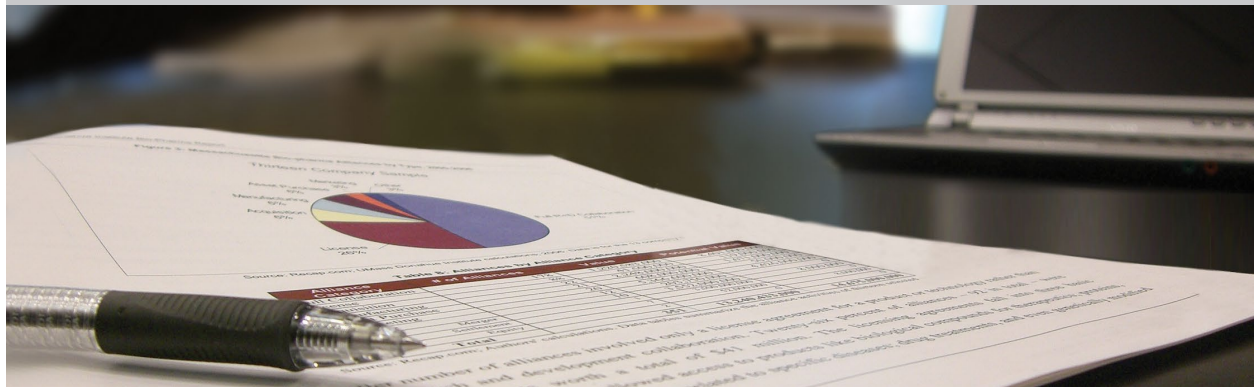


# Small Area Population Estimates for 2021 through 2030

Prepared for the Massachusetts Department of Public  
Health

May 8, 2025



UMassAmherst

Donahue Institute  
Economic and  
Public Policy Research

# Small Area Population Estimates for 2021 through 2030: Prepared for the Massachusetts Department of Public Health

Prepared by the UMass Donahue Institute's  
Economic & Public Policy Research Group

---

**Project Leader**

Susan Strate, Senior Manager,  
Population Estimates Program

**Research Assistants**

Aisling Donoghue  
Maxwell Williams

**Project Staff**

Meghan Flanagan, Senior Research Analyst  
Thomas Peake, Senior Research Analyst  
Jacob Harrington, Research Analyst II  
Denis McAuliffe, Research Analyst II

**Unit Director**

Mark Melnik, Director of Economic  
& Public Policy Research

---

Established in 1971, the UMass Donahue Institute is a public service, research, and economic development arm of the University of Massachusetts. Our mission is advancing equity and social justice, fostering healthy communities, and supporting inclusive economies that alleviate poverty and promote opportunity. We serve clients in the public, non-profit, and private sectors in the Commonwealth and throughout the nation and the world. For more information, [www.donahue.umass.edu](http://www.donahue.umass.edu).

The Institute's Economic & Public Policy Research (EPPR) group is a leading provider of applied research, helping clients make more informed decisions about strategic economic and public policy issues.

EPPR produces in-depth economic impact and industry studies that help clients build credibility, gain visibility, educate constituents, and plan economic development initiatives. EPPR is known for providing unbiased economic analysis on state-level economic policy issues in Massachusetts and beyond and has completed a number of industry studies on IT, defense industries, telecommunications, health care, and transportation. Their trademark publication is called *MassBenchmarks*, an economic journal that presents timely information concerning the performance of and prospects for the Massachusetts economy, including economic analyses of key industries that make up the economic base of the state.

---

UMassAmherst

Donahue Institute  
Economic and  
Public Policy Research

## Table of Contents

---

|  |    |
|--|----|
| List of Figures.....   | 5  |
| Acknowledgments.....   | 6  |
| Background.....  | 7  |
| Method Overview.....   | 9  |
| Note on Assumptions and Special Cases .....  | 11 |
| Detailed Methodology.....  | 13 |
| 1) Obtaining, structuring, and normalizing 2000, 2010, and 2020 census data.....   | 13 |
| 2) Reassigning “Some other race” .....   | 16 |
| 3) Adjusting Census population counts .....  | 18 |
| 4) Special adjustments: college tract overrides.....   | 20 |
| 5) Control race/ethnicity/age/sex launch to age/sex launch .....   | 21 |
| 6) Calculating cohort-change ratios and child-to-woman ratios .....  | 21 |
| Calculating Ratios.....  | 21 |
| Capping Ratios .....   | 22 |
| Alternative Race Ratios Testing .....  | 24 |
| 7) Applying ratios and caps to the modified Census 2020 base population to create<br>“uncontrolled” 2030 estimates ..... | 25 |
| 8) Distribute age/sex estimates to single years 2021-2030.....   | 26 |
| 9) Control the 2021 to 2030 estimates by tract and by town to the county-level annual<br>post-census estimates.....      | 26 |
| 10) Distributing 5-year age-group estimates and errors to single years of age for ages 0-<br>20.....                     | 27 |
| 11) Controlling race and ethnicity projections to age and sex projections.....   | 28 |
| 12) Assigning error and confidence intervals .....   | 28 |
| Mean Absolute Percent Errors.....  | 29 |
| Mean Percent Errors.....   | 29 |
| Error Calculations.....  | 30 |
| 13) Distributing error and confidence intervals over single years 2021-2030 .....  | 32 |
| 14) Tract-level “college flag” .....   | 33 |
| 15) Summing to total race, ethnicity, and sex.....   | 33 |
| Projections Output .....   | 35 |
| Appendix A .....   | 37 |
| Works Cited .....  | 51 |

