Methodology for the United States Resident Population Estimates by Age, Sex, Race, and Hispanic Origin and the State and County Total Resident Population Estimates (Vintage 2011): April 1, 2010 to July 1, 2011

Each year, the U.S. Census Bureau produces and publishes estimates of the population for each state and county, as well as the nation as a whole. We utilize administrative data from a number of sources to estimate 1) the change in population since the most recent decennial census and 2) the population for each year since the most recent decennial census. With each annual release of population estimates, the entire time series of estimates beginning on April 1, 2010 is revised and updated. The population estimates described in this document are the first estimates based on 2010 Census counts.

This document first describes the population estimates that we produce and publish. The second section describes the administrative data we utilize, our modifications and assumptions regarding these data, and the method we use to estimate the components of population change. The final section describes the methods we use to produce the population estimates from the administrative data and the most recent decennial census data.

Population Estimates Data

For the nation, we release monthly estimates of the resident population by age, sex, race, and Hispanic origin. Additionally, we release national estimates by demographic characteristics of four other populations: the resident plus Armed Forces overseas, civilian, civilian noninstitutionalized, and household populations. Each of these four additional populations is based directly on the resident population.

For each state and county, we release annual estimates of the resident population by age, sex, race, and Hispanic origin. This document describes the production of the total resident population for states and counties, which is the first step in the production of state and county estimates by demographic characteristics. The production of state and county population by demographic characteristics is described in a separate document.

Administrative Input Data

At the national level, the resident population is affected by births, deaths, and net international migration (NIM) only.

$$Pop\ Estimate = Base\ Pop + Births - Deaths + NIM$$

At the sub-national level (i.e., states and counties), the resident population is affected by an additional component of population change: net internal, or domestic, migration (NDM).

$$Pop\ Estimate = Base\ Pop + Births - Deaths + NIM + NDM$$

Base Population

The enumerated resident population from the 2010 Census is the starting point for all post-2010 population estimates. We modify this enumerated population in two ways to produce the April 1, 2010 population estimates base.

We reconcile the 2010 Census race categories with the race categories that appear in our administrative data by recoding the "Some other race" responses in the 2010 Census to one or more of the five 1997 Office of Management and Budget (OMB) race categories: White; Black or African American; American Indian and Alaska Native; Asian; and Native Hawaiian and Other Pacific Islander.[1]

We also update the population estimates base to reflect changes to the 2010 Census population due to the Count Question Resolution (CQR) program, legal boundary updates reported by January 1, 2011, and other geographic program revisions.[2] The base population and the estimates described in this document do not reflect any CQR changes because no CQR revisions had been determined in time for this cycle of estimates production.

Births

To estimate births, we utilize birth certificate data collected by the National Center for Health Statistics (NCHS). NCHS provides final individual birth records for births occurring before January 1, 2010 by date of birth, sex of child, residence and age of mother, and race and Hispanic origin of both mother and father. We modify these data in three main ways for use in estimates production.

First, not all states currently utilize the 1997 OMB race categories to request the parents' race on birth certificates. Some states still record race according to the 1977 OMB race categories of White; Black; American Indian; Eskimo or Aleut; and Asian or Pacific Islander, under the "mark one race" scenario. Therefore, we must convert the parents' reported race from the 1977 OMB race categories into the 1997 OMB race categories. We do this using the race bridging method designed by NCHS and the U.S. Census Bureau to make multiple-race and single-race data comparable.[3]

Second, birth certificates only require the race and Hispanic origin of the parents, not of the child directly. To impute the race and Hispanic origin of each child, we utilize the joint distribution of

[1] The OMB standards are detailed in Office of Management and Budget, "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity" Notice, Vol. 62, No. 210, Thursday, October 30, 1997 http://www.whitehouse.gov/omb/fedreg/1997standards.html.

[2] For more information on the 2010 Census Count Question Resolution (CQR), see http://2010.census.gov/2010census/about/cqr.php.

[3] For more information on the NCHS race-bridging factors, see http://www.cdc.gov/nchs/nvss/bridged_race.htm.

their parents' race and origin and decennial census information on race and origin reporting within family households.

The third modification to the NCHS birth data relates to the differences found in the race and Hispanic origin distributions of reported births and of the estimates base population under 1 year of age. We adjust the distribution of births by race and Hispanic origin to be consistent with the 2010 Census population under 1 year of age.

Once the birth certificate data from NCHS have been modified as needed for estimates production, we calculate age, sex, race, and Hispanic origin-specific birth rates for 2009 using the U.S. Census Bureau's intercensal population estimates for July 1, 2009.[4] With these rates, we estimate the number of births by sex, race, and Hispanic origin occurring from April 1, 2010 to June 30, 2011. This process is described in the national estimates production section below.

