

# **Methodology for Constructing Population Projections by Age, Sex, Race, and Hispanic Origin for Florida and Its Counties, 2020–2045, With Estimates for 2016**

Stefan Rayer and Ying Wang  
Bureau of Economic and Business Research  
University of Florida  
June, 2017

The composition of Florida's population has changed substantially in recent decades. Between 1950 and 2010, for example, the proportion of Florida's population younger than age 15 declined from 26.2 to 17.5 percent; the proportion age 65 and older rose from 8.6 to 17.3 percent; and the proportion black declined from 21.7 to 16.9 percent. The Hispanic population increased from 6.0 percent of the total population in 1970 to 22.5 percent in 2010. Changes in demographic composition have been even greater for many counties than for the state as a whole.

These changes have important implications for planning and public policy. They affect the demand for education, healthcare, housing, recreation, transportation, and many other goods and services. They affect the number and characteristics of persons in the labor force and in public and private retirement systems. They affect the allocation of many types of public funds. Consequently, there is a tremendous need for population estimates and projections by age, sex, race, and Hispanic origin. This report provides a description of the methodology used by the Bureau of Economic and Business Research (BEBR) to construct these estimates and projections.

## **Definitions of Race and Ethnicity**

The decennial census in the United States is based on self-enumeration. Residents of each household are asked to provide the responses they believe best describe their demographic characteristics, based on guidelines established by the U.S. Census Bureau and the U.S. Office of Management and Budget (OMB). These guidelines allow respondents to identify themselves as Hispanic or non-Hispanic and as belonging to one or more of several racial groups.

It should be noted that "Hispanic" is an ethnic classification rather than a racial category; that is, people can be identified both by Hispanic origin and by race. The OMB defines Hispanic or Latino as "a person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin regardless of race." For data collection and presentation purposes, federal agencies are required to use a minimum of two ethnicities: "Hispanic or Latino," and "Not Hispanic or Latino." We follow the same guidelines in this report and use the term "Hispanic" to refer to persons of Hispanic, Latino, or Spanish origin.

The three largest racial/ethnic groups in Florida are non-Hispanic whites, non-Hispanic blacks, and Hispanics. These three groups accounted for 97.0 percent of Florida's population in 2010. We made an initial set of estimates and projections by age and sex for these three racial/ethnic groups. Using these estimates and projections as a starting point, we constructed additional sets of estimates and projections for several other race/ethnicity combinations.

## **DATA**

Data from the 1990, 2000, and 2010 censuses formed the basis for these estimates and projections. Although census data are generally quite reliable, two issues regarding race and

ethnicity complicate their use. First, in 1990 and all previous censuses, respondents were required to identify themselves as belonging to a single race. Starting in 2000, they were permitted to identify themselves as belonging to one or more races. In Florida, 97.5 percent of the population identified themselves as belonging to a single race in 2010 and 2.5 percent identified themselves as belonging to two or more races. These proportions are very similar to those reported in 2000 (97.6 percent and 2.4 percent, respectively).

Second, although the U.S. Census Bureau defines “Hispanic origin” as an ethnic classification rather than a racial category, many respondents interpreted it as a racial category and listed their race as Hispanic, Latino, Mexican, Spaniard, or a similar response. In Florida, 3.6 percent of the total population in 2010 classified themselves as belonging to some race other than those listed on the census questionnaire; more than 90 percent of those respondents were of Hispanic origin. In 2000, 4.4 percent of the total population classified themselves as belonging to some race other than those listed on the census questionnaire; again, more than 90 percent of those respondents were of Hispanic origin.

Responding to these issues, the National Center for Health Statistics (NCHS) collaborated with the U.S. Census Bureau to create a set of modified 2000 census counts for every state and county in the United States. Using a variety of data sources and techniques, the NCHS assigned people who classified themselves as belonging to more than one race (or who marked “some other race” on the census questionnaire) to a single primary race. These modifications produced data that were consistent over time, prevented double-counting of people belonging to more than one race, and provided a racial classification for Hispanics who did not identify their race. The NCHS released a comparable set of modified census counts for 2010. For 1990, the U.S. Census Bureau made modifications to the 1990 census counts in which persons listing “some other race” were assigned to a specific race – the modified age/race, sex, and Hispanic origin (MARS) files. The estimates and projections described in this report were based on MARS data for 1990 and NCHS data for 2000 and 2010. For 2010, we used an updated April 1, 2010 population that includes Count Question Resolution (CQR) changes for Lake, Marion, and Miami-Dade counties from the Vintage 2013 NCHS bridged-race postcensal population estimates. We also made an additional adjustment for Gadsden County to correct for the institutional facility that the 2010 census failed to enumerate in the county.

