Methodology for the Intercensal Population and Housing Unit Estimates: 2000 to 2010

Revised October 2012

Introduction

The intercensal estimates for 2000-2010 for the United States and Puerto Rico populations and United States housing units are produced by modifying the 2000-2010 postcensal estimates prepared previously for the United States and Puerto Rico, to account for differences between the postcensal estimates for April 1, 2010 and the 2010 Census counts. The postcensal population estimates for 2000-2010 were produced by updating the resident population enumerated in Census 2000 by the estimates of the components of population change between April 1, 2000 and April 1, 2010.[1] The components include births to U.S. resident women, deaths to U.S. residents, domestic migration, international migration, and net movement of the U.S. armed forces. The postcensal housing unit estimates for 2000-2010 were produced by updating the number of housing units in the Census 2000 by adding new housing units and removing housing units that no longer existed.

The intercensal estimates reconcile the postcensal estimates with the 2010 Census counts and provide a consistent time series of population estimates that reflect the most recent census results. These intercensal estimates are used as survey controls for the American Community Survey and also serve as the bases for determining historical trends in birth and death rates, for projecting future populations, calculating incidence rates for cancer and other diseases, and for tracking changes in other population characteristics.

At the national, state, and county levels of geography, intercensal population estimates were produced by demographic characteristic (age, sex, race, and Hispanic origin). For Puerto Rico, intercensal population estimates were produced by age and sex at the Commonwealth and municipio levels. At the subcounty level of geography, and for housing units, intercensal estimates totals were produced. This document outlines the methods used to produce the 2000-2010 intercensal estimates for the resident population and for housing units. This document also outlines the procedures used to produce July 1, 2010 population and housing unit estimates using the 2010 Census results. This new estimate for July 1, 2010 was added to the intercensal estimates product to meet the growing need for a timely July 1, 2010 estimate to use with the intercensal population and housing unit estimates. The July 1, 2010 estimate is superseded by new population and housing unit estimates as they are released through the annual population estimates program, starting in December 2011.

[1]The postcensal population estimates used to produce the intercensal estimates differ from the Vintage 2010 population estimates released on the website. The postcensal estimates used to produce the intercensal estimates exclude challenges and special census results.

Methodology

There is no universal norm for producing intercensal estimates. The Census Bureau historically has used a method to produce intercensal estimates that was outlined by Prithwis Das Gupta in the early 1980s.[2] This method, hereafter referred to as the Das Gupta method, assumes that the ratio of the intercensal estimate to the postcensal estimate should follow a geometric progression over the decade (see Equation 1). It follows then, by modifying Equation 1, that the intercensal estimates can be produced (as we did) using Equation 2. Put simply, this formula produces intercensal estimates as a function of time and the postcensal estimates.

$$(1) \qquad P_t/Q_t = (P_{3652} / Q_{3652})^{(t / 3652)}$$

$$(2) \qquad P_t = Q_t (P_{3652} / Q_{3652})^{(t / 3652)}$$

Where t = time in days elapsed since April 1, 2000
P_t = population estimate at time t
Q_t = postcensal estimate at time t
P_{3652} = April 1, 2010 census count
Q_{3652} = April 1, 2010 postcensal estimate based on Census 2000

We used the above method for the 2000 to 2010 intercensal population and housing unit estimates.[3] However, as noted below, there were cases where we made exceptions.

Required Exceptions to the Base Das Gupta Interpolation Method

The intercensal estimates were produced for a substantial level of geographic and characteristic detail. There are some instances where the Das Gupta method does not work. For these instances, an alternative method was used.

[2]The 1990-2000 intercensal population estimates were produced using this method, see: http://www.census.gov/popest/methodology/intercensal_nat_meth.pdf.

[3]This method was applied separately for the household and total group quarters populations. Group quarters totals were produced for seven major group quarters types (e.g., correctional facilities for adults, skilled nursing facilities).