For the production of state and county population estimates, we also need to estimate the annual number of births by sex, race, and Hispanic origin occurring in each county. We utilize the information on residence of mother from the calendar year 2009 NCHS data along with more current birth reports from the state members of the Federal-State Cooperative on Population Estimates (FSCPE) to distribute the annual national-level births by sex, race, and Hispanic origin to each county.

Deaths

To estimate deaths, we utilize death data collected by NCHS. NCHS provides final individual death records for all deaths occurring before January 1, 2010 by residence, age, sex, race, and Hispanic origin of each decedent, as well as the place and date each death occurred.

As with birth data, the race of decedents from many states is still reported in the 1977 OMB race categories. Therefore, the first step in preparing death data for use in population estimates production is to convert the decedent's reported race into the 1997 OMB race categories using the NCHS race bridging. The race-bridging process produces the full age, sex, race, and Hispanic origin distribution of monthly deaths for the nation.

Research shows that there are discrepancies in the reporting of age between census counts and death registration, most notably in the oldest ages. To address this issue, we redistribute all deaths occurring to the aggregate population 70 years and older by sex, race, and Hispanic origin to single years of age (70 to 99 and 100+ years) using life-table based death rates.[5] We make no additional adjustments, beyond race-bridging, to the deaths occurring to the population under 70 years of age.

---

[4] The national intercensal estimates data and documentation are available at http://www.census.gov/popest/data/intercensal/index.html.

[5] To derive the death rates for the age 70 and older population, we use a life table produced by the Social Security Administration. This life table and associated documentation are located at http://www.ssa.gov/policy/docs/statcomps/supplement/2009/4c.html#table4.c6.

Once the death certificate data from NCHS have been modified as needed for estimates production, we calculate age, sex, race, and Hispanic origin-specific death rates for 2009 using the U.S. Census Bureau's intercensal population estimates for July 1, 2009. With these rates, we estimate the number of deaths by age, sex, race, and Hispanic origin occurring from April 1, 2010 to June 30, 2011. This process is described in the national estimates production section below.

For the production of state and county population estimates, we also need to estimate the annual number of deaths by age, sex, race, and Hispanic origin occurring in each county. We utilize the information on residence of decedent from the calendar year 2009 NCHS data along with more current death reports from the state members of the FSCPE to distribute the annual national-level deaths by age, sex, race, and Hispanic origin to each county.

Net International Migration

We estimate international migration in several parts: immigration of the foreign born, emigration of the foreign born, net migration between the United States and Puerto Rico, net migration of natives to and from the United States, and net movement of the Armed Forces population to and from the United States. For each component, we first estimate the total migration flow for the nation. To determine the demographic characteristics and geographic distribution of each component, proxy universes are developed that are assumed to be representative of the different components. The demographic characteristics and geographic distribution of these proxy universes are then applied to the totals for each component. For all components except net movement of the Armed Forces population to and from the United States, national characteristics and state-level totals and characteristics are based on the American Community Survey (ACS) three-year 2007-2009 file. County-level totals and characteristics distributions are based on data from the ACS five-year 2005-2009 file. County-level data are controlled to state-level data to ensure the component data sum as required. For the net movement of the Armed Forces population, demographic characteristics and state distributions are based on data collected by the Defense Manpower Data Center (DMDC) and Census 2000.

Immigration of the foreign born is estimated using the ACS question on residence in the prior year. The foreign-born population who indicated that they lived abroad in the prior year are considered immigrants. The number of foreign-born migrants who entered the United States between April 2010 and June 2010 is estimated as one quarter of the foreign-born population in the 2010 ACS who reported living abroad one year ago. Because this question is asked only of those aged one and higher, the estimate of foreign-born immigrants under the age of one is assumed to be equal to half the number of immigrants age one. Information from the 2010 ACS is used to estimate migration for the July 2010 to June 2011 period because more recent data are not available. The foreign-born population whose year of entry was within five years of the survey year is used as the proxy universe to estimate the age, sex, race, Hispanic origin and state distribution of foreign-born immigrants. Age in the ACS is adjusted for foreign-born immigrants to represent age at arrival to the United States.

Emigration of the foreign born is estimated using a residual method. The foreign-born household population in Census 2000 is aged forward (using NCHS life tables) to obtain the expected population in 2007 (2008, 2009, and 2010). The expected population is then compared to the population estimated in ACS 2007 (ACS 2008, ACS 2009, and ACS 2010). Subtracting the estimated from the expected populations provides us with the residual, which serves as the basis for our emigration rates for the 2000 to 2007, 2000 to 2008, 2000 to 2009, and 2000 to 2010 time periods. This calculation is performed for two period-of-entry groups: the foreign born who entered the United States between 1990 and 1999, and the foreign born who entered before 1990.

