Demographics & Workforce
Weldon Cooper Center for Public Service
University of Virginia

Virginia Population Projections Methodology

EXECUTIVE SUMMARY

In 2012, under a contract with the Virginia Employment Commission, the Demographics & Workforce Group of the University of Virginia's Weldon Cooper Center for Public Service developed the first set of post-2010 population projections for Virginia and its 134 localities.

The result includes projections for 2020, 2030, and 2040 of the following:

- Total population
- By age, sex, and race
- By age, sex, and Hispanic Origin

In addition –and for the first time – total population projections were produced for Virginia's large towns, those with populations greater than 5,000 in the 2010 census.

The methodologies used for producing these projections are a result of the Cooper Center demographers' extensive and thorough research, careful and thoughtful evaluation of options, and creative and practical applications of demographic and statistical theories. The methodologies selected are simple and elegant, requiring very few assumptions and using only census data. The procedures are straightforward and applicable to all localities. The results prove to be reasonable and highly consistent with historical trends.

1. Projecting total population for localities

(Autoregressive mixed-effects model)

Population data from 1950 – 2010 was used to identify the relationship between the past and future. The model was tested to see how well it would predict the population in 2010 statewide and across all localities. The model performed very well, with a low rate of error. This model was then used to project total population in 2020, 2030, and 2040.

2. Projecting age, sex, and race/ethnicity

(Hamilton-Perry method)

The methodology for producing age, sex, and racial/ethnic composition of the population began with trends observed in the last twenty years. Decade-to-decade change rates for each locality were assumed to remain the same over time and we used locality-specific rates wherever appropriate; this effectively reflected local demographic characteristics, especially for college towns, retirement communities, and military bases.

3. Projecting total population for large towns
(Linear extrapolation)

The optimal methodology for projecting the population of towns larger than 5,000 was to use each town's population growth trends from 1980-2010 as the assumed growth trends for each of the coming decades.

Projections at state and Planning District levels are aggregates from member localities.

The following pages detail the methods used for producing Virginia population projections. Should you have questions or need more information, please contact Dr. Rebecca Tippett at 434-982-5861, or rebecca.tippett@viriginia.edu.

LOCALITY PROJECTIONS

1. **Methodology for Projecting Total Population**

   Future population is modeled as a function of past population. Using Census population data from 1950 through 2010, the total population in each locality is regressed on the prior two decades' population counts. This allows the projection of future population to rely on only the most reliable information at our disposal – past population counts – without imposing a pre-determined relationship between the past and the future (e.g., that population changes linearly or exponentially).

   In addition, each locality's population is modeled simultaneously, but individually, within a multilevel model. The data from the prior six decades of Census counts do not, on their own, provide enough information to reliably estimate each locality's future population. By pooling observations across all of Virginia's localities we can produce much more robust estimates. The multilevel model, though, does not assume a common pattern of population change across localities, but estimates a random effect that accounts for variation in the relationship between past and future population across Virginia's counties and independent cities. A distinct relationship is estimated for each locality.

   Final Model – County-Total Population

$$Y_{it} = \beta_{0i} + \beta_{1i}Y_{it-1} + \beta_{2i}Y_{it-2} + \epsilon_{it}$$
$$\beta_{0i} = \gamma_{00} + u_{0i}$$
$$\beta_{1i} = \gamma_{10} + u_{1i}$$
$$\beta_{2i} = \gamma_{20} + u_{2i}$$

   In this model *Y* equals population, *i* indexes locality, and *t* indexes time. The βs and *γs* are parameters to be estimated, and *ϵ* and *u* are random effects.

   This model performed very well with two exceptions, Loudoun and Prince William counties. Their rapid growth over the past decade(s) was reflected in untenable exponential projected growth in all models considered, requiring expert judgment and manual adjustment of the projected total population. For these two localities, we substituted a linear extrapolation.

   These projections of the total population for 2020, 2030, and 2040 are used as control totals for the distribution of detailed population characteristics.

2. **Methodology for Projecting Age Distribution**

   The Hamilton-Perry method is a reduced form of the cohort-component method that requires less detailed data and captures the major components of population change (births, deaths, and migration) in a general way. Research has shown that the Hamilton-Perry method is an accurate methodology for projecting population characteristics. Birth rates in the prior decade are measured by a *child-population ratio* (CPR) in which the child population (0-4 and 5-9) is divided by the population of child-bearing age. For

example, the CPR for children 0 to 4 is

$$CPR_{0-4} = \frac{Pop\ {}^{2010}_{0-4}}{Pop^{2010}_{15-44}}$$

Deaths and migration are jointly captured in *cohort-change ratios* (CCR), that is, the ratio of total population in age group *a+10* in the launch year (2010) divided by the total population of age group *a* in the base year (2000)

$$CCR_a = \frac{Pop\ {}^{2010}_{a+10}}{Pop^{2000}_{a}}$$

For example, the CCR for the 25 to 29 year-old age group would be calculated as

$$CCR_{25-29} = \frac{Pop\ {}^{2010}_{35-39}}{Pop^{2000}_{25-29}}$$

CCRs greater than 1 indicate population growth due to net migration that outweighs deaths, whereas CCRs less than 1 indicate population loss due to deaths, out-migration, or a combination of the two. The population is projected forward by multiplying the current population at age *a* by the CCR for that age group. The young population (ages 0 to 9) is then projected by multiplying the CPR by the projected population of childbearing age.