## List of Tables

---

|  |           |
|--|-----------|
| <b>Table 1. Population Adjustment by Municipality and Challenge Type .....</b>                                       | <b>19</b> |
| Table 2. College Override Tracts and Populations .....   | 21        |
| Table 3. Single year of age as percent of 5-year age group and 10-year percentage point shift, 1990-2020 .....       | 27        |
| Table 4. Population Estimates Output Variable Names and Descriptions.....  | 35        |
| Table A1. Age/Sex Share of Census 2020 Massachusetts Group Quarters Population by Type .....                         | 37        |
| Table A2. Cohort-size categories used in error assignment .....  | 38        |
| Table A3. Mean Absolute Percent Errors by Age Group and Cohort Size, MCD-Level, All Massachusetts Geographies.....   | 39        |
| Table A4. Mean Absolute Percent Errors by Age Group and Cohort Size, MCD-Level, Cape and Island Geographies.....     | 40        |
| Table A5. Mean Absolute Percent Errors by Age Group and Cohort Size, MCD-Level, Balance of State ...                 | 41        |
| Table A6. Mean Absolute Percent Errors by Age Group and Cohort Size, Tract Level, All Massachusetts Geographies..... | 42        |
| Table A7. Mean Absolute Percent Errors by Age Group and Cohort Size, Tract-Level, Cape and Island Geographies.....   | 43        |
| Table A8. Mean Absolute Percent Errors by Age Group and Cohort Size, Tract-Level, Balance of State...                | 44        |
| Table A9. Mean Percent Errors by Age Group and Cohort Size, MCD-Level, All Massachusetts Geographies.....            | 45        |
| Table A10. Mean Percent Errors by Age Group and Cohort Size, MCD-Level, Cape and Island Geographies .....            | 46        |
| Table A11. Mean Percent Errors by Age Group and Cohort Size, MCD-Level, Balance of State .....                       | 47        |
| Table A12. Mean Percent Errors by Age Group and Cohort Size, Tract Level, All Massachusetts Geographies.....         | 48        |
| Table A13. Mean Percent Errors by Age Group and Cohort Size, Tract-Level, Cape and Island Geographies .....          | 49        |
| Table A14. Mean Percent Errors by Age Group and Cohort Size, Tract-Level, Balance of State .....                     | 50        |

## List of Figures

---

|  |    |
|--|----|
| Figure 1. Example of Census 2000 Block Cohort Population Distributed to Census 2010 Block..... | 15 |
| Figure 2. Example of Summing Distributed 2000 Population to Census 2010 Tracts.....            | 16 |

## Acknowledgments

---

The Small Area Population Estimates research team at UMass Donahue would like to express our gratitude to Chief Demographer Howard Hogan, Statistical Demographer Larry Sink, and Demographer Esther Miller, all of the U.S. Census Bureau's Population Division, for taking the time to talk with our team and to advise us on the technical background and uses of Census products including the modified race summary files and post-censal county-level estimates data as we developed the first iteration of this method in 2016.

We also wish to thank Dr. Henry Renski, Professor and Graduate Program Director of Landscape Architecture and Regional Planning at the University of Massachusetts Amherst for his contributions to the first iteration of this series in 2016.

We also wish to acknowledge the Minnesota Population Center at the University of Minnesota for making publicly available the U.S. Census time-series data used in our estimates through its National Historical Geographic Information System: Version 18.0. portal.

Finally, we wish to thank our esteemed colleague Dr. Stefan Rayer, Director of the Population Program in the Bureau of Economic and Business Research at the University of Florida, for his ongoing willingness to share his experience in and insight into population estimates and projections methods for Massachusetts.

Support for this project was provided by the Massachusetts Department of Public Health.

## Background

---

The Massachusetts Department of Public Health (MA DPH) publishes disease rates and other health-related incidence rates for specific populations at small-level geographies, including the city, town, and even neighborhood level. To calculate these rates, small-level geography population data, including population by age, sex, race, ethnicity, and geography, are needed on an annual basis to serve as denominators.<sup>1</sup> While the Census Bureau enumerates population by age, sex, race, and ethnicity at a very fine level of geography—the census block—it does so only every ten years. In the years following the census, or “post-censal” years, the Census Bureau's Population Estimates Program (PEP) produces annual estimates of the population down to the town level; however, they break out age, sex, race, and ethnicity only at the county level and higher. Researchers needing detailed population estimates for any sub-county geography must find another source.

Populations at small levels of geography and small sub-populations, such as specific age/sex cohorts, are notoriously difficult to estimate accurately. Baker et al., 2013 note:

As in previous studies, the overall magnitude of errors for tract-level... estimates is much larger than observed for larger geographic groupings such as counties or states; moreover, the large differences observed between mean and median percentage errors for both absolute and algebraic values speaks to the ubiquitous presence of outlying values in sets of demographic estimates for small areas. (932-936)

They report mean absolute percent errors of 83% to 92% overall for male and female 5-year cohorts, respectively, in a 2010 tract-level estimates test.