Large institutions (e.g., universities, prisons) account for a significant proportion of the total population in many counties in Florida. In such counties, it is important to account for the impact of these institutions when making population estimates and projections. Consequently, we used institutional records and data from the decennial census to estimate the non-institutional population by age, sex, race, and Hispanic origin for 1990, 2000, 2010, and 2016 in the following counties: Alachua, Baker, Bradford, Calhoun, Columbia, DeSoto, Dixie, Franklin, Gadsden, Gilchrist, Glades, Gulf, Hamilton, Hardee, Hendry, Holmes, Indian River, Jackson, Jefferson, Lafayette, Leon, Liberty, Madison, Okeechobee, Santa Rosa, Sumter, Suwannee, Taylor, Union, Volusia, Wakulla, Walton, and Washington counties. In these counties, we made separate projections for the institutional and non-institutional populations. The final estimates and projections for each county were constructed by adding together the institutional and non-institutional populations. The remainder of this report describes the methodology used for making estimates and projections of the non-institutional population.

## **METHODOLOGY**

### **2016 Estimates of Total Population by Race and Ethnicity**

We made estimates of the total number of non-Hispanic whites, non-Hispanic nonwhites, and Hispanics for 2016 using a variety of data sources and techniques. Some relied on extrapolations of previous population trends, whereas others incorporated data on births, deaths, and school enrollment by race and ethnicity. Some estimates were based on averages of several of the individual techniques. The final estimate for each racial/ethnic group in each county was based on our judgment regarding which technique was most likely to provide an accurate estimate of the non-institutional population. Estimates of total population by race/ethnicity were made by adding estimates of the institutional population to estimates of the non-institutional population. As a final step, estimates for the three racial/ethnic groups were controlled to the 2016 estimates of total population published in “Florida Estimates of Population: April 1, 2016,” Bureau of Economic and Business Research, December 2016. A more detailed description of the methodology can be found in an article by Stanley Smith and June Nogle published in the *Social Science Quarterly* in 2004 (volume 85, pp. 731–745).

### **Projections of Total Population by Race and Ethnicity**

Starting with the 2016 estimates, we made two sets of projections of the total non-institutional population of each county for non-Hispanic whites, non-Hispanic nonwhites, and Hispanics. For the first set, we used the following techniques:

LINE26: linear extrapolation of 1990–2016 non-institutional population change for each racial/ethnic group.

LINE16: linear extrapolation of 2000–2016 non-institutional population change for each racial/ethnic group.

SHARE26: each racial/ethnic group’s share of county non-institutional population change 1990–2016 is applied to projected county non-institutional population change.

SHARE16: each racial/ethnic group’s share of county non-institutional population change 2000–2016 is applied to projected county non-institutional population change.

EXPO26: exponential extrapolation of 1990–2016 non-institutional population change for each racial/ethnic group.

SHIFT26: the change in each racial/ethnic group’s share of county non-institutional population 1990–2016 is linearly extrapolated and applied to county projections of total non-institutional population.

CONST%: each racial/ethnic group’s share of the non-institutional population in 2016 is assumed to remain constant over time.

AVE7: an average of projections from the seven techniques described above.

AVE5: an average of these projections, excluding the highest and lowest.

AVE3: an average of these projections, excluding the two highest and the two lowest.

Based on results from our ongoing projection evaluation research, this year we also created a second set of projections, using the following techniques:

LINE16: linear extrapolation of 2000–2016 non-institutional population change for each racial/ethnic group.

LINE6: linear extrapolation of 2010–2016 non-institutional population change for each racial/ethnic group.

SHARE16: each racial/ethnic group's share of county non-institutional population change 2000–2016 is applied to projected county non-institutional population change.

SHARE6: each racial/ethnic group's share of county non-institutional population change 2010–2016 is applied to projected county non-institutional population change.

EXPO AVE: average of three exponential extrapolations of 1990–2016, 2000–2016, and 2010–2016 non-institutional population change for each racial/ethnic group.

SHIFT AVE: average of the three changes in each racial/ethnic group's share of county non-institutional population 1990–2016, 2000–2016, and 2010–2016, which are linearly extrapolated and applied to county projections of total non-institutional population.

CONST%: each racial/ethnic group's share of the non-institutional population in 2016 is assumed to remain constant over time.

AVE7: an average of projections from the seven techniques described above.

AVE5: an average of these projections, excluding the highest and lowest.

AVE3: an average of these projections, excluding the two highest and the two lowest.

These two sets apply the same projection techniques but differ with respect to the base data used. For the linear and share-of-growth techniques, the second set uses shorter base periods of 6 and 16 years (vs. 16 and 26 years in the first set). Furthermore, for the exponential and shift-share techniques, the second set averages projections made with 6, 16, and 26 year base periods (vs. a single base period of 26 years in the first set). We believe the combination of base period lengths and projection techniques in the second set to be preferable to those of the first set. The second set incorporates more recent data for the linear, share-of-growth, exponential, and shift-share techniques; it also results in more stable projections for the exponential and shift-share techniques by averaging past growth trends over three different base periods.

