsd1.dec22repwgtout (keep = qstnum occurnum GESTFIPS PEAGE PESEX PEHSPNON
prwtrace niwgt0-niwgt160 sswgt0-sswgt160)
                        fs.fsdec22 (keep = qstnum occurnum gereg hefaminc peeduca prcivlf);
        by qstnum occurnum;
    if pesex = -1 then pesex1 = 0; else pesex1 = pesex;
        *no pesex = -1;
    if hefaminc in (-3, -2, -1) then hefaminc1 = 0;                              *no
hefaminc < 0;
                else if hefaminc in (1:8) then hefaminc1 = 1;
                *under $30k;
                else if hefaminc in (9:13) then hefaminc1 = 2;
                *$30k <= inc < $75;
                else if hefaminc in (14:15) then hefaminc1 = 3;
                *$75k <= inc < $150k;
                else if hefaminc in (16) then hefaminc1 = 4;
        *$150k and above;
    if pehspnon = -1 then pehspnon1 = 0; else pehspnon1 = pehspnon;           *no pehspnon = -
1;
    if peeduca = -1 then peeduca1 = 0;
                else if peeduca in (31:38) then peeduca1 = 1;
        *less than high school diploma;
                else if peeduca = 39 then peeduca1 = 2;
                *hs diploma/GED;
                else if peeduca = 40 then peeduca1 = 3;
                *some college;
                else if peeduca in (41:46) then peeduca1 = 4;
        *college degree;
    if prwtrace = -1 then prwtrace1 = 0; else prwtrace1 = prwtrace;           *no prwtrace = -
1;
    if peage < 18 then peage1 = 1;
                        *no peage < 0;
                else if 18 le peage le 30 then peage1 = 2;
                else if 31 le peage le 64 then peage1 = 3;
                else peage1 = 4;
    if prcivlf = -1 then prcivlf1 = 0; else prcivlf1 = prcivlf;
        if hrsupint = 1 then nonresp = 1; else nonresp = 0;
run;


*data with test niwgt;
data fst.test_niwgt;
 set fs.cont_fs_dec22 (rename = (peage1 = peage hefaminc1 = hefaminc peeduca1 = peeduca));
 keep qstnum occurnum nonresp gereg gestfips pesex1 hefaminc1 peeduca1 prwtrace1 pehspnon1
prcivlf1 peage1 niwgt0-niwgt160;
    if peage < 18 then peage1 = 1;
                else if 18 le peage le 30 then peage1 = 2;
                else if 31 le peage le 64 then peage1 = 3;
                else peage1 = 4;
    if peeduca = 0 then peeduca1 = 0;
                else if peeduca in (31:38) then peeduca1 = 1;
        *less than high school diploma;
                else if peeduca = 39 then peeduca1 = 2;
                *hs diploma/GED;
                else if peeduca = 40 then peeduca1 = 3;
                *some college;
                else if peeduca in (41:46) then peeduca1 = 4;
        *college degree;
        if hefaminc in (-3, -2, -1) then hefaminc1 = 0;
                else if hefaminc in (1:8) then hefaminc1 = 1;
                *under $30k;
                else if hefaminc in (9:13) then hefaminc1 = 2;
                *$30k <= inc < $75;
```

```
                           else if hefaminc in (14:15) then hefaminc1 = 3;
                           *$75k <= inc < $150k;
                           else if hefaminc in (16) then hefaminc1 = 4;
              *$150k and above;
run;


/*compare niwgts at record level*/

proc sort data = fst.orig_niwgt;
 by qstnum occurnum;
run;

*rename niwgt so it can be compared to niwgt in test dataset;
data one;
  set fst.orig_niwgt (keep = qstnum occurnum nonresp niwgt0-niwgt160 pehspnon1 prwtrace1 peage1
pesex1);
    where nonresp = 1;
          array niwgt_orig[161] niwgt0 - niwgt160;
          array niwgt_orig1[161] orig_niwgt0 - orig_niwgt160;
          do i = 1 to 161;
            niwgt_orig1[i] = niwgt_orig[i];
          end;
  keep qstnum occurnum orig_niwgt0 - orig_niwgt160 pehspnon1 prwtrace1 peage1 pesex1 nonresp;
run;

proc sort data = fst.test_niwgt;
 by qstnum occurnum;
run;