A straight line interpolation was used to develop intercensal estimates for specific groups (age, sex, race, and Hispanic origin) for each specific time period when the following conditions arose:[4]

- $Q_t = 0$
- $Q_{3652} = 0$ or 1
- $P_{3652} = 0$ or 1
- Q_{3652} is less than one half of P_{3652}

Where straight line (or linear) interpolation was used, the difference between the Census 2000 count and the 2010 Census count was calculated and spread equally over the decade using Equation 3:[5]

$$(3) \qquad P_t = [P_{3652} * (t / 3652)] + [P_0 * ((3652 - t) / 3652)]$$

Where P_t = population estimate at time t
P_{3652} = April 1, 2010 census count
t = time in days elapsed since April 1, 2000
P_0 = April 1, 2000 census count

Geography-Specific Methods

The methods described above were applied independently at the national, state, and county levels. The national intercensal estimates were produced by characteristic and summed to obtain the total U.S. population. The county intercensal estimates were produced by characteristic and for the total population. State estimates were produced by characteristic, and totals were produced by summing the county total estimates within each state. The following section provides more detail on the specific methods employed for each of the individual levels of geography.

[4]Although these exceptions modify the Das Gupta method, for the sake of ease we still refer to our overall method as the Das Gupta method. In addition to these exceptions, all group quarters (GQ) intercensal population estimates by demographic characteristic were produced using straight line interpolation. The Das Gupta method was not applied because the population characteristics for the GQ postcensal estimates were based solely on the Census 2000 distribution of characteristics by GQ type (e.g., correctional facilities for adults, skilled nursing facilities). Finally, the Das Gupta method also was not applied for the totals by age and sex for the municipios of Puerto Rico.

[5]Throughout this work, the value used for April 1, 2000 reflects changes to the Census 2000 population from the Count Question Resolution program and geographic program revisions. Both the Census 2000 and the 2010 Census counts come from files where the race categories have been modified to reclassify the Some Other Race category. The procedures used to make these modifications are the same for Census 2000 and the 2010 Census and are available at: http://www.census.gov/popest/research/modified.html.

National, State, and County

Estimates at the national, state, and county levels were produced separately using the following approach:

1. Monthly national estimates by demographic characteristic were produced using the Das Gupta method;
2. National totals were created by summing the national population by characteristic;
3. Annual county totals were produced using the Das Gupta method and controlled to the sum of the national detail;
4. Annual state characteristics were produced and underwent a two-way control, to state totals (the sum of county totals) and national characteristics;[6]
5. Annual county characteristics were produced using the Das Gupta method and underwent a two-way control, to county totals and state characteristics.[7]

Puerto Rico Commonwealth and Municipios

Estimates for the Puerto Rico Commonwealth and its municipios were produced separately by single year of age and sex. The following approach was used:

1. Annual Puerto Rico Commonwealth population estimates by age and sex were produced using the Das Gupta method;
2. Population estimates for each municipio were produced using straight line interpolation of the age and sex distributions, and then controlled to the Puerto Rico Commonwealth totals by age and sex.

Subcounty

As with the levels of geography described above, estimates were produced separately for group quarters and household populations at the subcounty level and then summed to obtain the total population. No demographic detail was included with the subcounty estimates.

[6]This was an added control in the re-released intercensal estimates that was applied to better constrain the county characteristic intercensal estimates to the state-level. In the previous release, the county characteristic intercensal estimates were summed in order to produce the state-level intercensal estimates by characteristic.

[7]The rounding procedure used in the re-released intercensal estimates uses the same controls, but was modified to better maintain zeros within the data and results in more accurate county characteristic intercensal estimates.

For the Household Population:

1. Straight line interpolation between Census 2000 (vintage 2010 base population) and the 2010 Census was used to produce total household population estimates for primitive geography (the lowest level of mutually exclusive entities) with unrounded numbers.[8]
2. The primitive geography was controlled to the county intercensal totals;
3. Minor Civil division (MCD) place parts were summed to get non-primitive county place parts. MCD place parts in the balance of county and primitive MCDs were summed to get the non-primitive balance of county in states that have county place parts. MCD place parts were summed to produce non-primitive MCDs. County place parts were used to make incorporated places.

