

Methods for Estimating Intercensal (2010 to 2019) and Postcensal (2020 to 2024) Residential Population by Census Tract

Introduction to Intercensal Estimates, 2010 to 2019

Intercensal estimates of residential population by age (0, 1-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 74-79, 80-84, 85-89, 90-94, 95-99, and 100 and over), bridged race (White Non-Hispanic, Black Non-Hispanic, American Indian and Alaska Native Non-Hispanic, Asian and Pacific Islander Non-Hispanic, and Hispanic of any race) and gender were made for all U.S. census tracts (based on 2020 geographies) for the years 2010 (July 1) to 2019.

Intercensal Estimates, 2010 to 2019, Data Sources and Methods

The residential population estimates for all U.S. census tracts for the years 2010 to 2019 were made by assuming that the proportion of population by age, race, and gender for each tract in a particular county would change linearly from the 2010 and 2019 Census proportions. If a tract had 10% of the county male White Non-Hispanic population age 20-24 in 2010 and 12% in 2020 then for each year 2011 to 2019 the proportion would be a linear interpolation of the 2010 and 2020 proportions. The county total residential population by age, race and gender control estimates for the years 2010 to 2020 were from the NCI SEER Population Estimates database, Vintage 2024, with Alaska and Hawaii adjustments.

Census tract data for April 1 2020 are from Census 2020 PCT12H-O tables. The 2020 PCT12 data were summed to the 22 age cohorts and from the single year of age data. Asian Alone Non-Hispanic (PCT12L) and Native Hawaiian and Other Pacific Islander Alone Non-Hispanic (PCT12M) were summed to create Asian and Pacific Islander Alone Non-Hispanic. Other Non-Hispanic (PCT12N) and Two or More Races Non-Hispanic (PCT12O) population by age and gender were distributed proportionally to the four race groups based on the county proportion of Some Other Race Non-Hispanic and Two or More Races Non-Hispanic by race, gender and age for 2020.

The census tract definitions for all years, 2020 to 2019 are the 2020 Census census tract definitions. There are 84,414 census tracts in the United States based on 2020 census tract definitions. The 2010 Census data (in 2010 census tracts) were estimated from Census 2010 PCT12H-O tables using the same method to estimate 2020 census tract data. The 73,057 2010 census tracts for 2010 data were adjusted to 84,414 2020 census tracts using a 2010 to 2020 scalar census tract relationship factor from IPUMS.

The proportions of each race, gender, age cohort for 2010 and 2020 were made by dividing each census tract by the sum of all tracts in a particular county. These proportions were interpolated for 2011 to 2019 (as described above) and multiplied by county July 1 population for 2010 to 2020 to obtain July 1 census tract population by age, gender, and race for all years, 2010 to 2020.

Introduction to Postcensal Estimates, 2020 to 2024

Postcensal estimates of residential population by age (0, 1-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 74-79, 80-84, 85-89, 90-94, 95-99, and 100 and over), bridged race (White Non-Hispanic, Black Non-Hispanic, American Indian and Alaska Native Non-Hispanic, Asian and Pacific Islander Non-Hispanic, and Hispanic of any race) and gender were made for all U.S. census tracts (based on 2020 geographies) for the years 2020 (July 1) to 2024. The residential population estimates for all U.S. census tracts for the years 2020 to 2024 were made using a hybrid regression, demographic, and proportional model.

Postcensal Estimates Census Tract Data for 2020

Census tract data for April 1, 2020 were from Census 2020 PCT12H-O tables. The 2020 PCT12 data were summed to the 22 age cohorts from the single year of age data. Asian Alone Non-Hispanic (PCT12L) and Native Hawaiian and Other Pacific Islander Alone Non-Hispanic (PCT12M) were summed to create Asian and Pacific Islander Alone Non-Hispanic. Other Non-Hispanic (PCT12N) and Two or More Races Non-Hispanic (PCT12O) population by age and gender were distributed proportionally to the four race groups based on the county proportion of Some Other Race Non-Hispanic and Two or More Races Non-Hispanic by age, race, and gender for 2020.

Postcensal Estimates Proportional Model for Census Tracts

An estimate was made of total population by age and gender and of population by age, race, and gender of all years 2020-2024 for all census tracts. These proportional estimates were by age and gender (44 cohorts) were used to control the results of the demographic model to specific age gender cohorts. A second set of proportional model estimates were also made for census tract population by age, race, and gender (220 cohorts). The 220 cohort proportional estimates by age, race, and gender were used to evaluate the demographic model results.

The proportional model assumes that the 2020 proportion of a county (C) age, gender, race cohort for a particular census tract (CT) remain constant for all years 2021 through 2024. For example, for any year t , 2021 through 2024, White male population age 20-24 were estimated as:

$$WM-CT_{20-24}_t = (WM-CT_{20-24}_{2020} / WM-C_{20-24}_{2020}) \times WM-C_{20-24}_t$$

Total population for a census tract in the 220 cohort proportional model is not a constant proportion of the county for the year 2021 to 2024 because the age, race, and gender cohorts of the tract change at different rates based on Vintage 2024 county population changes by age, race, and gender for the years 2021 to 2024.

The proportional model results for all census tracts in a particular county for each age, race, and gender cohort exactly sum to the county totals without any constraining algorithms because the sum of the proportions always equal one.

Postcensal Estimates Regression Model for Census Tracts

For census tracts that were expected, based on contemporaneous data or observations, to grow significantly more rapidly, or more slowly, than the county total, then a lagged ordinary least squares (OLS) regression model was used to allow for divergent growth of particular census tracts.

