

1.4. HISTORICAL POPULATION

1.4.a. Overview

The HISTORICAL subprocess provides estimates of the Social Security area population for each year in the period December 31, 1940, through December 31, 2022. The Social Security area population consists of:

- U.S. resident population and armed forces overseas *plus*
- Net census undercount *plus*
- Civilian residents of Puerto Rico, the Virgin Islands, Guam, the Northern Mariana Islands, and American Samoa *plus*
- Federal civilian employees overseas *plus*
- Dependents of armed forces and federal civilian employees overseas *plus*
- Residual beneficiaries living abroad *plus*
- Other citizens overseas

The U.S. Census Bureau collects population data by age, sex, and marital status (and by other characteristics) every ten years for the decennial census. Generally, each subsequent year, the Census Bureau publishes a “post-censal” population estimate. This subprocess combines these census and post-censal estimates, along with the estimates of the other components of the Social Security area population listed above, and components of change described in sections 1.1 to 1.3, to develop historical estimates of the total Social Security area population ($P_{x,s}^z$) and temporary or unlawfully present population ($O_{x,s}^z$). Combining the total populations by single year of age and sex with an estimated marital status matrix results in the total Social Security area historical population by single year of age, sex, and marital status ($P_{x,s,m}^z$). Civilian noninstitutionalized populations by single year of age, sex, and marital status ($C_{x,s,m}^z$) are estimated using: (1) total populations by single year of age, sex, and marital status, (2) civilian noninstitutionalized population totals, and (3) assumptions about relationships of marital status between total population and civilian noninstitutionalized population. These estimates are then used as the basis for the PROJECTED POPULATION subprocess described in section 1.8. The primary equations for this subprocess are given below:

$$P_{x,s}^z = P_{x,s}^z(\cdot) \quad (1.4.1)$$

$$P_{x,s,m}^z = P_{x,s,m}^z(\cdot) \quad (1.4.2)$$

$$O_{x,s}^z = O_{x,s}^z(\cdot) \quad (1.4.3)$$

$$C_{x,s,m}^z = C_{x,s,m}^z(\cdot) \quad (1.4.4)$$

1.4.b. Input Data

Long-Range OASDI Projection Data -

Demography

1. Probabilities of death from the MORTALITY subprocess, by age last birthday and sex, for years 1941-2023. These data are updated every year.
2. The number of new arrival LPR immigrants plus adjustments of status LPR immigrants by age and sex for years 1941-2022. These data are from the LPR IMMIGRATION subprocess and are updated each year.
3. The number of legal emigrants by age and sex for years 1941-2022. These data are from the LPR IMMIGRATION subprocess and are updated each year.
4. The number of temporary or unlawfully present immigrants legalized under the Immigration Reform and Control Act of 1986 (IRCA) from the LPR IMMIGRATION subprocess and are updated each year if new data is available.
5. The number of non-IRCA adjustments of status by age and sex for years 1941-2022. These data are from the LPR IMMIGRATION subprocess and are updated each year.
6. Birth rates by single year of age of mother (14-49) for the years 1941-2023 from the FERTILITY subprocess. These data are updated each year.

U.S. Census Bureau Data -

7. Estimates of U.S resident population plus Armed Forces overseas (USAF) population as of each July 1 (1940-79) by sex and single year of age through 84, and for the group aged 85 and older. These data are generally not updated.
8. Estimates of the U.S. resident population for each decennial census (April 1) 1970-2020 by sex and single year of age 0 through 85+. New decennial census estimates come out about every ten years.
9. Estimates of total U.S. resident population and total U.S. resident population plus Armed Forces overseas population for each January of each decennial census year from 1990 to 2020. New decennial census estimates come out about every ten years.
10. Estimates of U.S resident population as of each July 1 (1980-2023) by sex and single year of age 0 through 99, and ages 100 and older. Each year, generally, the U.S. Census Bureau restates the data back to the most recent decennial census and includes one additional year of data.
11. Estimates of U.S. resident plus Armed Forces overseas (USAF) population as of each July 1 (1980-2023) by sex and single year of age 0 through 99, and ages 100 and older. Each year, generally, the U.S. Census Bureau restates the data back to the most recent decennial census and includes one additional year of data.
12. Estimates of U.S resident population as of each January 1 (1981-2023) by sex and single year of age 0 through 99, and ages 100 and older. Each year, generally, the U.S. Census Bureau restates the data back to the most recent decennial census and includes one additional year of data.

