

# Integrating Person and Housing Unit Weighting for the Current Population Survey

Andrew A. Zbikowski, U.S. Census Bureau; Phawn M. Letourneau, U.S. Census Bureau

## 1. Introduction\*

The Current Population Survey (CPS) is the primary source of labor force data for the United States. It is sponsored by the Bureau of Labor Statistics and the Census Bureau. Although the main purpose of the CPS is to produce estimates of employment status and other personal characteristics for the civilian noninstitutional population, it also produces estimates for householders and housing units. The estimates of households and householders should agree by definition, in that there is one householder for every household.<sup>1</sup> But there are persistent discrepancies in these estimates within the CPS. In addition to this difference within the CPS, there are differences in estimates of households between the CPS and other demographic surveys conducted by the Census Bureau. This paper discusses some preliminary research on weighting methods for the CPS to resolve these problems.

## 2. An Overview of The CPS

The CPS is conducted monthly on a sample of approximately 70,000 households, resulting in approximately 55,000 interviewed households. The United States monthly unemployment rate is probably the best-known statistic from the CPS.

CPS sample households are divided into eight groups of approximately equal size, called rotation groups, because they rotate in and out of the sample. Six of the eight groups are the same in any pair of consecutive months, and four of the eight groups are the same from year to year (i.e., identical months 1 year apart). The rotation of the groups is commonly referred to as the 4-8-4 pattern, in that all households in a given rotation group are in the CPS sample for four consecutive months, out of the sample for the next eight months, and back in sample for the following four months.

## 3. Nature of the CPS Weighting

---

\* This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed are those of the authors and not necessarily those of the Census Bureau.

<sup>1</sup> A household includes all individuals residing in a sample housing unit. The householder is the person (or one of the people) who owns or rents the unit.

The first two CPS weighting adjustments are based on households or occupied housing units, and are applied to all members of a household. These are the weighting control factor and the noninterview factor. The base weight is the initial weight or estimate for each CPS person. It is simply the inverse of the probability of selection for all persons in the civilian noninstitutional population in a geographic area, usually a state. The weighting control factor is applied when it is necessary to subsample in the field, when there are many more housing units than expected. In most cases, where no subsampling occurs, the factor has a value of one. The noninterview factor accounts for nonresponse, when an eligible household does not complete an interview.

The remaining steps of the CPS weighting are applied to persons within each household. These are the first-stage factor, the national and state coverage adjustment factors, and the second-stage factor. The first-stage factor is intended to reduce the variance that results from sampling of primary sampling units (i.e., the first stage of CPS sample selection). The national and state coverage adjustments and the second-stage adjustment apply independent population controls to adjust the CPS sample estimates. The result is that every person who is a member of an eligible, interviewed sample CPS household receives a weight for each step of the CPS weighting and estimation. For the discussion in this paper, the 'second-stage' weight can be considered the final person weight. See Tupek (2004) for more detail on the CPS weighting.

The CPS household and family weights are derived from the second-stage weight of a designated person in the household, and it is here that the discrepancy in the estimates of households and householders arises. In married-couple households, the second-stage weight of the wife is typically used as the household weight, even when the husband is designated as the householder. This is because the CPS coverage ratios<sup>2</sup> for women are usually higher and less variable, from month to month, than for men.

## 4. Background for Weighting Research

The discrepancy in estimates of the households and householders has occurred in other surveys conducted

---

<sup>2</sup> The coverage ratio measures how well the survey covers the target population. It is the ratio of estimates after the first-stage adjustment to independent population controls.

by the Census Bureau. A housing unit control working group was formed at the Census Bureau to examine this problem and make recommendations. This group recommended that all current surveys that produce housing-based estimates<sup>3</sup> use the same set of independent housing estimates, provided by the Census Bureau's Population Division, as controls. But for surveys that produce both population-based and housing-based estimates, like the CPS, the working group recommended research on alternative weighting methods, and noted that housing unit estimates could be affected considerably, depending on how the weighting methods were changed. In particular, the CPS currently uses "population-based" estimates of housing units or households, as indicated in section 3 above, and controlling these estimates to independent housing unit controls could result in large changes. See Love (2002) and Love (2003) for more details on the discussion and recommendations of this housing unit control working group.