Census tract total population was split into two age cohorts: 0-39 and age 40+. American Community Survey (ACS) 5-year census tract (based on 2020 geographies) total population age 0-39 and 40+ data for the years 2010 to 2024 were used to identify potential census tracts for the regression model. ACS data census tract data for the years 2010 to 2019 (based on 2010 geographies) were adjusted to 2020 census tract definitions using a 2010 to 2020 scalar census tract relationship factor from IPUMS.

The functional forms of all census tract (CT) OLS equations were:

$$\text{TPOP-CT}_{0-39_t} = f(\text{TPOP-CT}_{0-39_t-1})$$

$$\text{TPOP-CT}_{40+_t} = f(\text{TPOP-CT}_{40+_t-1})$$

The coefficients for each of the two age cohorts were applied identically to all age, race, and gender cohorts within the two OLS cohorts.

The OLS results were not constrained to county totals or to proportional model total population by age and gender census tract totals.

If the following conditions were met for a particular census tract, then the regression model was used to estimate population by age, race, and gender for that census tract:

1. the ACS population year to year change for both cohorts must be either all positive or all negative in at least 12 of the 14 years, 2010-2011 through 2023-2024;
2. all years of the ACS population data, 2010 to 2024, must be greater than zero for both cohorts, i.e. 0-39 and 40+;
3. if a coefficient was greater than 1.0 it must also be greater than the county population average annual rate of growth, 2010 to 2024; similarly, if a coefficient was less than 1.0 it must also be less than the county population average annual rate of growth, 2010 to 2024;
4. adjusted R-squared for both equations must be greater than 0.80.

In total, 897 regression models were used (1,794 equations); eight of the regression models had coefficients less than 1.0 and 1,786 had coefficients greater than 1.0. The average adjusted R-squared was 0.923 and the average T-statistic of the coefficients was 53.2.

Postcensal Estimates Demographic Model for Census Tracts

In models when the terminal census year for tract data is 2010 estimates were made for each of the 220 age, race, and gender cohorts for each census tract for the year 2020 with census tract data for 2000 and 2010 using a modified Hamilton-Perry method [1,2] with imposed constraints and exogenous boundaries. This is the method is described below even though for this model 2020 census tract data, from the 2020 Census, are known.

Postcensal Estimates Demographic Model for Census Tracts (continued)

Implicit survival rates and fertility rates were calculated for each of the 220 age, race, and gender cohorts for each census tract based on 2000 and 2010 data for each tract. The implicit survival and fertility rates were then used to estimate 2020 census tract population by age, race, and gender.

For the 18 age cohorts 10-14 through 95-99 by race and gender an implicit survival rate was used to estimate 2020 census tract population. For example, 2020 White male population age 20-24 for a particular census tract was estimated as:

$$WM_{20-24_2020} = (WM_{20-24_2010} / WM_{10-14_2000}) \times WM_{10-14_2010}$$

This was repeated for all 15 age cohorts by 5 race groups by gender.

The age cohort 85+ by race by gender was estimated using a similar implicit survival rate. For example, 2020 White male population age 85+ for a particular census tract was estimated as:

$$WM_{85+_2020} = (WM_{85+_2010} / WM_{75+_2000}) \times WM_{75+_2010}$$

This was repeated for 85+ population by 5 race groups by gender.

The age cohorts 0, 1-4, and 5-9 were estimated using implicit fertility rates by race. For example, 2020 White male population ages 0, 1-4, and 5-9 for a particular census tract were estimated as:

$$WM_{0_2020} = (WM_{0_2010} / WF_{15-44_2010}) \times WF_{15-44_2020}$$

$$WM_{1-4_2020} = (WM_{1-4_2010} / WF_{15-44_2010}) \times WF_{15-44_2020}$$

$$WM_{5-9_2020} = (WM_{5-9_2010} / WF_{20-49_2010}) \times WF_{20-49_2020}$$

Since in this model the census tract estimates for 2020 are from the 2020 Census and did not have to be estimated using the method described above. The population by age, race, and gender for 2021 to 2024 were estimated by linearly interpolating and extrapolating 2010 and 2020 estimates for each of the 220 age, race, and gender groups for each census tract.

In some cases the initial results of the demographic model, prior to constraining, yielded dramatic population changes for specific census tracts. Approximately 8.2% of all census tract age, race, and gender implicit survival rate estimates for 2020 were either greater than 8.0 or less than 0.5 and were adjusted to these limits. However, even with these limits imposed some census tracts had total population results greater than 1.7 times or less than 0.7 times total population results from the proportional model. If the demographic model exceed these bounds (greater than 1.7 times or less than 0.7 times total population results from the proportional model) for a particular census tract, then the 220 cohort proportional model was used for that census tract. A total of 19,596 census tracts used the 220 cohort proportional model in place of the demographic model because the demographic model yielded results that were out of bounds.

Constraining Postcensal Estimates to County Totals

An iterative bi-directional constraining, or raking, algorithm was used to constrain, or control, the census tract estimates to county totals. The results of both the demographic model and the 220 cohort proportional model (which was used in place of demographic model results for some census tracts), were constrained to both the county total for the cohort and to the census tract total population by age and gender cohort determined in the 44 cohort proportional model for total population by age and gender. The county total to which the 220 cohorts were constrained was the difference of county population less any census tract regression model estimates so that the regression model estimates were not constrained to county totals; similarly, regression model estimates were not constrained to the 38 cohort proportional model results for a particular tract.

References

- [1] Hamilton, C.H., and Perry, J., 1962, "A Short Method for Projecting Population by Age from One Decennial Census to Another", *Social Forces*, vol. 41, pgs. 163-170.
- [2] Swanson, D.A., 2010, "Forecasting the Population of Census Tracts by Age and Sex: An Example of the Hamilton-Perry Method in Action", *Population Research and Policy Review*, vol. 29, pgs. 47-63.

Woods & Poole Economics, Inc.
Washington, D.C.
February 12, 2026