13. Estimates of U.S. resident population plus Armed Forces overseas (USAF) population as of each January 1 (1981-2023) by sex and single year of age 0 through 99, and ages 100 and older. Each year, generally, the U.S. Census Bureau restates the data back to the most recent decennial census and includes one additional year of data.
14. Estimates of U.S civilian population as of each July 1 (2010-23) by sex and single year of age for 0 through 99, and ages 100 and older. Each year, generally, the U.S. Census Bureau restates the data back to the most recent decennial census and includes one additional year of data.
15. Estimates of U.S. civilian noninstitutionalized population as of each July 1 (2010-23) by sex and single year of age for 0 through 99, and ages 100 and older. Each year, generally, the U.S. Census Bureau restates the data back to the most recent decennial census and includes one additional year of data.
16. Estimates of U.S civilian population as of each January 1 (2011-23) by sex and single year of age for 0 through 99, and ages 100 and older. Each year, generally, the U.S. Census Bureau restates the data back to the most recent decennial census and includes one additional year of data.
17. Estimates of U.S. civilian noninstitutionalized population as of each January 1 (2011-23) by sex and single year of age for 0 through 99, and ages 100 and older. Each year, generally, the U.S. Census Bureau restates the data back to the most recent decennial census and includes one additional year of data.
18. Estimates of the population by marital status, sex, and age from the American Community Survey (ACS) public use microdata sample (PUMS) files for years 2000-23. In general, an additional year of data is available each year.
19. Estimates of the civilian noninstitutionalized population by marital status, sex, and age from the American Community Survey (ACS) public use microdata sample (PUMS) files for years 2006-23. In general, an additional year of data is available each year.
20. Undercount factors by single year of age (0-85+) and sex, estimated using post-censal survey or demographic analysis data. These data are updated after each decennial census.
21. The total annual population estimates for Puerto Rico, Virgin Islands, Guam, Northern Mariana Islands, and American Samoa for years 1951-2023. For each Trustees Report, an additional data year is downloaded from the U.S. Census Bureau's international database. Historical data back to 1951 is also obtained if any changes have occurred.
22. Decennial census population estimates, by varying degree of age detail and sex, for decennial censuses from 1950-2000 for territories and components outside the 50 states, D.C., and armed forces overseas. Most data are aggregated into 18 age groups for each sex, though single year of age data is available for young ages in the territories for 1960 and 1970 and for all ages starting in 1980. New estimates are added as they become available.
23. July populations of the territories by single year of age and sex from 2000-2023. An additional year of data is available each year.
24. From the ACS PUMS, number of existing marriages from 2006-2023 by age of husband crossed with age of wife. In addition, starting with the 2012 ACS, number of existing marriages by age of each spouse is available for same-sex couples. Final grids for same-sex couples are adjusted based on reported same-sex marriages from the states through 2014. Generally, an additional year of data is available each year.

25. From ACS PUMS, flows of foreign-born ins and Cuban ins by single year of age, sex, and year of entry used to produce most total temporary or unlawfully present population estimates for January 1, 2013-2023.⁹
26. From decennial census PUMS (via the University of Minnesota's IPUMS website), number of existing marriages for decennial census years 1940-2000 by age group of husband crossed with age group of wife. New estimates are added as they become available.
27. From decennial census PUMS (via the University of Minnesota's IPUMS website), estimates of the population by marital status, age, and sex for decennial census years 1940-2000. New estimates are added as they become available.
28. Estimates of net immigration by age and sex of the U.S. resident population plus armed forces overseas (USAF) population from April 1, 2000, through July 1, 2023. In general, an additional year of data is available each year.
29. Total Americans overseas estimate based on international data sources and estimates of federal employees and military in Iraq and Afghanistan. The data from the various international sources are derived from different years but center around the year 2003. Additional data will be updated as they become available.
30. Total civilian population in Alaska and Hawaii for years 1940-49. These data are not updated.
31. Census residential population in Alaska and Hawaii for years 1940 and 1950 by sex and age group and total armed forces in Alaska and Hawaii for 1940 and 1950. These data are not updated.

Other input data -

32. From the Department of State, old historical total estimates of outside area populations (federal employees overseas, overseas dependents of federal employees and military, and other Americans overseas) for various years between 1951 and 1990.
33. The SSA Annual Statistical Supplement provides estimates of the total number of OASDI beneficiaries living abroad as of December 31, for most years 1957-2022. Age group data is also available for 0 – 17, 18 – 54, and 55 – 61 starting in 1997. For each Trustees Report, an additional year of data is available. Age group data is updated for each decennial census year.
34. The SSA Annual Statistical Supplement provides estimates of the total number of OASDI beneficiaries living abroad as of December 31 for age groups 62 – 64, 65 – 69, 70 – 74, 75 – 79, 80 – 84, 85 – 89, 90 – 99, and 100+ starting in 1997. Data is updated decennially so it is available by age group for each census year.
35. From the National Center for Health Statistics (NCHS), births by month and sex of child, for years 1931-2023. Each year, NCHS provides another year of data.
36. From the NCHS, National Survey of Family Growth (NSFG) public-use data that help split up the population eligible for same-sex marriage into marital statuses starting in 2013. These survey data are from 2011-15. This data is updated as they become available.