So *within* the CPS, there are differences in estimates of households and householders, but there are also differences between the CPS estimates and those from other Census Bureau surveys. The case of the Housing Vacancy Survey (HVS) is an illustration of this difference. The HVS is administered for sample housing units that are currently vacant, or occupied by people with a usual residence elsewhere. The HVS is the basis for estimates of rental and owner-occupied housing occupancy and vacancy rates, which are analyzed and presented in quarterly and annual reports.

The HVS is actually a supplement of the CPS, and prior to 2003, it used CPS household estimates to adjust its estimates of occupied and vacant housing units. Since 2003, the HVS has controlled these estimates to the independent housing estimates provided by the Census Bureau's Population Division. Thus, the HVS estimates are now consistent with other Census Bureau surveys that produce housing-based estimates, but the difference in estimated households between the HVS and the CPS is accentuated, with the CPS estimates for the United States larger by five to six million.

### 5. Preliminary Analysis of Current Weights

Prior to developing or examining any alternative procedures for household weighting in the CPS, we wanted to examine current estimates of households and householders in the CPS; aside from knowing that the total national estimates are different, we wanted to

---

<sup>3</sup> These include the American Housing Survey, the New York City Housing and Vacancy Survey, the Housing Vacancy Survey and the American Community Survey.

examine these differences by social and demographic characteristics. Were the inconsistencies between households and householders greater for householders of a certain age or race/ethnicity group? Or did household type and composition (e.g., presence of children, married couple versus single person households) or household size make a difference?

We examined data from the July, August and September 2005 CPS. We chose these months for a couple of reasons. July 2005 was the first month consisting completely of sample households from the sample design based on the 2000 Census<sup>4</sup>. We also have data from the HVS and from the independent housing unit estimates for this quarter, for additional comparisons.

Table 1 presents the estimated numbers of households and householders, and the household/householder ratios, by race/ethnicity of the householder<sup>5</sup>. Note that all of the ratios are close to one, and range from approximately .98 to 1.03. It is notable that the White NonHispanic and Other groups have the highest household/householder ratios, and that these groups also have the highest CPS coverage ratios. The overall ratio is .994, but this still represents about 674,000 more householders than households.

Table 2 presents the household/householder ratios by sex and age of householder. Most notably, the ratios for all female age groups round to 1, 1.001 or 1.002. This is consistent with the fact that, in the current determination of household weights, the second-stage weight for a civilian adult female is typically used for the household weight. This would obviously be true for single-person households, but it is also true for married-couple households; the wife's final weight is most often the assigned household weight, even when the husband is the householder. Exceptions to this rule would include a married-couple household where the wife is in the armed forces and the husband is a civilian. In this case, the husband's final weight would be used for the household weight. Among male householders, the ratios were highest for those under 25.

Finally, Table 3 looks at the data by type of household. We classified CPS sample households into seven groups:

---

<sup>4</sup> The transition from the 1990 Census sample to the 2000 Census sample occurred over a 15-month period, from April 2004 to June 2005.

<sup>5</sup> All comparisons presented in this paper were tested for statistical significance at the 90-percent confidence interval. When we report significant differences in this paper, they are at the 90-percent confidence interval.

1. married couples (MC) with at least one child 6 years old or younger;
2. married couples with at least one child under 18, but none 6 or younger;
3. other households with at least one child 6 years old or younger;
4. other households with at least one child under 18, but none 6 or younger;
5. married couples without children;
6. single adults 18 and over;
7. two or more adults 18 and over.

The households with one adult have, by definition, the same number of households and householders and a household/householder ratio of 1. Married- couple households with at least one child under 18 had the lowest ratios, around .982; married- couple households with no children under 18 had a household/householder ratio of .993. All other groups were at or slightly above 1.

Among the household and householder characteristics we've examined for levels of discrepancies in household and householder estimates, we found no statistical significance. Neither the race/ethnicity of the householder, nor the sex and age of the householder were associated with significant differences in the household/householder ratios. Similarly, the number of persons in the household and the household type (e.g., households with children, adults with no children) were not associated with significant differences in the household/householder ratio. We did note nonsignificant differences by the race/ethnicity of the householder, and among males, the age of the householder. As noted earlier, these findings are consistent with what we know about coverage ratios for the CPS, in that White NonHispanic and Other race/ethnicity groups, and female and older age groups have higher coverage ratios.