To provide for greater continuity with our previous county projections by race and ethnicity, we decided to average projections from the first and second set. The final projection of the total population for each racial/ethnic group in each county was based on our judgment regarding which technique was most likely to provide an accurate forecast of the future non-institutional population. In 66 counties, the final projection was based on the average of AVE3 from the two sets, controlled to the medium county projection of total non-institutional population; in Hardee County, it was based on AVE5 from the second set, controlled to the medium county projection of total non-institutional population.

In counties with institutional adjustments, projections of the institutional population were based on institutional records and our judgment regarding future institutional growth. Projections of the racial/ethnic breakdown of the institutional population were made by applying the racial/ethnic distribution from the 2010 census to the projections of the total institutional population, which were

adjusted to reflect changes in the racial/ethnic distribution of the non-institutional population over the projection horizon.

Finally, projections of total population by race/ethnicity were made by adding projections of the institutional population to projections of the non-institutional population. In all counties, projections for the three racial/ethnic groups were controlled to the medium projections published in “Projections of Florida Population by County, 2020–2045, with Estimates for 2016,” Florida Population Studies, Bulletin No. 177, Bureau of Economic and Business Research, April 2017.

### **Projections by Age, Sex, and Race/Ethnicity**

Projections by age and sex for each of the three racial/ethnic groups were made using a cohort-survival rate methodology. Age was calculated in five-year groups from 0–4 to 85+. Projections were made in five-year intervals, starting with the 2015 estimates published last year; each projection served as the base for the following projection.

Using modified census and institutional population data for 2000 and 2010, and intercensal population estimates and institutional population data for 2005, we subtracted the institutional population from the total population for each age, sex, racial, and ethnic group to derive estimates of the non-institutional population in each demographic subgroup. We calculated cohort-survival rates by sex for the non-institutional population by dividing the 2010 modified census count for each age, racial, and ethnic group by the 2005 intercensal population estimate for the corresponding group 5 years younger. We also calculated cohort-survival rates by sex for the non-institutional population by dividing the 2005 intercensal population estimate for each age, racial, and ethnic group by the 2000 modified census count for the corresponding group 5 years younger. From these we calculated an average of 2000–2005 and 2005–2010 cohort-survival rates. We chose an average of those two periods because population growth in the first half of the decade was quite different from population growth in the second half. Averaging has generally been found to increase the accuracy of population projections.

Using cohort-survival rates averaged over 2000–2005 and 2005–2010, we made several additional adjustments. First, we applied weighting factors to account for higher survival rates among the older age groups. For many counties, we further adjusted the resulting cohort-survival rates to account for apparent data errors and to smooth out differences among age groups, or between males and females. These adjustments were most frequent in counties with small populations, especially for the non-Hispanic nonwhite and the Hispanic populations.

We applied the adjusted cohort-survival rates to the 2015 non-institutional population by age, sex, race, and ethnicity to produce projections for 2020 for the population age 5 and older. For the population less than age 5, we used child-woman ratios based on 2010 NCHS data (i.e., population aged 0–4 divided by females aged 15–44). We applied those ratios to the projected female population in 2020 to provide projections of children aged 0–4. The population age 0–4 was divided between males and females using proportions of 0.51 and 0.49, respectively. In some instances, we adjusted the child-woman ratios to account for expected changes in fertility rates. For each of the three racial and ethnic groups, we controlled the non-institutional age and sex projections to the independent projections of the total non-institutional population for 2020.

We repeated the process to produce projections for 2025, 2030, 2035, 2040, and 2045. These projections were controlled to the independent projections of the non-institutional population described above. As a final step, we added the independent projections of the institutional

population, providing projections by age and sex for non-Hispanic whites, non-Hispanic nonwhites, and Hispanics. Projections at the state level were calculated by adding up the county projections.

### **Projections for other Racial/Ethnic Groups**

We developed projections for several additional racial/ethnic groups. Using the 2010 NCHS data, we calculated the white/nonwhite proportion of the Hispanic population for each county and applied those proportions to the Hispanic projections to provide a white/nonwhite breakdown of the Hispanic population (in Florida, approximately 76 percent of the Hispanic population identified themselves as white alone in the 2010 census). Adding the Hispanic white population to the non-Hispanic white population provided projections of the total white population by age and sex for each county.

Using the 2010 NCHS data, we calculated blacks as a proportion of nonwhites for both the Hispanic and non-Hispanic populations. We made those calculations separately for each county and – based on historical trends and the 2010 values – projected those proportions into the future. By applying these proportions to projections of the nonwhite population (for both Hispanics and non-Hispanics), we developed projections of the non-Hispanic black population and the total black population by age and sex for each county.

### **Estimates for 2016**

The methodology described above was used to make projections by age, sex, race, and Hispanic origin in five-year intervals from 2015 to 2045 for each county in Florida. Estimates for 2016 were made by interpolating between the 2015 estimates and the 2020 projections for each age/sex/racial/ethnic group and controlling those interpolations to the 2016 estimates of total population by race and ethnicity described above.

### **ACKNOWLEDGEMENT**

Funding for these estimates and projections was provided by the Florida Legislature.