⁹ Because the 2020 ACS had data collection issues due to the beginning of the COVID-19 pandemic, the 2020 flow data is ignored and the January 1, 2021, stock estimate is an average of the January 1, 2020, and January 1, 2022, stock estimates.

37. From the NCHS, National Survey of Family Growth (NSFG) public-use data that split up the total population, by sex, into the population eligible for same-sex marriage and the population eligible for opposite-sex marriage. These survey data are from 2011-15 and 2015-17. This data is updated as they become available.
38. From the Department of Homeland Security (DHS), the number of unauthorized immigrants and nonimmigrants from January 1, 2005-12. In addition, other estimates from the Unauthorized Immigrant Population Reports and emails from DHS are used to estimate January 1, 2012, through January 1, 2023, unauthorized immigrants and nonimmigrants including an adjustment for shift in reference date, undercount of nonimmigrants, undercount of unauthorized immigrants, legally resident immigrants, foreign-born flows from 1980 and later of LPR immigrants, mortality, and emigration.
39. From the Office of Personnel Management (OPM), total estimates of the number of federal employees overseas from July 1, 1998-2013. These estimates are updated as they become available on the OPM website.
40. From the OPM, the number of federal employees overseas by single year of age and sex from a subset of the OPM data source above. Years 1980-2023 are available. These estimates are updated as they become available.
41. From the Department of Defense, total numbers of armed forces in Puerto Rico, the Virgin Islands, Guam, and American Samoa each decennial census year starting in 1990. These data are updated as they become available. 2020 data is assumed to be the same as 2010 until the data is available to OCACT.
42. Using 2015 TR death rates and historical populations, an assumed December 31, 1940, 85+ distribution. These data are not updated.
43. Assumed January total populations added to the Social Security for years 1951, 1957, and 1961 when new territories were added to the Social Security area. These data are not updated.
44. Armed forces counts, by age group and sex, from 1940-57. These data are not updated.

1.4.c. *Development of Output*

Equation 1.4.1 – Historical Population by age and sex ($P_{x,s}^z$)

The Census Bureau's estimate of the residents of the 50 States, D.C., and U.S. Armed Forces overseas is used as a basis for calculating $P_{x,s}^z$. The base estimate is adjusted for net census undercount and increased for other U.S. citizens living abroad (including residents of US territories) and for non-citizens living abroad who are insured for Social Security benefits.

The estimates of the number of residents of the fifty States and D.C. and Armed Forces overseas, as of July 1 of each year, by sex for single years of age through 84, and for the group aged 85 or older, are obtained from the Census Bureau. January 1 and April 1 estimates by sex for single years of age through 84, and for the group aged 85 or older for selected years starting in 1990 and 1970, respectively, are also obtained from the Census Bureau. Adjustments for net census undercount are estimated using post-censal survey data from the Census Bureau. Population counts over age 65 after the last Census year (2020) are modified to be consistent with OCACT mortality and Census USAF net immigration data. The numbers of persons in the other components of the Social Security area as of July 1 are estimated by sex for single years of age through 84, and for the group aged 85 or older, from data of varying detail. Numbers of people residing in Puerto Rico, the

Virgin Islands, Guam, American Samoa, and the Northern Mariana Islands are estimated from data obtained from the Census Bureau. Numbers of Federal civilian employees overseas are based on estimates from the Office of Personnel Management (OPM). Dependents of Federal civilian employees and Armed Forces overseas are based on the stock of Federal civilian employees from OPM and the stock of armed forces overseas from the Census Bureau. Other citizens overseas covered by Social Security are also based on estimates compiled by the Census Bureau. The overlap among the components, believed to be small, is ignored.

The first step of the process is to estimate $P_{x,s}^z$ as of December 31st for certain “tab years” (1940, 1950, 1956, 1960, each December from 1969 through 2009, and the last year of historical data [2022 for the 2025 Trustees Report]). For ages 0-84, $P_{x,s}^z$ for each tab year, is set equal to an undercount adjustment plus other component populations plus:

- The averaged surrounding July 1 U.S. population and armed forces overseas counts from the Census Bureau prior to 1970
- Modified April 1 U.S. populations from the Census Bureau for decennial census years from 1970 through 2000.
- The January 1 U.S. population and armed forces overseas counts from the Census Bureau for tab years after 2000. For 2023, these populations are modified by OCACT mortality rates and Census USAF net immigration for ages over 65.