Relying on these potentially inaccurate small area estimates to produce specific incidence rates will likely produce unreasonable rates and unrealistic findings. Given this knowledge, epidemiologists need to know not only the estimated population for a given place and time, but also the estimated error associated with the estimate. Considering both the estimate and its full range of values—including its error, or confidence intervals—allows researchers to assess where a rate may be significant or reliable and where it will not.

To meet the MA DPH project needs, UMDI has produced several vintages of small area population estimates by age, sex, race, and ethnicity at both the town and census tract levels over the years. The first vintage was produced in 2016 for the years 2011 through 2019.<sup>2</sup> In 2022, UMDI developed an updated series that extended the municipal-level estimates to 2020 by incorporating new Census 2020 population counts of the population under 18 and 18-plus.<sup>3</sup> In May of 2023, the U.S. Census Bureau

---

<sup>1</sup> “Ethnicity” in this report refers to Hispanic or non-Hispanic origin.

<sup>2</sup> Tract-level estimates conformed to Census 2010 tract boundaries. Estimates for years 2011 through 2015 were controlled to the latest, June 2016 release of the U.S. Census Bureau's *Annual Estimates of the Resident Population by Sex, Race, and Hispanic Origin Population Estimates for Massachusetts Counties, 2010-201*. Estimates 2016-2020 were uncontrolled, to be later controlled to county estimates to be released by the Census Bureau in future years.

<sup>3</sup> *UMDI Interim 2020 Population Estimates by Age, Sex, Race, and Municipality*, UMass Donahue Institute, March 16, 2022.

---

released 2020 population counts by sex and by 5-year age groups, allowing for a full refresh of the estimates in the V2024 series described in our June 2024 report.<sup>4</sup>

The current report describes methods used to develop UMDI's V2025 series. The series starts with the V2024 age/sex series, which includes male and female age cohorts in 5-year age groups from 0-4 through 80-84 plus an 85+ group as well as estimates by single year of age for ages 0-20. To these estimates, the V2025 series adds race and ethnicity dimensions, resulting in estimates by age, sex, race, and ethnicity for all Massachusetts municipalities and Census 2020 tracts. Ethnic groups in the series include the Census defined ethnicity groups Hispanic ("H") and Not Hispanic ("NH"), as well as a category for combined Hispanic and Non-Hispanic group ("ALL"). Race groups include the Census defined major types: American Indian or Alaska Native ("AIAN"), Asian ("A"), Black or African American ("B"), Native Hawaiian or Other Pacific Islander ("NHPI"), White ("W"), Two or More Races ("TOM"), and, again, a category for combined race groups ("ALL"). Additionally, the series includes error ranges by age and cohort size, which are incorporated into the estimates as low-to-high values for each cohort estimate, making each a range-of-values rather than a single-value estimate. To distribute the "Some other race" response values that appear in the decennial Census data, the V2025 age/sex/race/ethnicity series applies the 2010 version of the U.S. Census Bureau Modified Summary Race File to the 2010 Census count data<sup>5</sup>, and the Census 2020 Modified Age and Race Census file (MARC) to the 2020 Census count data.<sup>6</sup>

---

<sup>4</sup> *Small Area Population Estimates for 2021 through 2030*, UMass Donahue Institute, June 26, 2024.

<sup>5</sup> 2010 File: 4/1/2010 Modified Race Characteristics Source: U.S. Census Bureau. Released: July 2010.  
[https://www.census.gov/popest/research/modified/STCO-MR2010\\_AL\\_MO.csv](https://www.census.gov/popest/research/modified/STCO-MR2010_AL_MO.csv)

<sup>6</sup> 2020 File: Modified Age and Race Census (MARC) file: Source: U.S. Census Bureau. Released: March 6, 2025.  
<https://www.census.gov/data/datasets/2020/demo/popest/modified-race-data-2020.html>

---



## Method Overview

---

To produce small area population estimates for the purposes of this project, UMDI uses a modified *Hamilton-Perry* method and data from the 2010 and 2020 decennial Census counts conducted by the U.S. Census Bureau. The Hamilton-Perry model is a simplified version of the cohort-component method of estimating population change. In general, a cohort-component model considers basic components of population change, including births, deaths, and migration for a specific cohort and geography. It takes the starting, or base, population for the cohort, ages it forward, and adds or subtracts births, deaths, and migration to create a future estimate. These components of change are often expressed as cohort-specific rates—for example, birth rates by age and race of mother, death rates by age and sex, etc.—which are then applied back to the appropriate cohort population in the base year to model the future population. Rates can be held constant, increased, or decreased depending on the assumptions that a particular model makes.

Smith, Tayman, and Swanson explain:

Hamilton and Perry (1962) proposed the use of cohort-change ratios as a short-cut way to apply the cohort-component method... The major difference [in the Hamilton-Perry model] is that it treats mortality and migration as a single unit rather than separately. In addition, the fertility component is often simplified by using child-woman ratios rather than ASBRs [age-specific birth rates]... The Hamilton-Perry method projects population by age and sex using cohort-change ratios (CCRs) based on data from two consecutive censuses. (176-177)

More specifically, the model creates a ratio between a specific cohort population (by age, sex, geography, and sometimes additional variations such as race) for age  $a$  in year  $y$  to its corresponding cohort age  $a-10$  in year  $y-10$ . The ratio is then applied to the corresponding base population to create a forecast of the cohort population ten years later. In this way, each individual group is aged up by ten years over a ten-year period and is also increased or decreased by the effects of mortality and migration experienced over the same period.