We then calculate three-year averaged rates for each period of entry group and apply the rates to the population at risk of emigrating each year to obtain estimates of emigration of the foreign-born population who entered the United States within the last ten years and of those who entered more than ten years ago. The average of the rates from the 2000 to 2007, 2000 to 2008, and 2000 to 2009 residuals are applied by period of entry to ACS 2009. The annual estimates are then divided by four to obtain estimates of foreign-born emigration by period of entry from April 2010 through June 2010. The average of the rates from the 2000 to 2008, 2000 to 2009, and 2000 to 2010 residuals are applied to ACS 2010 to obtain estimates of emigration by period of entry from July 2010 through June 2011. The proxy universe we use for foreign-born emigrants who entered the United States within ten years of the estimate year is the foreign-born population in the ACS who entered the United States within ten years of the survey year. The proxy universe we use for foreign-born emigrants who entered the United States more than ten years before the estimate year is the foreign-born population in the ACS who entered the United States more than ten years before the survey year.

Data from the American Community Survey and the Puerto Rico Community Survey (PRCS) allow us to estimate the annual migration flows between the United States and Puerto Rico directly using the question on place of prior residence. People who indicated on the ACS that they lived in Puerto Rico one year ago are considered immigrants. People who indicated on the PRCS that they lived in the United States one year ago are considered emigrants. The proxy universe for the net migration between the United States and Puerto Rico is the population born in Puerto Rico whose year of entry was five or fewer years before the survey year.

The net migration of natives is based on research by Schachter (2008) using data from over 80 countries.[6] This work compared estimates of the U.S. born or U.S. citizen population living overseas measured at two consecutive time periods and used the difference to develop estimates of net native migration. The proxy universe we use for the net native migrant component is the native household population residing in the United States.

We derive the estimate of the net overseas movement of the Armed Forces population from data collected by DMDC. DMDC provides monthly tabulations of military personnel stationed or deployed outside the United States by age, sex, Hispanic origin, and individual branches of service within the Department of Defense. We assume that change in the overseas military population, excluding deaths, indicates movement of personnel in and out of the United States. To derive the estimates of net movement by race, we apply the race and geographic distribution

[6] Schachter, Jason. 2008. "Estimating Native Emigration from the United States," Memorandum dated December 24, delivered to the U.S. Census Bureau.

of the active-duty military population from Census 2000 to DMDC estimates by age, sex, Hispanic origin, and branch of service.

Net Domestic Migration

We estimate net domestic migration separately for two population universes (household and group quarters) and two age groups (0 to 64 years and 65 years and older).

For the 0 to 64 year old household population, we use person-level data on filers and dependents aged 0 to 64 years from Federal income tax returns supplied by the Internal Revenue Service (IRS). We match two years of IRS tax returns and compare the addresses to identify the number of individuals (represented by exemptions) who moved from one county to another between tax filings. Since every U.S. resident may not file or be claimed as an exemption on a tax return, we cannot use these data to directly estimate the number of county-to-county migrants. Instead, we calculate net domestic migration rates by subtracting the number of out-migrant exemptions from the in-migrant exemptions for each county to produce the number of IRS-based net migrant exemptions. Then, we divide the number of IRS-based net migrant exemptions by the sum of non-migrant exemptions and out-migrant exemptions for each county.

$$NDM\ Rate_{0-64} = \left(\frac{In\ Migrants_{0-64} - Out\ Migrants_{0-64}}{Non\ Migrants_{0-64} + Out\ Migrants_{0-64}}\right)$$

For the 65 years and older household population, we use annual Medicare enrollment data for each county from the Centers for Medicare and Medicaid (CMS). As with the IRS data, we know that not all U.S. residents aged 65 and older receive, or are eligible to receive, Medicare benefits. Therefore, we use the year-to-year change in the Medicare enrollment to calculate a domestic net migration rate. We assume that the year-to-year change in enrollment (benchmarked to the resident population on April 1, 2010 to adjust for undercoverage) represents the total change in the aged 65 and older population in each county. We utilize the estimates of deaths and international migration of the aged 65 and older population, as well as the number of individuals turning 65 years old in each period, to identify the amount of population change that, by default, must be due to net domestic migration. We calculate a Medicare-based net migration rate for each county by dividing the net domestic migration estimate by the total number of Medicare enrollees at the beginning of the time period.

$$NDM\ Rate_{65+} = \left(\frac{Change\ in\ Medicare\ Enrollment - (Pop\ turning\ 65 - Deaths_{65+} - NIM_{65+})}{Beginning\ Medicare\ Enrollment_{65+}}\right)$$

