

## Confidentiality Procedures for the Public Use Files

All of the editing and imputation procedures described in the preceding sections are part of the process of preparing the data for internal Census Bureau use. Before the files are released for public use, they undergo additional editing to protect the confidentiality of respondents. Two procedures are used: topcoding of selected variables (income, assets, and age) and suppression of geographic information. As a result of these procedures, estimates based on data from the public use files will differ slightly from the Census Bureau's published estimates

### Topcoding

One piece of information that might reveal a respondent's identity is a very high income. For that reason, the Census Bureau topcodes income before making that information publicly available, recoding any income amounts over a certain maximum value to that maximum. In other words, income on the public use data files has a ceiling value. Although income is the primary variable that is topcoded, other variables that may disclose a respondent's identity, such as age, are also topcoded. A few variables, such as starting dates for employment, may be bottomcoded if they pose a disclosure risk.

### Suppression of Geographic Information

Geographic information that can be used to directly identify survey respondents, such as an address, is removed from the public use files. In addition, states and metropolitan areas with populations less than 250,000 are not identified. Specific nonmetropolitan areas (such as counties outside of metropolitan areas) are never identified. In certain states, when the nonmetropolitan population is small enough to present a disclosure risk, a fraction of that state's metropolitan sample is recoded to nonmetropolitan status. For that reason, the SPD data cannot be used to estimate characteristics of the population residing outside metropolitan areas.

For the SPD, state-level geography is shown for 41 individual states and the District of Columbia; the nine other states are combined into three groups:

1. Maine, Vermont;
2. Iowa, North Dakota, South Dakota; and
3. Alaska, Idaho, Montana, Wyoming