Because it relies solely on decennial census count data, which is available at a very detailed level, the Hamilton-Perry method is particularly well suited for estimating small geographies and specific sub-populations. Smith, Tayman, and Swanson explain:

Large data requirements preclude the use of some forms of the cohort-component method at some levels of geography. Although seldom a problem for states and large counties, the lack of data presents a formidable challenge for small counties and subcounty areas. Birth and death data are routinely available for counties but not for most subcounty areas. Migration data are an even greater problem. ACS [American Community Survey] migration data are quite limited, especially at the subcounty level. Although PUMS [Public Use Microdata Sample] files provide detailed migration information, they are often based on a small sample size and are available only for places with at least 100,000 residents. IRS migration data are not tabulated below the county level and do not provide breakdowns of demographic characteristics. Because of these data problems, the Hamilton-Perry method is often the best cohort-component model to use for subcounty projections. (181-182)

While in Massachusetts birth and death data by age and sex are available at the sub-county level, they are not available by tract. Additionally, direct measures of migration are not available by any means at the sub-county level in Massachusetts. Researchers who have tested the use of vital statistics data that has been geocoded by street address have found that method to be unreliable (Baker et al., 2012).

While the Hamilton-Perry method assumes that the cohort-change ratios observed between recent censuses will carry forward into the future, the UMDI model modifies this assumption by incorporating current, and sometimes unanticipated, changes in population by age, sex, race, and ethnicity. In our method, we control the town- and tract-level estimates produced by the cohort-change ratio method to the most recent annual county-level estimates of age/sex/race/ethnicity populations produced by the U.S. Census Bureau.<sup>7</sup> In this way, if a particular race group begins to grow in an area at a rate higher than what past censuses would have predicted, or if migration within a particular age group suddenly changes in a region, this change will be picked up by the estimates and distributed to the sub-county geographies. While the UMDI model still assumes that the *sub-county* distribution of population by age/sex/race/etc. is the same as that indicated by the Hamilton-Perry forecast, the total populations by age, sex, race, and ethnicity are updated for each county.<sup>8</sup>

For their annual county-level estimates, the Census Bureau uses administrative records data to capture birth and death counts at the county level and uses IRS, Medicare enrollment, and American Community Survey data to capture recent county-to-county and international migration. While these data can lag by up to two years, they are still effective in capturing trends that have emerged since the last census count. For instance, they can reflect increased migration to and from specific counties or changes in overall birth rates since the last census. In assessing the accuracy of their post-censal county estimates, the Census Bureau reported a high level of accuracy in accounting for population change between 2010 and 2020, reporting an average mean absolute difference of 2.9 percent between the population estimates and the actual Census 2020 count across all counties.<sup>9</sup> The Census Bureau now also uses a records-matching method to assign characteristics such as race and age to their annual estimates, matching IRS records to Census household responses wherever possible to capitalize further on the integration of current administrative data and to improve the accuracy of their estimates.

While the Hamilton-Perry method is considered a standard model for population forecasting, for the purposes of this product UMDI tested and applied various modifications to maximize performance for Massachusetts geographies when developing the methodology in 2016. Our primary tests included analyzing 10-year versus 20-year cohort-change ratios (CCRs); customized child-to-woman-ratios; and the application of CCR caps. In the testing process, UMDI applied each model variation to create 2010 population estimates and compared them to Census 2010 population counts.<sup>10</sup> The differences between the estimated and actual 2010 populations-by-cohort were recorded as errors, mean absolute percent

---

<sup>7</sup> U.S. Census Bureau's *Annual Estimates of the Resident Population by Sex, Race, and Hispanic Origin Population Estimates for Massachusetts Counties*.

<sup>8</sup> Estimates by race and ethnicity are scheduled for production in a subsequent, V2025 series, after the U.S. Census Bureau releases an updated version of its Modified Summary Race File to correspond with the 2020 Census count

<sup>9</sup> Methodology for the United States Population Estimates: Vintage 2023. *Nation, States, Counties, and Puerto Rico – April 1, 2020 to July 1, 2023*. <https://www2.census.gov/programs-surveys/popest/technical-documentation/methodology/2020-2023/methods-statement-v2023.pdf>

<sup>10</sup> This testing method assumes that future estimates errors for each method will be the same as historic estimates errors for each model variation, however past performance is our best evaluation tool given the unavailability of current, precise cohort-level data.

errors (MAPEs), and weighted absolute percent errors (WAPEs). Lower MAPEs or WAPEs indicate a smaller average absolute difference between the resulting 2010 estimate and the actual 2010 Census data; therefore, model variations with lower MAPEs or WAPEs were considered the better performers. As a result of our model testing, we made the following three decisions for our final model: 1) to use a 10-year cohort-change ratio over a 20-year; 2) to use a modified child-to-woman ratio for estimating children aged 0-9; and 3) to use capped CCR values for small cohort groups. More detailed information on the tests performed in the development of this model is available in the Vintage 2016 methodology report.<sup>11</sup>

## Note on Assumptions and Special Cases

It is important to note that the Hamilton-Perry method, like all estimation methods, relies on particular assumptions; the prime assumption being that future population trends will resemble trends observed in the past. UMDI's modifications to the Hamilton-Perry method also include a number of other assumptions that are described in the methods section of this report. These assumptions should be carefully considered by researchers using or evaluating the resulting population estimates. Where feasible, UMDI tested which of the various candidate models and assumptions performed best in predicting historic population counts. However, even these analyses rely on the assumption that the performance of the model in the current decade will be like that of the past decade.

