Census 2000 Special EEO Tabulation
Technical Documentation for SAS Datasets

Appendix 3: Things to Know Before Using the Special EEO Tabulation Datasets

1.   **Dataset descriptions:** The Special EEO Tabulation is a set of 24 separate datasets (actually each dataset is a set of two datasets or versions, one version covers all areas except Hawaii and a second version covers Hawaii only; the only difference between the versions is that the Hawaii version has 15 race/ethnicity categories while the version for all areas except Hawaii contains only 12 – for further explanation see #2 below).

   - Datasets 1 through 7 are <u>residence</u> tabulations. The tabulations are presented for the civilian **labor force** 16 years and older by occupation categories, race and ethnicity, and sex based on the location where they live.

   - Datasets 8 through 19 are <u>workflow</u> tabulations. The tabulations are presented for a central worksite (county or place of 50,000 or more population) showing the flow to that worksite of workers from up to nine individual counties, places, or balance of counties. The data are shown for civilian **workers** 16 years and older by occupation categories, race and ethnicity, and sex based on the location of the worksite.

   - Datasets 20 through 24 are <u>worksite</u> tabulations. The tabulations are presented for civilian **workers** 16 years and older by occupation categories, race and ethnicity, sex, and one of the following: educational attainment, older age groups, industry, and earnings. The tabulations are based on the location of the worksite.

      See Appendix 4.1 for a detailed description of the 24 datasets involved with the Special EEO Tabulation.

2.   **Race and Ethnicity:** For all geographies outside of Hawaii (including the U.S. total), data are shown for 12 categories of race and ethnicity (see Appendix 4.2 for a description of those categories). For Hawaii only, three additional race and ethnicity categories are added to identify people of multiple races, one of which is Native Hawaiian and Other Pacific Islander (see Appendix 4.2 for a description of the three additional categories).

3.   **County Subdivisions in New England:** In those datasets that present data below the county level, the county subdivision shown is the place in all states outside of New England. For the New England states, the county subdivision shown is the minor civil division (MCD). Below is a list of the MCDs of 50,000 or more population that are shown for New England.

Connecticut

<u>Fairfield County</u>
Bridgeport town
Danbury town
Fairfield town
Greenwich town
Norwalk town
Stamford town

<u>Hartford County</u>
Bristol town
Hartford town
Manchester town
New Britain town
West Hartford town

<u>New Haven County</u>
Hamden town
Meriden town
Milford town
New Haven town
Waterbury town
West Haven town

Maine

<u>Cumberland County</u>
Portland city

Massachusetts

<u>Bristol County</u>
Fall River city
New Bedford city
Taunton city

<u>Essex County</u>
Haverhill city
Lawerence city
Lynn city

<u>Hampden County</u>
Chicopee city
Springfield city

<u>Middlesex County</u>
Cambridge city
Framingham town
Lowell city
Malden city
Medford city
Newton city
Somerville city
Waltham city

<u>Norfolk County</u>
Brookline town
Quincy city
Weymouth town

<u>Plymouth County</u>
Brockton city
Plymouth town

<u>Suffolk County</u>
Boston city

<u>Worcester County</u>
Worcester city

New Hampshire

<u>Strafford County</u>
Manchester city
Nashua city

Rhode Island

<u>Kent County</u>
Warwick City

<u>Providence County</u>
Cranston City
Pawtucket City
Providence City

4. **Results of Disclosure Avoidance:** Because of the need to implement various disclosure avoidance techniques to protect respondent privacy, users of the Special EEO Tabulation should be aware of the following:

a. Seven metropolitan areas (MSAs and PMSAs) are not included in datasets 1 and 2 because to identify them in conjunction with identifying County Sets would result in showing data for an area of under 50,000 population. The seven areas are –

Arizona: Flagstaff AZ-UT MSA - FIPS MSA Code #2620
Arkansas: Fort Smith AR-OK MSA - FIPS MSA Code #2720
Maryland: Baltimore MD PMSA - FIPS PMSA Code #0720
New York: Newburgh NY-PA PMSA - FIPS PMSA Code #5660
North Dakota: Grand Forks ND-MN MSA - FIPS MSA Code 2985
Pennsylvania: Pittsburgh PA MSA - FIPS MSA Code #6280
Wisconsin: La Crosse WI-MN MSA - FIPS MSA Code #3870

However, all metropolitan areas (MSAs and PMSAs) are shown in datasets 3, 4, 5, and 7.

b. Sixty places with populations of 50,000 or more are not shown in datasets 14 through 19 because showing them in conjuction with data for the county they are in would result in the identification of a balance of the county that is under 50,000 population. However, all places of 50,000 or more are shown in datasets 3 through 6. The sixty places not shown are –

Alabama:
> Dothan city
> Montgomery city

Arkansas:
> Fort Smith city
> Jonesboro city
> Pine Bluff city

Colorado:
> Pueblo city

Florida
> Jacksonville city

Georgia:
> Albany city
> Athens-Clarke County (balance)
> Augusta-Richmond County (balance)
> Columbus city (balance)