*calcuate ratio of test niwgt to original niwgt;
data fst.niwgt_rec;
  merge one (in = a) fst.test_niwgt (keep = qstnum occurnum niwgt0 - niwgt160);
    by qstnum occurnum;
    if a;
    array niwgt_orig[161] orig_niwgt0 - orig_niwgt160;
    array niwgt_test[161] niwgt0 - niwgt160;
    array ratio[161] ratio_niwgt0 - ratio_niwgt160;
          do i = 1 to 161;
          ratio[i] = niwgt_test[i] / niwgt_orig[i];
          end;
run;

*rsubmit;
*calculate the standard error and perform significance test for ratio different from 1;
data fst.niwgt_rec_t (keep = qstnum occurnum ratio_niwgt0 sterr stat sig pehspnon1 prwtrace1
peage1 pesex1);
 set fst.niwgt_rec;
    array ratio[0:160] ratio_niwgt0 - ratio_niwgt160;
    array repdifsq[1:160] repdifsq1 - repdifsq160;

          do i = 1 to 160;
          repdifsq[i] = (ratio[i] - ratio_niwgt0)**2;
          end;

          sterr = sqrt(.025 * sum(of repdifsq:));

          stat = (ratio_niwgt0 - 1) / sterr;                          *calculate test statistic
for ratio higher than 1;

    if stat > 1.645 then sig = "sig";                                 *significantly
different at .1 significance level?;
          else sig = "not sig";

run;

rsubmit;
*pre-Table 6;
*how many ratios are significantly from 1?;
proc freq data = fst.niwgt_rec_t;
 tables sig;
```

```sas
run;

*Table 6;
*look at niwgt and replicates summary statistics;
proc means data = fst.niwgt_rec n median mean min max std;
 var ratio_niwgt0;*/ - ratio_niwgt160;
run;


*what is the distribution of cases that are significantly higher than one?;
proc freq data = fst.niwgt_rec_t;
 where sig = "sig";
 tables pehspnon1 / chisq testp = (.1564 .8346);
run;

proc freq data = fst.niwgt_rec_t;
 where sig = "sig";
 tables prwtrace1 / chisq testp = (.7964 .1036 .0571 .0428);
run;

proc freq data = fst.niwgt_rec_t;
 where sig = "sig";
 tables peage1 / chisq testp = (.2086 .1472 .4322 .212);
run;

proc freq data = fst.niwgt_rec_t;
 where sig = "sig";
 tables pesex1 / chisq testp = (.4868 .5132);
run;


*find the 5% tails of the distribution of the ratio of NIWGT0;
proc univariate data = fst.niwgt_rec;
 var ratio_niwgt0;
 *histogram;      *doesn't work;
 output out = tails p5 = p5 p95 = p95;
run;

data fst.niwgt_rec;
 set fst.niwgt_rec;
  comb = 1;
run;

data tails;
 set tails;
  comb = 1;
run;

data fst.niwgt_rec (drop = comb);
 merge fst.niwgt_rec tails;
  by comb;
run;

data fst.extremes_high fst.extremes_low;
 merge fst.niwgt_rec (keep = qstnum occurnum ratio_niwgt0 p5 p95 in = a)
          fst.orig_niwgt (keep = qstnum occurnum gestfips gereg peage1 pesex1 prwtrace1
pehspnon1 prcivlf1 hefaminc1 peeduca1);
 by qstnum occurnum;
 if a;
 if ratio_niwgt0 >= p95 then output fst.extremes_high;
 if ratio_niwgt0 <= p5 then output fst.extremes_low;
run;

endrsubmit;


*find distribution of all cases by demo/geo variables;

proc freq data = fst.orig_niwgt;
 title 'Distribution of All Cases - Response and Nonresponse';
```

```
 tables nonresp pehspnon1 prwtrace1 peage1 pesex1 gereg;*gestfips;
run;

%macro dist(in, var);
proc freq data = &in;
 title "unweighted distibution of &var. - no blanks";
 tables &var;
 where &var ne 0;
run;
%mend;

%dist(fst.orig_niwgt, prcivlf1);
%dist(fst.orig_niwgt, hefaminc1);
%dist(fst.orig_niwgt, peeduca1);


*find distribution of demo/geo variables with high and low ratios of test to orig NIWGT;

proc freq data = fst.extremes_high;
 title 'Distribution of Cases with High Ratio of Test to Orig NIWGT - Respondents Only';
 tables pehspnon1 prwtrace1 peage1 pesex1 gereg;*gestfips;
run;

%dist(fst.extremes_high, prcivlf1);
%dist(fst.extremes_high, hefaminc1);
%dist(fst.extremes_high, peeduca1);


proc freq data = fst.extremes_low;
 title 'Distribution of Cases with Low Ratio of Test to Orig NIWGT - Respondents Only';
 tables pehspnon1 prwtrace1 peage1 pesex1 gereg; *gestfips;
run;