It is also critical to note that while we attempt to reduce the effect of outlier events in our model, and account for them in our confidence intervals, when working with small-level geographies there are inevitably a number of special or extreme cases of population change that can and do occur. Some of these are past events that are picked up and perpetuated by our model, and some are current events over which our model has no predictive power.

For example, in the *Detailed Methodology* section of this report, we describe a special adjustment we make for the Town of Lincoln, which was subject to a Census 2010 count that did not reflect the true population trend in that geography. While Lincoln is one special case about which we have direct and detailed knowledge, it should be noted that there may exist other special cases around the state for which our model does not account. Especially at the tract level, an off-trend change occurring from one census to the next can significantly exaggerate population in one direction or another. We attempt to minimize the effect of special cases by adding CCR caps, as described in the *Detailed Methodology* section of this report. We also capture the effect of the unusual situations that have occurred in the past in the error ranges and confidence intervals assigned to our estimates, to the extent that the present decade is similar to the last in terms of numbers and magnitude of special cases. Nonetheless it is worth noting that there exist a number of typical "special case" geographies that researchers should be sensitive to when evaluating the accuracy of estimates at a refined level. These include:

- Geographies in which student housing or enrollment changes significantly from one census to the next or since 2010. For example, an all-male college or dormitory becomes co-ed; a large new residence hall is built; or a graduate studies program is added to an undergraduate campus.

---

<sup>11</sup> *Small Area Population Estimates for 2011 through 2020. Prepared for the Massachusetts Department of Public Health's Bureau of Environmental Health. UMass Donahue Institute, October 2016.*

- Geographies in which an assisted living or nursing home residence or some other large group quarters facilities is opened or closed from one census to the next or post-Census 2020.<sup>12</sup>
- Geographies that have undergone major new construction or demolition of residential housing that is out-of-trend with the usual historical or regional (county) construction.
- Geographies that include concentrations of shifting seasonal or international workers, such as the Cape and Islands region.
- Geographies for which post-census count corrections have been made. In these cases, the Census Bureau approved corrections to the total population but made no revisions to the detailed Summary File data used in our estimates model, which is required for age/sex distributions.
- Tracts which have experienced boundary changes from one Census to the next are subject to error inherent in the assumptions made when cross-walking population into new tracts based on land area.
- Cohorts affected by the disclosure avoidance system imposed by the U.S. Census Bureau on all Census 2020 population counts to protect privacy, which may particularly perturb the accuracy of smaller cohorts and smaller geographies.<sup>13</sup>

In the V2025 series, UMDI has attempted to correct for special cases that we discovered in the course of our Census review activities, including the closure of Mt. Ida college, the downsizing of Hampshire College, the addition of new dormitories at UMass Boston, undercounted college dormitory residents in Boston in the 2020 Census count, and a number of smaller group quarters populations and housing units we identified as undercounted in the 2020 Census. Nonetheless, these probably do not represent all of the possible special cases in the state and do not capture additional cases that will emerge between 2020 and 2030.

Additionally, we make certain assumptions about normalizing populations within geographies that change over time. Due to the fact that tract geographies are redrawn at each decennial Census, in order for UMDI to project at a tract level, all launch populations must be redistributed to a common geography for the development of CCRs and error ranges. NHGIS geographical crosswalks provide ratios which describe how total population is redistributed from one decade's geography to another. For our model, we assume that all demographics (age/sex/race/ethnicity) are uniformly distributed within each tract and apply these general ratios to individual cohorts. See Section 1 in "Detailed Methodology: Obtaining, structuring, and normalizing 2000, 2010, and 2020 census data" for more detail on this process and its assumptions.

---

<sup>12</sup> UMDI tracks changes in large group quarters facilities around the state on an annual basis. DPH researchers may consult with UMDI to determine whether a specific geography in their analysis may be experiencing facility-related changes that affect the accuracy of their estimates.

<sup>13</sup> For more information on the U.S. Census Bureau's Census 2020 privacy protection applications, see the UMDI report at: <https://donahue.umass.edu/our-publications/donahue-data-dash-the-effects-of-differential-privacy-on-massachusetts-pl-94-census-data> or the U.S. Census Bureau's website at: <https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/differential-privacy.html>

## Detailed Methodology

---

As described in the *Method Overview* section of this report, the UMDI *Small Area Population Estimates for 2021 through 2030* are the product of a modified Hamilton-Perry model that uses Census 2010-2020 10-year cohort-change-ratios (CCRs) with customized child-to-woman ratios and CCR limits or “caps.”