For the 2020-2040 projections by age, the following procedure was developed:

a. CCRs and CPRs were calculated for each locality for 1990-2000 and 2000-2010 and averaged together. This helps smooth decade-to-decade variability that might result in unreasonable future projections if projected forward 30 years. This is also consistent with the two-lag autoregressive model used to project total population.
b. The averaged CCRs and CPRs were applied to the 2010 base population to project 2020, to the projected 2020 population to project 2030, and to the projected 2030 population to project 2040.
c. The age projections are controlled to the locality total population projection.

3. Methodology for Projecting Race and Ethnicity

We first considered using locality-specific age-race/ethnicity CCRs. However, preliminary analysis of detailed population projections revealed substantial shifts toward the end of the projection horizon that were inconsistent with historical trends. This was largely due to high cohort-change ratios for Asian, other, and Hispanic groups, compared to more stable CCRs for white, black, and non-Hispanic Virginians. The high CCRs for these groups reflected high immigration among an initially smaller population; changes in the way that individuals are able to identify themselves on the Census form; social changes that increase the prevalence of multi-racial individuals; and increasing individual willingness to identify and claim “other race” on Census forms.

To address these issues, we developed an approach to implementing the Hamilton-Perry method to project the proportion of each racial and ethnic group within each age group:

1) Used state-level age-race/ethnicity-specific CCRs (based on 2000 to 2010) and CPRs (held constant to 2010) for each locality.

2) Distributed race within each age category in two ways:

a. *If the sum of the four race projections at age a was **less than or equal to** the total population projected for age a,* the population proportion of race *r* at age *a* for the locality was determined by

$$\frac{\boldsymbol{r_a}}{(white_a + \ black_a + Asian_a + other_a)}$$

b. *If the sum of the four race projections at age a was **greater than** the total population projected at age a*, the population proportion of white and black residents in a given locality was determined by

$$\frac{\boldsymbol{r_a}}{Total\ Pop_a}$$

The population proportion of Asian and other race residents was determined by

$$\frac{(Total\ Pop_{\ a} - \ White_a - Black_a) \times \frac{\boldsymbol{r_a}}{(Asian_a + other_a)}}{Total\ Pop_a}$$

Distributed ethnicity within each age category in two ways:

c. *If the sum of non-Hispanic and Hispanic projections was **less than or equal to** the total population projected for age a*, the population proportion of ethnic group *e* at age *a* for the locality was determined by

$$\frac{\boldsymbol{e_a}}{(nonHispanic_a + \ Hispanic_a)}$$

d. *If the sum of the non-Hispanic and Hispanic projections was **greater than** the total population projected at age a*, the population proportion of non-Hispanic residents in a given locality was determined by

$$\frac{nonHispanic_{\boldsymbol{a}}}{Total\ Pop_a}$$

The population proportion of Hispanic residents was determined by

$$\frac{(Total\ Pop_{\ a} - \ nonHispanic_a)}{(Total\ Pop_a)}$$

e. In six age/race categories ("other race" at ages 80 and 85 in 2020 and 2030, and Asian at age 80 (2020) and age 85 (2030)), this method resulted in negative population projections for selected localities, due to quite small initial populations of Asian and other race individuals at younger ages in 2010. In these cases, the negative projection was set to

zero and the population proportion of the remaining racial groups was re-calculated (using the sum of the remaining projected racial categories as the basis for the ratio).

3) Raked the race/ethnicity proportion to the locality-specific age distribution developed in step 2.

4. Methodology for Projecting Sex

Our methodology holds the sex ratio constant to the total population's age-specific sex ratio within each locality. Since sex ratios are historically stable, this enables localities with unique sex structures (prisons, military barracks, etc.) to retain their local characteristics.

$$Prop_{i-a-s} = \frac{N_{i-a-s}^{2010}}{N_{i-a}^{2010}}$$

In the above equation, *i* indexes locality, *a* indexes age group, and *s* indexes sex.

5. Calculating Detailed Characteristics

The final distribution of detailed population characteristics is calculated by the following:

$$N_{ti-a-r-s} = Y_{ti} \times Prop_{i-a} \times Prop_{i-a-r} \times Prop_{i-a-s}$$

Where the number of individuals in a given age group *a* of race or ethnic group *r* and sex *s* living in locality *i* at time *t* are equal to the projected total population of locality Y (from AR2 model, step (1)), times the proportion within that age group (step (2), HP model), the proportion of the race or ethnic group within that age group (step (3), HP model), and the proportion of the sex within that age group (step (4), 2010 ratio).

State Projections

Projections for Virginia are equal to the sum of the projections for each of the 134 localities.

Planning District Commission Projections

Planning District Commission (PDC) projections are equal to the sum of their member localities.

Town Projections

Total population for 2020, 2030, and 2040 was projected for towns with populations of 5,000 or more in 2010. Projected total population is calculated by linear extrapolation:

$$Pop_t = Pop_{2010} + \left[\frac{(Pop_{2010} - Pop_{1980})}{30} \times (t - 2010)\right]$$