%dist(fst.extremes_low, prcivlf1);
%dist(fst.extremes_low, hefaminc1);
%dist(fst.extremes_low, peeduca1);



*find distribution of demo/geo variables weighted by each NIWGT;

proc freq data = fst.orig_niwgt;
 title 'Distribution of Respondent Cases Weighted by Original NIWGT';
 tables pehspnon1 prwtrace1 peage1 pesex1 gereg;* gestfips;
 weight niwgt0;
 *where nonresp = 1;
run;

%macro dist_wgt(in, var, weight);
proc freq data = &in;
 title "weighted distibution of &var. from &in (no blanks)";
 tables &var;
 where &var ne 0;
 weight &weight;
run;
%mend;

%dist_wgt(fst.orig_niwgt, prcivlf1, niwgt0);
%dist_wgt(fst.orig_niwgt, hefaminc1, niwgt0);
%dist_wgt(fst.orig_niwgt, peeduca1, niwgt0);

proc freq data = fst.test_niwgt;
 title 'Distribution of Respondent Cases Weighted by Test NIWGT';
 tables pehspnon1 prwtrace1 peage1 pesex1 prcivlf1 hefaminc1 peeduca1 gereg;* gestfips;
 weight niwgt0;
run;

%dist_wgt(fst.test_niwgt, prcivlf1, niwgt0);
%dist_wgt(fst.test_niwgt, hefaminc1, niwgt0);
%dist_wgt(fst.test_niwgt, peeduca1, niwgt0);
```

SAS Code for Evaluation of Test Noninterview Weights   **Appendix A**

```sas
title;
```
Program 2: eval_macro.sas

```sas
rsubmit;


%macro comp_niwgt(vars, n, run, high, low);

proc sort data = fst.orig_niwgt out = orig_niwgt_srt;
 by &vars;
run;

proc sort data = fst.test_niwgt out = test_niwgt_srt;
 by &vars;
run;


%do i = 0 %to &n;

*original niwgts - sum of all niwgts for each replicate by demo group;
proc means data = orig_niwgt_srt sum noprint;
  by &vars;
  where nonresp = 1;
  var niwgt&i;
  output out = orig_niwgt_test&run.&i (drop = _type_ rename = (_freq_ = orig_count&run)) sum =
orig_sum_niwgt&run.&i;
run;

*test niwgts - sum of all niwgts for each replicate by demo group;
proc means data = test_niwgt_srt sum noprint;
  by &vars;
  where nonresp = 1;
  var niwgt&i;
  output out = test_niwgt_test&run.&i (drop = _type_ rename = (_freq_ = test_count&run)) sum =
test_sum_niwgt&run.&i;
run;

%end;

*Calculate ratio of test niwgt to orig niwgt, standard error of ratio using replicates,
        and test statistic against the constant 1;

data comp&run (keep = &vars ratio_niwgt&run.0 - ratio_niwgt&run.160 sterr&run stat&run sig&run
diff&run);
 merge orig_niwgt_test&run.0 - orig_niwgt_test&run.160 test_niwgt_test&run.0 -
test_niwgt_test&run.160;
  by &vars;

  array niwgt_orig[0:160] orig_sum_niwgt&run.0 - orig_sum_niwgt&run.160;
  array niwgt_test[0:160] test_sum_niwgt&run.0 - test_sum_niwgt&run.160;
  array ratio[0:160] ratio_niwgt&run.0 - ratio_niwgt&run.160;
  array repdifsq[1:160] repdifsq&run.1 - repdifsq&run.160;

        do i = 0 to 160;
        ratio[i] = niwgt_test[i] / niwgt_orig[i];
        end;

        do i = 1 to 160;
        repdifsq[i] = (ratio[i] - ratio_niwgt&run.0)**2;
        end;

        sterr&run. = sqrt(.025 * sum(of repdifsq&run.:));

        stat&run. = (ratio_niwgt&run.0 - 1) / sterr&run.;

    if stat&run. > 1.645 then sig&run = "sig";
        *significantly different at .1 significance level?;
        else sig&run = "not sig";

        diff&run = test_count&run - orig_count&run;
```

```
run;


*how many ratios are significantly different from 1?;
proc freq data = comp&run;
 tables sig&run;
run;

*distribution of ratios that are significantly different from 1;
/*
proc freq data = comp&run;
        where sig&run = "sig";
        tables &vars;
run;
*/