The implementation of both our age/sex model and age/sex/race/ethnicity model include similar processing steps, with the age/sex/race/ethnicity model requiring a few additional steps. The sequence of the steps are as follows:

1. Obtaining, structuring, and normalizing 2000, 2010, and 2020 census data
2. Reassigning “Some other race” to the six defined races (race/ethnicity model only)
3. Adjusting Census population counts to reflect missed housing units and populations identified by UMDI in post-Census 2020 review and research
4. Making special adjustments in select “college tracts”
5. Controlling age/sex/ethnicity/race launch populations to age/sex projections (race/ethnicity model only)
6. Calculating CCRs and child-to-woman ratios
7. Applying ratios and capped ratios to the modified Census 2020 base population to create “uncontrolled” 2030 estimates
8. Distributing estimates over single years from 2021 to 2030
9. Controlling sub-county estimates to current county-level estimates by age/sex or age/sex/race/ethnicity released by the Census Bureau for years 2021 through 2023
10. Distributing 5-year age group estimates to single years of age for ages 0-20
11. Controlling age/sex/race/ethnicity projections to age/sex projections (race/ethnicity model only)
12. Assigning historic error and confidence intervals
13. Distributing error and confidence intervals over single years 2021-2030
14. Assigning tract-level “college flag”
15. Summing estimates by age, race, and ethnicity to “total” race and ethnicity categories and summing male and female cohorts by age to total by age.

The above production methods themselves involve a number of detailed processing steps and assumptions, which we describe in more detail below.

### 1) Obtaining, structuring, and normalizing 2000, 2010, and 2020 census data

While the geographies and categories counted in the decennial census can change over time, our 2030 estimates model requires data from the 2000, 2010, and 2020 Censuses to be normalized to Census 2020 geographies. Although city and town boundaries in Massachusetts have not changed during the time period used in our model, census tract boundaries are revised every ten years. Therefore, a first step in our production method is to normalize census data from 2000 through 2020. This step is necessary because for each 2020 cohort population, we create a CCR based on a corresponding 2010

cohort population. In order to calculate and apply the proper ratio, the two associated groups must correspond precisely in terms of their geography, age groups, and sex categories.

Corresponding 2000 Census data is also required in our method for the purpose of creating historic errors associated with each cohort estimate. To create error ranges, we run the estimates model to create 2020 age/sex estimates for comparison against 2020 Census counts. 2000 and 2010 data are used to develop CCRs that are applied to 2010 cohorts to develop the 2020 projections used in the comparison. This process requires data from the 2000 Census normalized to 2010 and 2020 census data; that is, the data must be standardized across geographic boundaries. The difference between the estimates and the actual counts forms the basis of our confidence interval assignments.

To normalize the Census datasets for age and sex projections, UMDI downloaded U.S. Census Summary File 1 (SF1) data for 2000 and 2010 and Census 2020 DHC-A from the National Historical Geographic Information System (NHGIS), which provides online access to aggregate census data and GIS-compatible boundary files for United States geographies.<sup>14</sup> Because population data at the tract level had not been normalized by NHGIS at the time of this production cycle, UMDI normalized the time series data using block-level population data in combination with block-to-block correspondence files also available from NHGIS.<sup>15</sup> These correspondence files included the NHGIS 2000 to 2010, 2010 to 2020 and 2020 to 2010 crosswalks, as follows:<sup>16</sup>

1. 2000 → 2010 (nhgis\_blk2000\_blk2010\_gj.csv)
2. 2010 → 2020 (nhgis\_blk2010\_blk2020\_gj.csv)
3. 2020 → 2010 (nhgis\_blk2020\_blk2010\_gj.csv)

*As explained by NHGIS,*

“In a block-to-block crosswalk, each record identifies a possible intersection between a single source block and a single target block, along with an interpolation weight (ranging between 0 and 1) identifying **approximately** what portion of the source zone's population and housing units were located in the intersection. These weights can be used to estimate how **any counts** available for source blocks (e.g., females aged 75 and over, single-member households, owner-occupied housing units, etc.) are distributed among target blocks.”<sup>17</sup>

*NHGIS explains their method for developing their crosswalks as follows:*

“The interpolation weights in NHGIS block crosswalks are primarily based on ‘target-density weighting’ (TDW) (Schroeder 2007). TDW assumes that characteristics within each source zone have a distribution proportional to the densities of another characteristic among target zones. For example, if a 2020 block intersects two 2010 blocks, one of which was 10 times as dense as the other in 2010, then TDW assumes that the same 10:1 ratio holds within the 2020 block in 2020. The interpolation weights in the crosswalks

---

<sup>14</sup> Minnesota Population Center. IPUMS National Historical Geographic Information System: Version 18.0 [dataset]. Minneapolis, MN: IPUMS. 2023. <http://doi.org/10.18128/D050.V18.0>.