#### **6. Exploring Alternatives: the Noninterview Weight and the Regional Control Weight**

As described earlier, the CPS estimates of households are "person-based" – that is, they are derived from the final weight of a person in the household. One of the areas for research recommended by the Housing Unit Control Working Group was to consider basing the household weight for CPS (and other surveys) on a weight that is more closely related to the household. Remember that in the CPS, the weighting and estimation steps through the noninterview stage are housing unit or household based. The noninterview weight is one of the easier alternatives to examine for the household weight.

The noninterview weight accounts for eligible sample CPS households<sup>6</sup> that are not interviewed. Eligible households for which we fail to obtain an interview are classified as type A noninterviews. Reasons for a type A noninterview include respondent refusal, no one at home, and temporary absence. The interviewers for CPS strive to minimize the number of such noninterviews. In the CPS in July-September 2005, type A noninterviews comprised about 7.8 percent of the eligible sample households. The type A noninterview factor, NI, is the following ratio:

$$NI = \frac{Z + A}{Z}$$

where Z is the weighted total of interviewed households through the weighting control factor weight, and A is the weighted total of type A noninterviewed households through the weighting control factor weight.

To compute the noninterview adjustment, the interviewed and type A noninterviewed households are grouped by state, metropolitan statistical area (MSA) classification (yes/no) and household location (central city- urban/not central city-rural). So for each state, there are four separate cells or groups for which the noninterview adjustment is computed. For the July 2005 CPS, the median type A factor was 1.075, and the highest was 1.239.

Another alternative to consider is the use of the independent housing unit controls produced by the Census Bureau's Population Division. The Population Division produces housing unit estimates as of July 1<sup>st</sup> of each year, using state and county data on building permits and mobile homes, and estimates of housing unit loss. The division produces other monthly estimates by linear interpolation. See U.S. Census Bureau (2004) for more details on how the Census Bureau obtains estimates of housing units.

We noted earlier in this paper that the Bureau's housing-based surveys are using these estimates as controls. How would estimates of CPS household characteristics be affected, if they were controlled to these independent housing unit estimates? For our comparisons, we obtained independent housing unit estimates, (from the Bureau's Population Division) as of July 1, 2005, aggregated to the geographic regional level. We applied HVS 3<sup>rd</sup> quarter occupancy rates to these estimates to obtain estimates of occupied housing units.

---

<sup>6</sup> Of approximately 70,000 CPS sample households each month, about 60,000 are determined to be eligible for an interview.

The use of regional controls is consistent with the weighting process for the HVS. We obtained the regional control weight by computing the following factor,  $HUC_R$ , and applying it to the CPS household weight:

$$HUC_R \equiv \frac{\text{Independent HU estimate : region R}}{\text{Aggregated CPS weights : region R}}$$

where R is one of the four geographic regions into which CPS classifies every sample household: 1. Northeast; 2. Midwest; 3. South; and 4. West.

We examined the noninterview weight and regional control weight as alternatives for the household weight in two ways: 1. comparing the total estimates of households or occupied housing units resulting from each method for the period 2003-2005, and 2. comparing estimates of selected household/householder characteristics among the three methods using CPS data for July-September 2005.

Figure 1 compares the estimates of households or occupied housing units for the current CPS household weight with the CPS noninterview weight and the regional control weight, from January 2003 to October 2005. We noted in section 4 of this paper that the CPS estimates of households are considerably higher than the independent estimates of housing units from the Bureau's Population Division, and that fact is illustrated in Figure 1. For July 2005, the CPS estimates of households is 114.355 million, compared to 108.651 million occupied housing units based on the regional control weights. The estimate based on noninterview weights is lower still, 105.263 million, and each of these estimates is significantly different from the others. But we note that the independent estimate is closer to the noninterview weight estimate than to the current household weight estimate.

Table 4 compares the estimates of households by the householder race/ethnicity, using the current CPS household weight, with the CPS noninterview weight and the regional control weight. For all three estimation methods, slightly more than 70 percent of householders were White NonHispanic, but there was enough of a difference in the distributions for statistical significance. There is a significantly higher proportion of White NonHispanic householders based on the noninterview weight than for the other two methods. There are also significantly lower proportions of Black NonHispanics and Asian NonHispanics based on the noninterview weight. These results are consistent with the fact that the current household weights have been controlled to race/ethnicity population estimates in

several steps, while the noninterview estimates come from the weighting stage prior to those controls.

