

Adjusting 2000 census SF3 housing data to fit SF1 totals for counties, cities and towns

The Office of Financial Management (OFM) requires 2000 Census housing data by structure type as benchmarks for its county, city and town population estimates during the post-census decade. For reasons explained below, the raw detailed census data for population and housing were not suitable for estimation purposes. Adjustments had to be made.

The 2000 U.S. Census collected two basic sets of data. One set was from a complete enumeration of households. Its results were released in what the Census Bureau calls *Summary File 1* (SF1). The other set came from a one-in-six sample of households and was released in *Summary File 3* (SF3).

SF1 does not include detailed information on population, housing units, and housing unit occupancy classified by type of structure (number of housing units per structure, mobile homes versus conventional construction, etc.). These data are found in SF3 only. In most cases, the sum of units by type as reported in SF3 does not agree with the total number of housing units found in SF1. The same applies to population and occupied units. The Census Bureau recommends that SF1 data should be used in preference to SF3, because the former is from a complete count and the latter from a sample. Therefore, OFM adjusted the SF3 detailed data to conform to control totals from SF1.

The adjustment procedure can be summarized as follows:

- 1. SF3 detailed structure types were aggregated into the broader categories of structure types used in OFM's estimation database.**
- 2. The category-type-specific data were forced to agree with totals from SF1.**
For example, if SF1 reported the population in housing units as 1,473 and the sum of populations by structure type from SF3 was 1,496, the SF3 populations were reduced (proportionally) so that their sum was now 1,473, as reported in SF1. This was done for every county and each city and town. Where local boundaries crossed county lines, the parts within separate counties were dealt with individually.
- 3. Other selected adjustments were made to correct for logical inconsistencies.**
Some anomalies were found in the adjusted data for various structure types when derived ratios were calculated. That is, there might be more occupied units than total units. There might be population, but no occupied units. This occurred because some of the original SF3 data were logically inconsistent (i.e., 19 people living in 20 occupied housing units) and because the forcing was done separately for population, housing units and occupied units. The only "logical" inconsistency allowed to persist was that the number of housing units in duplexes might not be an even number.

A software system developed by OFM corrected for such anomalies by shifting small numbers of people and housing units from one type to another so that household size, occupancy rates, and so forth were logically consistent. In nearly all cases, only one housing unit and one or two people were shifted from one structure type category to another. Similar adjustments were made for the unincorporated county balances; county-level structure data were adjusted so that the results of subtracting incorporated data from county data yielded logically consistent results.

4. **No adjustments were made for “odd” or “unlikely” results most likely due to SF3 content derived from quirks in the sample.**

There are a few instances where “odd” or “unlikely” results can be found—such as an average household size greater than 15. No adjustments were made. Only logically inconsistent data were adjusted.