<sup>15</sup> Available at: <https://www.nhgis.org/geographic-crosswalks>

<sup>16</sup> For more info, see <https://www.nhgis.org/user-resources/geographic-crosswalks>

<sup>17</sup> Geographic Crosswalks, Minnesota Population Center. IPUMS National Historical Geographic Information System. Accessed at <https://www.nhgis.org/geographic-crosswalks#basics> on June 21, 2024.

from 1990 and 2000 blocks to 2010 blocks involve some more advanced modeling as documented in these pages:

- [2000 Block Data Standardized to 2010 Geography](#)
- [1990 Block Data Standardized to 2010 Geography.](#)<sup>18</sup>


By applying the general block weights developed by NHGIS to all cohorts, we can estimate block populations in the new geography. Then, the resulting block-level populations are summed to the new geographic tract level.<sup>19</sup>

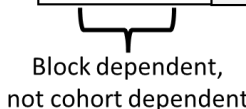
Figure 1 below provides an example of how block-level cohort populations within a Census 2000 tract are distributed to Census 2010 blocks using the NHGIS weights. In Figure 2, the 2000 population is then summed to 2010 tracts.<sup>20</sup>


*Example tract G2500270731202 originally had 300 people total in 2 blocks.*

**Figure 1. Example of Census 2000 Block Cohort Population Distributed to Census 2010 Block**

| 2000 Blocks        | 15 to 19 Females |   | Block Weight | 2010 Blocks        |   | 2000 15 to 19 Females in 2010 blocks |
|--------------------|------------------|---|--------------|--------------------|---|--------------------------------------|
| G25002707312021000 | 100              | × | 1            | G25002707312021001 | = | 100                                  |
| G25002707312021001 | 200              | × | 0.07         | G25002707312021002 | = | 14                                   |
| G25002707312021001 | 200              | × | 0.90         | G25002707312021003 | = | 180                                  |
| G25002707312021001 | 200              | × | 0.03         | G2500270777771000  | = | 6                                    |







*The original 300 people in 1 tract in 2000 geography are distributed to 2 tracts in 2010 geography.*

<sup>18</sup> *ibid*

<sup>19</sup> Due to some adjustments to block and tract boundaries from decade to decade, very small proportions of these smaller geographies crossed state lines in one decade or another, resulting in “slivers” of population and very small weights attributed to out of state population. Because of this, some weights were re-calculated allow in-state tracts and blocks to sum to 100% in Massachusetts geographies.

<sup>20</sup> Note that in Census geographic hierarchy, blocks nest within tracts and tracts are the sum area of multiple blocks.



**Figure 2. Example of Summing Distributed 2000 Population to Census 2010 Tracts**

| 2010 Blocks        | 2000 15 to 19 Females in 2010 blocks |   | 2010 Tracts   | 2000 15 to 19 Females |
|--------------------|--------------------------------------|---|---------------|-----------------------|
| G25002707312021001 | 100                                  | } | Tract 0731202 | 294                   |
| G25002707312021002 | 14                                   |   |               |                       |
| G25002707312021003 | 180                                  |   |               |                       |
| G25002707777771000 | 6                                    | } | Tract 0777777 | 6                     |

It is important to note that in the NHGIS crosswalk weighting, no distinctions are made for different subgroups of population or housing. In effect, the model assumes that, *within each block*, for example, the 15-19-year-old population's *spatial* distribution is the same as every other age-cohort's spatial distribution; the Asian population's spatial distribution is the same as the spatial distribution of any other race; and the spatial distribution of owner-occupied housing is proportionally the same as renter-occupied housing's, etc. In actuality, particular age or race groups etc. may be more heavily clustered in one area of the block over other areas. This assumption of consistent within-block distributions across subgroups is one source of error in the interpolation model, but it greatly simplifies the model and ensures that all interpolated subtotals will correctly sum to totals.<sup>21</sup> The error ranges that we generate for each age/sex cohort capture the error introduced by the distribution assumptions embedded in the tract normalization process because these same assumptions are made when generating the 2020 estimates used to assign error.

The tract-to-tract crosswalk for 2010 to 2020 became available from NHGIS while age/sex/race/ethnicity projections were in production. To develop 2010 base populations by age/sex/race/ethnicity in 2020 tract geographies, UMDI utilized the race-bridged 2010 data file created for the V2017 series, then used the NHGIS 2010 to 2020 tract level crosswalk to distribute race/ethnicity/age/sex to 2020 tract geographies.<sup>22</sup> The result was controlled to the V2024 age/sex launch populations for consistency between the age/sex-only V2024 and age/sex/race/ethnicity V2025 series (see Section 3) .

## 2) Reassigning “Some other race”

To develop race and ethnicity projections to 2030, we re-assign the “Some other race” population counted in the Census 2010 and 2020 data to the six race categories comprising the post-Census 2020 county-level age/sex/race/ethnicity estimates. We do this because “Some other race” is not technically a race category recognized by the Office of Management and Budget (OMB), but rather a survey response category. Re-assigning the “Some other race” category also allows our model to correspond

<sup>21</sup> <https://nhgis.org/documentation/time-series/2000-blocks-to-2010-geog>

<sup>22</sup> Geographic Crosswalks, Minnesota Population Center. IPUMS National Historical Geographic Information System. Accessed at <https://www.nhgis.org/geographic-crosswalks#download> on October 4, 2024.



fully to the county-level age/sex/race/ethnicity population estimates that the Census Bureau updates on an annual basis and, more critically, to conform to the race categories reported by MA DPH. Finally, inconsistencies in “Some other race” reporting from census to census had caused large errors in model testing in our V2017 projections to 2020, the highest among all race groups tested, indicating instability in this reporting category from census to census.

