1990 CENSUS RACE AND AGE MODIFICATION

Source: US Bureau of the Census 1990 CPH-L-74

AGE, SEX, RACE, AND HISPANIC ORIGIN INFORMATION FROM THE 1990 CENSUS: A COMPARISON OF CENSUS RESULTS WITH RESULTS WHERE AGE AND RACE HAVE BEEN MODIFIED

Essentially the race statistics were modified to be consistent with the classification used in data sets other than the census, while the age data were adjusted to correspond with the April 1, 1990 census date. **These "modified" data are consistent with the counts of the 1990 census as enumerated, which do not include any adjustments for census undercount or overcount.**

RACE MODIFICATION

There were 9,804,847 *other race* persons included in the 1990 census. These people were not included in one of the 15 racial categories listed on the census form. The existence of this group is inconsistent with the race categories defined by the Office of Management and Budget in *OMB Directive 15*. Such "non-specified" race persons also are not found in data sources other than the census. In order to serve the needs of some portions of the user community, it is necessary to assign each of these persons to a specified race.

After evaluating many alternatives, the following race assignment rule was used: Assign each *other race* person to the specified race reported by a nearby person with an identical response to the Hispanic origin question. The only alternative specifications which yielded significantly different national results were those where responses to the Hispanic origin question were ignored during the assignment of race. This work was done separately for each of the 449 district offices.

Background Information on the Race Assignment Rule

1. The specific Hispanic origin of each *other race* person in the 1990 census was taken into account when assigning them to a specified race. This was considered appropriate because over 95 percent of the *other race* persons were of Hispanic origin. Their origin response was used, whether or not it had been allocated, in order to preserve the race distribution within each type of origin. The specific Hispanic origin responses were "not Spanish/Hispanic, Mexican, Puerto Rican, Cuban, and other Spanish/Hispanic".

2. Virtually every person who reported both a specified race and an origin was included in the "donor pool" of eligible persons. The sole exception was the exclusion of several non-specific American Indian codes from the donor pool since: 1) preliminary 1990 research suggested questionable reporting in the American Indian category and; 2) previous research showed that a high proportion of such persons were much less likely to be American Indians than those who actually provided a specific tribe response as instructed on the census form. These codes were: 548--Amerind White, 549--Amerind Black, 597--American Indian (no tribe reported), 598--American

Indian (tribal responses not elsewhere classified), and 973--FOSDIC circle with no write-in response. These were excluded because of evidence from the 1980 census that misreporting of race was much higher in these codes than it was in codes representing specific American Indian tribes. Consistent with advisory committee recommendations, any person assigned to the American Indian race through allocation was given code 973 rather than a specific tribal code.

3. The assignment of a specified race was made on an individual basis. That is, no effort was made to minimize racial heterogeneity within households. Any such attempt would have made it difficult to assign race in a manner which approximated the specified-race distribution reported by persons with the same Hispanic origin response.

4. The race, origin, or sex of some persons also changed as a result of the assignment of a different age to them during the application of the age modification procedures. Their changed age sometimes caused the person to be allocated a different relationship and/or sex, which resulted in the person receiving their race or origin from a different person in the household (since those items were allocated according to a hierarchy of relationships).

5. The results of the race modification procedures were overridden in four counties where the American Indian population grew by more than 100 percent and also became at least one percentage point more of the county's population: Adams County, Washington; Harmon County, Oklahoma; Clark County, Idaho; and Washington County, Idaho. In each case the persons were made white.

6. In most census allocation procedures, acceptable data from eligible persons (donors) are far more common than are the cases where the value is assigned to persons without the characteristic (the donees). This means information from any given donor is rarely used more than once. However, such large donor-to-donee ratios were not usual here. There were a number of occasions where those needing a specified race outnumbered those who reported the same origin as well as a specified race.

Special procedures were adopted to minimize the distortions produced by such low ratios of donors to donees in some cells of these matrices. The most important change was to expand each cell of every allocation matrix to hold 64 values, rather than the 8 stored in most 1990 census matrices. For each district office, the initial race data were then obtained by running the actual census file and storing up to the first 64 acceptable cases in each data cell. If the district office did not have 64 cases, the remaining values were assumed to be white. After the matrices were so "warmed," the program went back to the beginning of the data and began the race assignments. The actual race assignment was a random selection from the unused values within each cell. If all 64 values had been used, then assignment was random from among all 64 until new acceptable cases were available. Used values in a cell were replaced with acceptable data in reverse of the order of use.

AGE MODIFICATION

The following is a portion of the text of a user note which is incorporated in 1990 census products. It explains why the data were modified.

Age Reporting--Review of detailed 1990 information indicated that respondents tended to provide their age as of the date of completion of the questionnaire, not their age as of April 1, 1990. In addition, there may have been a tendency for respondents to round up their age if they were close to having a birthday. It is likely that approximately 10 percent of persons in most age groups are actually 1 year younger. For most single years of age, the misstatements are largely offsetting. The problem is most pronounced at age 0 because persons lost to age 1 may not have been fully offset by the inclusion of babies born after April 1, 1990 and because there may have been more rounding up to age 1 to avoid reporting age as 0 years. (Age in completed months was not collected for infants under age 1.)

The reporting of age 1 year older than age on April 1, 1990 is likely to have been greater in areas where the census data were collected later in 1990.

About 95 percent of the population provided acceptable birth year responses which were adjusted with the following procedures. The age data for individuals in households were modified by adjusting the reported birth year data by race and sex for each of the 1990 Census' 449 district offices to correspond with the national level quarterly distribution of births available from the National Center for Health Statistics. The data for persons in group quarters were adjusted on a state basis to minimize the number of matrix cells with very small numbers. The central assumption in this procedure is that there is no reason for the residents of any subnational area to have a different quarter of birth distribution from that found at the national level. It was also assumed that there are no significant birth place-sex-race-origin differences in annual birth distributions by quarter, that all those born before 1920 have the same quarter of birth distribution, and that mortality is not selective by quarter of birth. Approximately 100 million persons have an age in this modified file which is one year different from that they had in the 1990 census.

The modification procedure was done separately for each birth year, by sex, for the white; black; Asian or Pacific Islander; and American Indian, Eskimo, or Aleut populations. For every birth year the program was provided with the number of cases out of 10,000 where the birth year plus the person's age should equal 1989 (most years that number was close to 7,500). These overall control values were calculated from the available monthly birth statistics for the 1920 to 1989 period. Earlier birth years were each assumed to have the same seasonal pattern as was exhibited by the unweighted average of the 1920-24 birth cohorts. Each sex-race-origin cell was next randomly assigned a value of 0, 0.25, 0.5, or 0.75. Then, each time that birth year cell was encountered, a test was made to see if that birth year plus the person's age should equal 1989 or 1990. The choice of an age was dependent on whether its acceptance moved the cell's actual population distribution toward the control value distribution.