It is interesting to look at the comparisons among the three weighting methods for other characteristics. Tables 5 and 6 compare the weighting/estimation methods for the number of household members and for household type, respectively. Table 5 shows that there were no statistically significant differences in the distributions of households, based on household size. But Table 6 shows that there is a significantly higher proportion of married-couple households with no children based on the CPS noninterview weight, compared to the other methods

Table 7 compares household estimates based on the labor force status of the householder. We emphasize that the CPS is primarily intended to measure the labor force status of the population; for example, any proposal to include supplemental questions on a monthly administration of the CPS is evaluated to insure that the extra questions will not adversely affect the collection or quality of labor force data. So it is reasonable to assert that of all the characteristics examined in this paper, the labor force data is the most critical. There is a higher proportion of people not in the labor force, based on the CPS noninterview weight: 33.1 percent of householders compared to 32.5 percent for each of the other two methods. But the unemployment rates among the three methods were not significantly different. The unemployment rate for householders using the current household weight was 4.26 percent, 4.16 percent using the noninterview weight, and 4.26 percent based on the regional control weight.

## 7. Discussion and Further Research

Our results indicate that both of the alternatives we examined produce distributions of household and householder characteristics that, with a few exceptions, correspond closely to those of the current CPS household weight. The proportions for categories for a given characteristic are generally consistent among the three methods, although the absolute estimates are not. This is the primary consideration for any change in the household weighting: that the estimates for critical household and householder characteristics will not change significantly. For either of these alternatives, would we consider the household and householder weights interchangeable? Using either the noninterview weight or the regional control weight for the household/householder estimation would not affect the person weighting and estimation. But the facts that these alternatives would mean a 'break' from the person-based weighting process, and a significant drop

in the resulting household estimates from the current method, would need to be considered carefully.

The Housing Unit Control Working Group had other suggestions for survey weighting and estimation research that have not yet been pursued extensively, and are not covered in this paper.

Within the CPS, the possibility of raking to get agreement between household and householder estimates could be investigated. If household and householder estimates were raked to achieve consistency, we would still need to consider differences between the estimates from the CPS and from other Census sources. Another approach to investigating the household vs. householder difference would be to assign household weights differently. For example, we could use an average of the weights of all members of a household instead of a single person weight, to evaluate the effect on our household estimates. The ‘family equalization’ method used in weighting for the CPS Annual Social and Economic Supplement (ASEC) could be investigated for possible application in the basic CPS weighting. This step equalizes the estimates of people in married and unmarried partnerships. In particular, estimates of males and females in male/female partner households are made equal. The estimates of households and householders are also consistent after the family equalization step. See Tupek (2005) for more on the family equalization step in the CPS ASEC weighting.

Much of this paper has focused on the differences between the current CPS household estimates and estimates from other Census sources. The regional control weight that we discussed as an alternative is a one-step ratio adjustment of the current household weight. Although we adjusted the regional controls by applying vacancy rates from the HVS to obtain estimates of households, we could investigate further use of HVS vacancy estimates. For example, we could adjust the weights of CPS households, so that the sum of these and the HVS estimates of vacant units equal the independent estimates provided by the Bureau’s Population Division.

The topics for further research discussed in this section illustrate the point that the work presented in this paper is preliminary to further research on alternative household weighting in the CPS and other Census Bureau demographic surveys. The variety and extent of work that could still be done is substantial.

**8. References**

Love, Susan. “Recommendations by the Housing Unit Control Working Group for Research on the Adjustment of Survey Housing Controls and Their Use”, September 3, 2002. Internal Census memorandum.

Love, Susan. “Recommendations by the Housing Unit Control Working Group for Research on Weighting Occupied Housing Units, Households and Householders in Current Surveys”, January 28, 2003. Internal Census memorandum.

Tupek, Alan. “Basic Weighting Specifications for the Current Population Survey Including Replicate Weighting (ESTWGT2000-2)”, January 22, 2004. Internal Census memorandum.

Tupek, Alan. “CPS Supplements: 2005 Weighting Specification for the CPS Annual Social and Economic (ASEC) Supplement”, undated. Internal Census memorandum.

Tupek, Alan. “Weighting Specifications for the Housing Vacancy Survey (HVS2000-1)”, June 1, 2004. Internal Census memorandum.

U.S. Census Bureau, Current Population Survey Design and Methodology. Technical Paper 63RV, March 2002.