For ages 85 and over, $P_{x,s}^z$ for each tab year is set equal to [Built Up Pops Age x, Sex s] * [Total 85+ for Sex s]/[Total Built Up 85+ for Sex s], where the built up estimates are created by taking into account deaths and immigration data from the previous tab year and [Total 85+ for Sex s] is the sum of the calculated U.S. population and armed forces overseas calculated using the same method listed above for each year for ages 0-84.

For years between the tab years, populations are estimated taking into account the components of changes due to births, deaths, legal emigration, adjustments of status, and net LPR immigration (or total net immigration, if known) during that time period. These estimates are then multiplied by the appropriate age-sex-specific ratios so that the error of closure at the tab years is eliminated.

Equation 1.4.2 – Historical Population by age, sex, and marital status ($P_{x,s,m}^z$)

Since eligibility for auxiliary benefits is dependent on marital status, the Social Security area population is disaggregated by marital status. The four marital states are defined as single (having never been married), married, widowed, and divorced.

The distribution of the number of existing marriages are available for decennial census years 1940-2000 from Census public use microdata sample (PUMS) files. These data are aggregated by age group of husband crossed with age group of wife. Additional tabulations from the American Community Survey from 2000-23 are incorporated to adjust these marital prevalence grids for changes since 2000. The grids are transformed from age grouped numbers to single year of age figures from ages 14 to 100+ for husband and wife using the two-dimensional H.S. Beers method of interpolation.

Percentages of single, married, widowed, and divorced persons are calculated by using the H.S. Beers method of interpolation on compiled data by age group and sex based on Census and/or ACS PUMS. Data is converted to a December basis for each year by taking a weighted average of surrounding Census and/or ACS data. Note that the ACS is not used until 2006, the first year grouped quarters were included. These percentages are multiplied by the total populations calculated in Equation 1.4.1 for each age, sex, and year to get a preliminary population for each age, sex, and marital status.

To keep the marriage prevalence grids and the marital status percentages smooth and consistent, several algorithms are used. First, the married population is adjusted so that the number of married men equals the number of married women (though, this is not forced to be ultimately true once same-sex marriages were federally recognized as described in the next paragraph). Then, the number of married persons for each age and sex is set equal to the marginal total of the associated year's marital prevalence grid. Finally, the other marital statuses population totals are adjusted to keep the total number of people in all marital statuses the same as calculated before splitting into marital statuses.

The population is modeled to include the following population statuses for December 31, 2013, and later: heterosexual, gay, and lesbian. Gay and lesbian populations are broken out assuming 2.5% of the male population is gay and 4.5% of the female population is lesbian. (Note that this means we assume 2.5% of the male population and 4.5% of the female population are in an existing same-sex marriage, or would same-sex marry, versus opposite-sex marry.) Marriage grids of age of older spouse crossed with age of younger spouse for same-sex couples are needed starting December 31, 2013. The grids and populations were produced using data from the American Community Survey, National Survey of Family Growth, and state-level same-sex marriage data.

Equation 1.4.3 – Historical Temporary or Unlawfully Present Immigrant Population by age and sex ($O_{x,s}^z$)

This subprocess also estimates historical levels of temporary or unlawfully present immigrants in the population, by age and sex. For each year, an initial net residual estimate by single year of age and sex is backed out from estimates of beginning and end of year populations, births, deaths, LPR immigrants, adjustments of status, and legal emigrants. This net residual equals the implied initial other in minus other out. These residuals are then modified to ensure reasonableness. Next, using these modified net residuals, along with deaths among adjustments of status immigrants and temporary or unlawfully present immigrants (using the same death rates as for the total population), an initial temporary or unlawfully present immigrant stock is built. These stocks are then modified to ensure reasonableness. After 2000, one further adjustment is done to these stocks. From January 2001 through January 2004, the total temporary or unlawfully present immigrant populations are set equal to the values that linearly grade from the January 2000 total temporary or unlawfully present immigrant population to the January 2005 total temporary or unlawfully present immigrant population. From January 2005 through January 2023, the total temporary or unlawfully present immigrant population is forced to match estimates based on the latest methods used by DHS.

Equation 1.4.4 – Civilian Noninstitutionalized Population by age, sex, and marital status ($C_{x,s,m}^z$)

This subprocess also estimates historical levels of civilian noninstitutionalized population, by age, sex, and marital status, beginning with year 2010. Total populations, by age and sex, come directly from the Census Bureau. Then, the subprocess splits these total populations into marital status using data from the ACS PUMS. Unlike with $P_{x,s,m}^z$, the married marital status is split into “married, spouse present” and “separated”.