*find high and low ratios based on percentiles in the distribution;
proc univariate data = comp&run noprint;
 var ratio_niwgt&run.0;
 *histogram;
 output out = tails p&low = p&low._&run p&high = p&high._&run;
run;

data comp&run;
 set comp&run;
  comb = 1;
run;

data tails;
 set tails;
  comb = 1;
run;

data comp&run (drop = comb);
 merge comp&run tails;
  by comb;
run;

/*
proc print data = comp&run;
 where ratio_niwgt&run.0 > p&high._&run or ratio_niwgt&run.0 < p&low._&run;
 var &vars ratio_niwgt&run.0 diff&run sterr&run stat&run sig&run p&high._&run p&low._&run;
run;
*/

proc means data = comp&run n sum mean median min max;
 title 'ratio of test to orig niwgts for &vars';
 *by &vars;
 var diff&run ratio_niwgt&run.0;
run;
/*
proc freq data = comp&run;
 where ratio_niwgt&run.0 > p&high._&run;
 tables &vars;
run;

proc freq data = comp&run;
 where ratio_niwgt&run.0 < p&low._&run;
 tables &vars;
run;
*/
%mend;


*rsubmit;
%macro comp_dist(run, crit, var, p);

*testing distribution of significant ratios vs unweighted;
proc freq data = comp&run;
 where &crit;
 tables &var / chisq testp = (&p);            *test against unweighted proportions;
```

```
run;

proc freq data = comp&run;
        where &crit;
        tables peage1 / chisq testp = (.2086 .1472 .4322 .212);
run;

proc freq data = comp&run;
        where &crit;
        tables pesex1 / chisq testp = (.4868 .5132);
run;
%mend;


*rsubmit;
%comp_niwgt(vars = gestfips pehspnon1 peage1 pesex1, n = 160, run = a, high = 90, low = 10);
%comp_dist(a, siga eq "sig", pehspnon1, 15.64 83.46);
%comp_dist(a, ratio_niwgta0 > p90_a, pehspnon1, 15.64 83.46);
%comp_dist(a, ratio_niwgta0 < p10_a, pehspnon1, 15.64 83.46);

%comp_niwgt(gestfips prwtrace1 peage1 pesex1, 160, b, 90, 10);

%comp_dist(b, sigb eq "sig", prwtrace1, 79.64 10.36 5.71 4.28);

%comp_dist(b, ratio_niwgtb0 > p90_b, prwtrace1, 79.64 10.36 5.71 4.28);          *chisq
test in excel bc missing level;
%comp_dist(b, ratio_niwgtb0 < p10_b, prwtrace1, 79.64 10.36 5.71 4.28);

%comp_niwgt(gereg pehspnon1 peage1 pesex1, 160, c, 75, 25);
%comp_dist(c, sigc eq "sig", pehspnon1, 15.64 83.46);
%comp_dist(c, ratio_niwgtc0 > p75_c, pehspnon1, 15.64 83.46);
%comp_dist(c, ratio_niwgtc0 < p25_c, pehspnon1, 15.64 83.46);

%comp_niwgt(gereg prwtrace1 peage1 pesex1, 160, d, 75, 25);
%comp_dist(d, sigd eq "sig", prwtrace1, 79.64 10.36 5.71 4.28);
                ;
%comp_dist(d, ratio_niwgtd0 > p75_d, prwtrace1, 79.64 10.36 5.71 4.28);          *chisq
test in excel bc missing level;
%comp_dist(d, ratio_niwgtd0 < p25_d, prwtrace1, 79.64 10.36 5.71 4.28);          *chisq
test in excel bc missing level;




*record level distribution tests against unweighted distribution;
%macro comp_dist_rec(data, crit);

*testing distribution of significant ratios vs unweighted;
proc freq data = &data;
 &crit;
 tables pehspnon1 / chisq testp = (.1564 .8346);          *test against unweighted
proportions;
run;

proc freq data = &data;
 &crit;
 tables prwtrace1 / chisq testp = (.7964 .1036 .0571 .0428);          *test against unweighted
proportions;
run;

proc freq data = &data;
        &crit;
        tables peage1 / chisq testp = (.2086 .1472 .4322 .212);
run;

proc freq data = &data;
        &crit;
        tables pesex1 / chisq testp = (.4868 .5132);
```

```
run;
%mend;


*rsubmit;
title 'record level chisq dist test significant ratio cases against unweighted dist';
%comp_dist_rec(fst.niwgt_rec_t, where sig eq "sig");
title 'record level chisq dist test top 10% ratio cases against unweighted dist';
%comp_dist_rec(fst.extremes_high, );
title 'record level chisq dist test bottom 10% ratio cases against unweighted dist';
%comp_dist_rec(fst.extremes_low, );
title;


title 'compare distributions of ratios in top and bottom 5th and significant higher than one';

rsubmit;