U.S. Census Bureau, “Methodology: State and County Housing Unit Estimates”, 2004  
[www.census.gov/popest/topics/methodology/2004\\_HU\\_meth.html](http://www.census.gov/popest/topics/methodology/2004_HU_meth.html)

Table 1. July-September 2005 CPS: Ratio of Estimated Households to Estimated Householders, by Race/Ethnicity of Householder (estimates in thousands, ratio standard errors in parentheses)			
Race/Ethnicity of Householder	# Households	# Householders	Ratio: Households/ Householders
White NonHispanic	82,062	82,272	0.997 (.0056)

Black NonHispanic	13,781	13,951	0.988 (.0120)
Hispanic	12,395	12,708	0.975 (.0139)
Asian NonHispanic	4,184	4,232	0.989 (.0247)
Other	2,066	1,999	1.034 (.0491)
Total	114,488	115,162	0.994 (.0039)

Table 2. July-September 2005 CPS: Ratio of Estimated Households to Estimated Householders, by Sex and Age of Householder (estimates in thousands, ratio standard errors in parentheses)

Sex and Age of Householder		# Households	# Householders	Ratio: Households/ Householders
Male	< 20	439	436	1.007 (.0960)
	20-24	2,671	2,662	1.003 (.0406)
	25-34	10,091	10,271	0.983 (.0195)
	35-44	12,728	12,989	0.980 (.0170)
	45-54	12,901	13,104	0.985 (.0170)
	55+	20,335	20,398	0.997 (.0131)
	Total	59,165	59,860	0.988 (.0051)
Female	< 20	531	531	1.000 (.0837)
	20-24	3,217	3,213	1.001 (.0356)
	25-34	9,191	9,177	1.002 (.0203)
	35-44	10,947	10,944	1.000 (.0184)
	45-54	10,841	10,840	1.000 (.0185)
	55+	20,596	20,596	1.000 (.0126)
	Total	55,323	55,301	1.000 (.0055)
Total		114,488	115,162	0.994 (.0039)

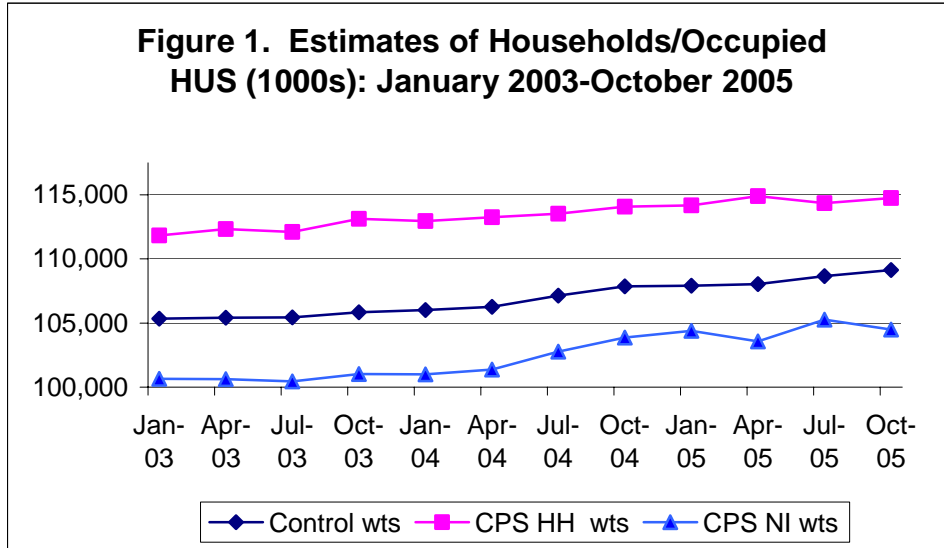
Table 3. July-September 2005 CPS: Ratio of Estimated Households to Estimated Householders, by Type of Household (estimates in thousands, ratio standard errors in parentheses)

Type of Household	# Households	# Householders	Ratio: Households/ Householders
MC <sup>7</sup> with at least one child 6 or under	13,581	13,834	0.983 (.0174)
MC with at least one child between 7 and 17	12,811	13,054	0.981 (.0179)
Other HH <sup>8</sup> with at least one child 6 or under	5,737	5,731	1.001 (.0280)
Other HH with at least one child between 7 and 17	7,260	7,253	1.001 (.0248)
MC no children	31,423	31,647	0.993 (.0110)
Single adult, no children under 18	30,733	30,733	1.000 (.0112)