Idaho:
> Idaho Falls city
> Pocatello city

Illinois:
> Decatur city

Indiana:
> Terre Haute city

Iowa:
> Ames city
> Council Bluffs city
> Dubuque city
> Iowa City city
> Sioux City city

Kansas:
> Kansas City city
> Lawrence city
> Topeka city

Kentucky:
> Owensboro city

Louisiana:
> Bossier City city

Minnesota:
> Rochester city

Missouri:
> Kansas City city
> St. Joseph city

Montana:
> Billings city
> Great Falls city
> Missoula city

Nebraska:
> Lincoln city

New Mexico:
> Rio Rancho city

North Carolina:
> Durham city
> Rocky Mount city

North Dakota:
> Bismarck city
> Fargo city

Oklahoma:
> Lawton city

Oregon:
> Gresham city
> Salem city

South Dakota:
> Rapid City city
> Sioux Falls city

Tennessee:
> Clarksville city
> Jackson city

Nashville-Davidson (balance)
Texas:
    Abilene city
    Amarillo city
    Bryon city
    Corpus Christi city
    Laredo city
    Longview city
    Lubbock city
    Midland city
    Odessa city
    San Angelo city
    Victoria city
    Wichita Falls city
West Virginia:
    Huntington city
Wisconsin:
    Eau Claire city
Wyoming:
    Cheyenne city

c.   Thirty-four places with populations of 100,000 or more are not shown in datasets 21 through 24 because showing them in conjuction with data for the county they are in would result in the identification of a balance of the county that is under 100,000 population.
The thirty-four places not shown are:

Alabama:
    Montgomery city
Colorado:
    Pueblo city
Florida:
    Jacksonville city
    Tallahassee city
Georgia:
    Athens-Clarke County (balance)
    Augusta-Richmond County (balance)
    Columbus city (balance)
Illinois:
    Peoria city
    Springfield city
Indiana:
    Evansville city
    Indianapolis city (balance)
Iowa:
    Cedar Rapids city

Kansas:
     Kansas City city
     Topeka city
Louisiana:
     Lafayette city
     Shreveport city
Michigan:
     Lansing city
Mississippi:
     Jackson city
Missouri:
     Independence city
     Kansas City city
     Springfield city
Nebraska:
     Linclon city
     Omaha city
North Carolina:
     Durham city
South Dakota:
     Sioux Falls city
Tennessee:
     Clarksville city
     Nashville-Davidsonville (balance)
Texas:
     Abilene city
     Amarillo city
     Corpus Christi city
     Laredo city
     Lubbock city
     Waco city
     Wichita Falls city

5. **Effects of Required Rounding:** The Census Bureau's internal Disclosure Review Board (DRB) is responsible for reviewing specifications for all census data products, and determining that no product format is approved that contains any degree of disclosure risk. To meet the DRB's disclosure avoidance standards, the Special EEO Tabulation was created using approved rounding rules, that are described in the "Protecting Privacy" appendix. Basically, all counts in these datasets are rounded.

Specifically, in the EEO residence and worksite datasets (#1-7 and #20-24), each occupational group contains combinations of characteristics that are individually rounded (for example, the sex – race group variables in dataset #1, such as White non-Hispanic males) and also a total civilian labor force variable that was constructed based on the unrounded total and was then rounded (named "total" on

the datasets). This approach provides users with occupational group totals that are closer to the actual unrounded estimates. However, the total civilian labor force value may not equal the sum of the individually rounded characteristic variables for each occupational group. (This rounding procedure occurs for each observation in the dataset; for example, each geographic unit's occupational group.) In addition, the totals for a particular summary level of geography (state) will likely not equal the sum of the geographies within that summary level (counties).

In the EEO worksite/residence (also known as workflow) datasets (#8-19), the rounding of the total number of people employed at work (named "people_r" on the datasets) occurs for each observation in the dataset, which is a combination of occupational groups by race by sex for every place of work by place of residence occurrence. Some totals may not agree between datasets for the same geographic areas and/or occupational groups, because of the rounding of estimates.

Data from the Special EEO Tabulation will not agree with estimates from other Census 2000-based products (such as Summary Files 3 and 4) because of universe differences and rounding. Use published SF-3 and SF-4 data products to obtain estimates of the total civilian labor force, total number of employed, or high-level occupational groups for the civilian employed population. The Special EEO Tabulation gives users access to more detailed occupational group data.

6. The Special EEO Tabulation reflects corrections to a number of county to county flows that were discovered after the release of Census 2000 SF3 and SF4 data to have errors. In general, the errors were the result of name similarity/duplication/confusion between counties, MCDs, and places, oftentimes the results of respondents mis-fielding their answers. The major exception to this, and the largest correction made, was the incorrect geocoding of the Warner-Robbins Air Force Base. This base was geocoded to a tract and block in the wrong county. About 7,800 workers were moved as a result of the correction. Most of the other corrections involved flows of less than 150 workers, although a few were between 150 and 500 workers.