*compare the three distributions together in Table 7;
data three;
 set fst.niwgt_rec_t (keep = sig pehspnon1 prwtrace1 peage1 pesex1 in = sig1)
         fst.extremes_high (keep = pehspnon1 prwtrace1 peage1 pesex1 in = high)
         fst.extremes_low (keep = pehspnon1 prwtrace1 peage1 pesex1 in = low);
         if sig1 then do;
                 if sig = "sig" then source = 'sig';
                 else delete;
         end;
         else if high then source = 'high';
         else if low then source = 'low';
         else delete;
run;

proc freq data = three;
        tables prwtrace1 * source / chisq;
run;

proc freq data = three;
        tables pehspnon1 * source / chisq;
run;

proc freq data = three;
        tables peage1 * source / chisq;
run;

proc freq data = three;
        tables pesex1 * source / chisq;
run;


title 'orig and test weights against dec22 pop controls using replicate weights';

rsubmit;

*compare distributions to controls using replicates for original and test NIWGT;

%macro comp_dist_rec_pop(data);
proc surveyfreq data = &data varmethod=brr (fay=0.5);
weight niwgt0;
repweights niwgt1-niwgt160;
        tables pehspnon1 /chisq testp = (19.223179 80.776821);
        title 'hisp orig niwgt against pop control Dec 22 with replicate weights';
run;

proc surveyfreq data = &data varmethod=brr (fay=0.5);
weight niwgt0;
repweights niwgt1-niwgt160;
        tables peage1 /chisq testp = (22.036261 16.791284 43.646813 17.525642);
        title ' niwgt against pop control Dec 22 with replicate weights';
run;
```

```
proc surveyfreq data = &data varmethod=brr (fay=0.5);
weight niwgt0;
repweights niwgt1-niwgt160;
        tables pesex1 /chisq testp = (49.195589 50.804411);
        title 'hisp orig niwgt against pop control Dec 22 with replicate weights';
run;

%mend;

%comp_dist_rec_pop(fst.orig_niwgt);
%comp_dist_rec_pop(fst.test_niwgt);

endrsubmit;

title 'compare orig v test weighted distributions with replicate weights';

rsubmit;

*compare weighted distributions using both weights;
data both;
 set fst.orig_niwgt (keep = niwgt0-niwgt160 pehspnon1 prwtrace1 peage1 pesex1 in = orig)
         fst.test_niwgt (keep = niwgt0-niwgt160 pehspnon1 prwtrace1 peage1 pesex1 in = test);
         if orig then source = 'orig_niwgt';
         else if test then source = 'test_niwgt';
run;

proc surveyfreq data = both varmethod=brr (fay=0.5);
weight niwgt0;
repweights niwgt1-niwgt160;
        tables prwtrace1 * source / chisq;
        title 'race orig niwgt v test niwgt with replicate weights';
run;

proc surveyfreq data = both varmethod=brr (fay=0.5);
weight niwgt0;
repweights niwgt1-niwgt160;
        tables pehspnon1 * source / chisq;
        title 'hisp orig niwgt v test niwgt with replicate weights';
run;

proc surveyfreq data = both varmethod=brr (fay=0.5);
weight niwgt0;
repweights niwgt1-niwgt160;
        tables peage1 * source / chisq;
        title 'age orig niwgt v test niwgt with replicate weights';
run;

proc surveyfreq data = both varmethod=brr (fay=0.5);
weight niwgt0;
repweights niwgt1-niwgt160;
        tables pesex1 * source / chisq;
        title 'sex orig niwgt v test niwgt with replicate weights';
run;
```

SAS Code for Evaluation of Test Noninterview Weights **Appendix A**

Program 3: dist_test_all.sas

```sas
title 'ORIGINAL NIWGT: calculate standard errors of percentages using replicate weights';

rsubmit;

proc surveyfreq data = fst.orig_niwgt varmethod=brr (fay=0.5);
        where nonresp = 1;
        weight niwgt0;
        repweights niwgt1-niwgt160;
    tables pehspnon1 prwtrace1 peage1 pesex1;
run;

endrsubmit;

title 'TEST NIWGT: calculate standard errors of percentages using replicate weights';

rsubmit;

proc surveyfreq data = fst.test_niwgt varmethod=brr (fay=0.5);
        where nonresp = 1;
        weight niwgt0;
        repweights niwgt1-niwgt160;
    tables pehspnon1 prwtrace1 peage1 pesex1;
run;
```