<sup>7</sup> MC is married-couple

<sup>8</sup> HH is household

Two or more adults, no children under 18	12,943	12,910	1.003 (.0183)
Total	114,488	115,162	0.994 (.0039)



**Table 4. July-September 2005 CPS: Comparison of Distribution of Households by Race/Ethnicity of Householder (in thousands, standard errors in parentheses)**

Race/Ethnicity of Householder	CPS Household Weight		CPS Noninterview Weight		CPS Weight Adjusted by HU Estimates	
	N	%	N	%	N	%
White NonHispanic	82,062	71.7 (0.2)	76,285	*72.9 (0.2)	78,035	71.7 (0.2)
Black NonHispanic	13,781	12.0 (0.2)	11,433	*10.9 (0.2)	13,106	12.0 (0.2)
Hispanic	12,395	10.8 (0.1)	11,294	10.8 (0.2)	11,719	10.8 (0.1)
Asian NonHispanic	4,184	3.7 (0.1)	3,562	*3.4 (0.1)	3,959	3.6 (0.1)
Other	2,066	1.8 (0.1)	2,073	*2.0 (0.1)	1,955	1.8 (0.1)
Total	114,488	100.0	104,647	100.0	108,775	100.0

\* Proportions are significantly different from other weighting methods.

**Table 5. July-September 2005 CPS: Comparison of Distribution of Households by Number of Persons in Household (in thousands, standard errors in parentheses)**

Persons in Household	CPS Household Weight		CPS Noninterview Weight		CPS Weight Adjusted by HU Estimates	
	N	%	N	%	N	%
1	30,741	26.9 (0.2)	27,959	26.7 (0.2)	29,218	26.9 (0.2)
2	38,283	33.4 (0.2)	35,228	33.7 (0.2)	36,372	33.4 (0.2)
3	18,282	16.0 (0.2)	16,603	15.9 (0.2)	17,371	16.0 (0.2)
4-5	23,152	20.2 (0.2)	21,170	20.2 (0.2)	21,991	20.2 (0.2)
6+	4,031	3.5 (0.1)	3,687	3.5 (0.1)	3,823	3.5 (0.1)
Total	114,488	100.0	104,647	100.0	108,775	100.0

Type of Household	CPS Household Weight		CPS Noninterview Weight		CPS Weight Adjusted by HU Estimates	
	N	%	N	%	N	%
MC with at least one child 6 or under	13,581	11.9 (0.1)	12,312	11.8 (0.2)	12,895	11.9 (0.1)
MC with at least one child between 7 and 17	12,811	11.2 (0.1)	11,923	11.4 (0.1)	12,169	11.2 (0.1)
Other HH with at least one child 6 or under	5,737	5.0 (0.1)	4,967	*4.7 (0.1)	5,450	5.0 (0.1)
Other HH with at least one child between 7 and 17	7,260	6.3 (0.1)	6,559	6.3 (0.1)	6,896	6.3 (0.1)
MC no children	31,423	27.4 (0.2)	29,360	*28.1 (0.2)	29,860	27.5 (0.2)
Single adult, no children under 18	30,733	26.8 (0.2)	27,952	26.7 (0.2)	29,211	26.9 (0.2)
Two or more adults, no children under 18	12,943	11.3 (0.1)	11,575	11.1 (0.1)	12,293	11.3 (0.1)
Total	114,488	100.0	104,647	100.0	108,775	100.0

\*Proportions are significantly different from other weighting methods.

Householder Status	CPS Household Weight		CPS Noninterview Weight		CPS Weight Adjusted by HU Estimates	
	N	%	N	%	N	%
Employed	73,737	64.6 (0.2)	66,882	64.1 (0.2)	70,050	64.6 (0.2)
Unemployed	3,281	2.9 (0.1)	2,901	2.8 (0.1)	3,118	2.9 (0.1)
Not in Labor Force	37,045	32.5 (0.2)	34,493	*33.1 (0.2)	35,204	32.5 (0.2)
Total**	114,062	100.0	104,276	100.0	108,371	100.0
Unemployment Rate (%)	4.26 (0.1)		4.16 (0.1)		4.26 (0.1)	

\* Proportions are significantly different from other weighting methods.

\*\* Householders missing labor force data are excluded from Table 9.