For the Group Quarters Population:

1. The same straight line interpolation methodology described for household population was applied.

Housing Units

Housing units for all levels of geography were estimated using the assumptions and methodologies outlined above:

1. National-level housing units were produced using Das Gupta 6;
2. County-level housing units were produced using Das Gupta 6 and controlled to the national-level housing unit estimate;
3. State-level housing units were produced by summing the controlled county-level estimates;
4. Subcounty housing units were produced using straight line interpolation and then controlled to the county estimates.

[8]There were 21 cases in which a Census 2010 tabulation area did not exist in the Vintage 2010 Estimates GUSSIE base. In those cases we used some information from the Census 2010 Census Tabulation GUSSIE. Specifically, if the 2010 Census was zero, the 2000 population was set to zero, for all other values, the 2000 population was pulled from the 2010 Census Tabulation GUSSIE file. That happened for 10 cases. In all but one case only the value for the new geography was replaced. For Saint Louis County in Minnesota, all 2000 population values were pulled from the 2010 Census GUSSIE tabulation file. This is due to the number of changes in that county.

This is a variation from the Population Estimates' normal procedures for housing unit estimates, but was necessary for two reasons: first, the national control prevents unnecessary shifts in the number of housing units caused by summing up from primitive geography; and second, we needed to have the county-level housing unit estimates available and reviewed in a short timeframe to provide internally to ACS for controls.

Estimating the July 1, 2010 Population

This section describes the method used to estimate the July 1, 2010 population for the same geographic levels and demographic detail outlined above, three months beyond the 2010 Census. The July 1, 2010 population estimates were developed separately for the household population and the group quarters (GQ) populations and were summed to create the resident population.

Estimates at the national, state, county, and subcounty levels were produced separately. State totals were produced by summing the county population estimates for each state; national totals were created by summing the national population by characteristic. The household populations were produced separately, and then summed with the GQ population estimates from the 2010 Census to obtain the total resident population. GQ estimates were held constant at the 2010 Census values because there were no reliable indicators to use for estimating change over the three-month period.

For the household population, the following methods were used for each age, sex, race, and Hispanic origin group:

National Characteristics

$$P = P_{3652} + (Q_{3743} - Q_{3652})$$

Where P = July 1, 2010 population estimate based on the 2010 Census
P_{3652} = April 1, 2010 census count
Q_{3743} = July 1, 2010 postcensal estimate based on Census 2000
Q_{3652} = April 1, 2010 postcensal estimate based on Census 2000

County Totals

$$P = P_{3652} + \{P_{3652} * [((Q_{3743} - Q_{3652}) / Q_{3652})]\}$$

Results were controlled to the national total estimate for July 1, 2010 and summed to create state totals.

State and County Characteristics

State level characteristics were created by controlling the 2010 Census state populations by characteristic to the newly-produced national-level characteristics estimates and the state total population (sum of the newly-produced county totals) estimates for July 1, 2010. Similarly, county level characteristics were produced by controlling the 2010 Census county characteristics to the newly-created state level estimates by characteristic and the county total population for July 1, 2010.

Subcounty Totals

Subcounty level totals were produced by controlling the 2010 Census subcounty household population (primitive geographies) to the new county July 1, 2010 total household estimates. The other subcounty geographic levels were calculated by summing the primitive geographies.

Puerto Rico Commonwealth and Municipios

July 1, 2010 estimates for the Puerto Rico Commonwealth were calculated in the same manner as the national characteristics for July 1, 2010, by age and sex only.

The Puerto Rico municipio estimates were produced by controlling the 2010 Census counts (P_{3652}) to the Puerto Rico Commonwealth estimates by age and sex for July 1, 2010.

Housing Units

To produce county-level housing unit estimates for July 1, 2010, county-level numeric change between April 1, 2010 and July 1, 2010 (Census 2000-based) housing unit estimates was calculated and then applied to the 2010 Census estimates base. The new July 1, 2010 county housing unit estimates were summed to the national and state levels, as needed.

To produce subcounty-level housing unit estimates for July 1, 2010, the 2010 Census subcounty housing units (primitive geographies) were controlled to the new county total July 1, 2010 housing unit estimates. The other subcounty geographic levels were calculated by summing the primitive geographies.