To estimate the net domestic migration of the group quarters (GQ) population, we first estimate the annual group quarters population for each county. To do this, we use group quarters population data from two sources: (1) 2010 Census group quarters population by single year of age, sex, race, Hispanic origin, and facility type for each subcounty area (e.g., cities, towns, etc.)

and (2) a time series of total population in individual GQ facilities from the Group Quarters Report (GQR) prepared by the FSCPE members.[7]

From these data sources, we first estimate a time series of annual total GQ population by type beginning at the subcounty level. We separately sum the GQ populations from the 2010 Census and the GQR to the subcounty level by facility type for each estimate date in the time series. We calculate a time series of subcounty GQ population by facility type by adding the year-to-year change indicated by the GQR data to the Census 2010 GQ populations for the same subcounty area by facility type.

Once the total GQ estimates are produced, we sum the subcounty estimates to the county level. We use the 2010 Census GQ population by county and facility type as the proxy universe to estimate the age, sex, race, and Hispanic origin distribution of the total county-level GQ population by facility type.

From these data, we aggregate the GQ population to the national level by age, sex, race, and Hispanic origin for use in the production of national estimates. We also aggregate the GQ population to the county level by age group (0 to 64 years and 65 years and older) for use in the production of state and county total estimates.

Net domestic migration of the GQ population is simply estimated by calculating the annual change between the final GQ population estimates for each county and age group.

Production of Population Estimates

National Population by Age, Sex, Race, and Hispanic Origin

There are three main steps in the production of monthly national population estimates: estimating the quarterly national resident population; estimating the population for the other months of each quarter; and estimating the monthly population for the other four universes.

First, we calculate the births, deaths, and population for one quarter at a time using a cohort component method starting with the base population by age, sex, race, and Hispanic origin as of April 1, 2010. We begin by estimating the population at risk of giving birth or dying within the quarter of interest. Then, we multiply this population by the birth and death rates (by age, sex, race, and Hispanic origin) to estimate the births and deaths that occur in that quarter. Finally, we use those births and deaths, along with the net international migration component, to estimate the population in each birth cohort for the first day of the next quarter. We repeat this process one quarter at a time to the end of the time series (i.e., July 1, 2011).

With final quarterly estimates in hand, the second step in the process is to estimate the population for the other two months of each quarter. We assign the calculated quarterly births

[7] The seven major GQ facility types utilized in estimate production are: correctional institutions, juvenile institutions, nursing homes, other institutional facilities, college dormitories, military housing, and other noninstitutional facilities.

and deaths to specific months within each quarter based on the monthly distribution of births and deaths from 2009. Then, we use these monthly components of change and the balancing equation to estimate the population by age (not cohort), sex, race, and Hispanic origin for August 1, 2010, September 1, 2010, November 1, 2010, December 1, 2010, February 1, 2011, March 1, 2011, May 1, 2011, and June 1, 2011.

The last step in the production of monthly national estimates is to calculate the four additional population universes by demographic characteristics. To calculate the resident plus Armed Forces overseas population, we add the monthly overseas military population, based on data collected by DMDC, to the resident population. To estimate the civilian population, we subtract the monthly resident military population, also estimated from data collected by DMDC, from the resident population. The civilian noninstitutionalized population is produced by subtracting the institutionalized group quarters population from the civilian population.[8] Finally, we estimate the household population by subtracting the total group quarters population from the resident population.

State and County Total Population

To produce annual state and county total population estimates, we begin by producing county-level estimates for two age groups: population aged 0 to 64 and population aged 65 and older.[9] Starting with the county-specific base population by age group as of April 1, 2010, we calculate the population as of July 1, 2010 using a component of change method.

We first estimate the household population at risk of migrating into or out of each county between April 1 and June 30, 2010. Then, we multiply this population by the net domestic migration rates for the household population in each age group to estimate the county-specific net household domestic migration. We use the births, deaths, net international migration, and net domestic migration (of both the household and GQ populations) to estimate the resident population in each age group. Finally, we control the county population estimates by age group to the national resident population estimates for these age groups and repeat this process to estimate the July 1, 2011 resident population for each county.

The total resident population for each county is the sum of the populations in the two age groups. The resident population estimates and components of change for each state are simply the sum of the populations and components of change for its counties.

[8] The institutionalized population is defined as people under formally authorized, supervised care or custody in institutions including correctional institutions, juvenile institutions, nursing homes, skilled nursing facilities, psychiatric hospitals, and facilities for the disabled.

[9] We estimate the county population separately for ages 0 to 64 and ages 65 and older because these groups are consistent with the populations at risk in our two sources of domestic migration data: IRS tax return data and Medicare enrollment data.