We re-assign the “Some other race” populations by age/sex/ethnicity/geography to the six major race categories using percentage allocations created from the Census Bureau’s Modified Race Summary File<sup>23</sup> for 2010 data and the Modified Age and Race Census (MARC) file<sup>24</sup> for 2020 data. Our method assumes that each sub-county geography’s race reassignment percentages are the same as the parent county’s reassignment percentage. Regarding this assumption, Census Bureau staff advised that within a state the reassignment will not vary much by geography and that the even finer county-level files should be sufficient barring any unusual circumstances.

To apply the 2010 modified race file, we compare each age/sex/race/ethnicity population in the Census 2010 SF1 and Census 2020 DHC-A county files to the age/sex/race/ethnicity population in the corresponding modified county file, in which all “Some other race” populations have been re-assigned, and calculate a ratio between the two<sup>25</sup>. We then apply each age/sex/race/ethnicity/county-specific ratio to the corresponding age/sex/race/ethnicity cohorts within each county. Because processing these ratios within statistical software includes some decimal rounding, there is some difference between the final “modified” race totals and the original SF1 race totals, however these differences due to rounding average less than one person total per town.

At the tract level, after redistributing “Some other race”, each 2010 age/sex/race/ethnicity population is distributed to 2020 geographies by using the NHGIS 2010-2020 tract crosswalk published in June 2024.<sup>26</sup>

Applying the 2020 MARC file follows a similar process but includes some additional steps due to changes in the content and processing of this modified race file from versions the Census Bureau produced for previous Census years. Unlike in previous decades’ Modified Race Summary files, which re-codes Census “Some other race” data for 5-year age groups, the 2020 MARC file was published in single-year age groups and also corrects for “age heaping”. This is a phenomenon where some reported ages—usually ones that end in a 5 or a 0—have artificially high populations in the Census data. It can occur when Census takers rely on neighbors to estimate the age of a person who has not filled out their own Census or when a responder approximates the age of a household member rather than specifying a birth date. To rectify this in the MARC file, the Census Bureau identified and mathematically distributed the “age only” response records that were contributing to age heaping. This age redistribution means that the county age/sex cohort populations and age/sex/ethnicity for the 5-year of age in Census 2020 and the

---

<sup>23</sup> 2010 File: 4/1/2010 Modified Race Characteristics Source: U.S. Census Bureau. Released: July 2010.

[https://www.census.gov/popest/research/modified/STCO-MR2010\\_AL\\_MO.csv](https://www.census.gov/popest/research/modified/STCO-MR2010_AL_MO.csv)

<sup>24</sup> 2020 File: Modified Age and Race Census (MARC) file: Source: U.S. Census Bureau. Released: March 6, 2025.

<https://www.census.gov/data/datasets/2020/demo/popest/modified-race-data-2020.html>

<sup>25</sup> For the 2010 launch data used to develop cohort change ratios, we relied on the race re-assignment of the Census SF1 file previously done for the V2017 projections to 2020. We utilized the same bridging method to reassign races for the 2020 launch data, using the 2010 Modified Race Summary File.

<sup>26</sup> Geographic Crosswalks, Minnesota Population Center. IPUMS National Historical Geographic Information System. Accessed at <https://www.nhgis.org/geographic-crosswalks#download> on October 4, 2024.

2020 MARC file are, unlike previous decades, no longer equal. Our usual process to develop the race-bridging multipliers used to re-code “some other race” at the tract level is based on having equivalent age/sex/ethnicity cohorts between the race bridging file and Census 2020 at the county level for use as denominators. Because these populations no longer match, we first must control the 2020 MARC populations to the Census 2020 county populations by age and sex. Next, we calculate the share of each race cohort within each ethnicity in the controlled MARC file and multiply this by the original Census 2020 ethnicity populations for each age/sex group. This estimates the population of each race within the Hispanic/non-Hispanic populations based on the ethnicities as represented in the 2020 Census. From here, we follow the same steps as we do for the 2010 data to develop the multipliers for each age/sex/race/ethnicity combination.

Because age and ethnicity no longer perfectly align between the Census 2020 and 2020 MARC files, the 2020 race bridging introduces more error, and ethnicities and races may become decoupled to a certain extent when dealing with the redistribution of age categories in the MARC file. For this reason, population estimates of ethnicities or races (e.g., Black OR Hispanic) on their own may be more reliable than estimates in combination (e.g., Black AND Hispanic).

### 3) Adjusting Census population counts

After the 2020 Census count, UMDI assisted cities and towns with population challenges and supporting research to address missing housing units or “group quarters” populations. Towns that warranted corrections according to this research included Middlefield, Erving, Chester, Boston, Bourne, Chicopee, Dartmouth, Dedham, Franklin, Middleton, Randolph, Springfield, Wareham, and Wenham. In the V2024 series, Census 2020 corrections prepared for or submitted to the Census Bureau (regardless of acceptance status) are incorporated into the 2020 launch population.

First, the missing population is distributed by age and sex based on the type of challenge. Populations associated with missing housing units are developed by using the persons per household counted in the 2020 Census for the MCD and the resulting populations are distributed by age and sex based on the Census 2020 age profile of the MCD. Group quarters populations are distributed according to the state-level age/sex structure associated with the group quarters “type” (prisons, student dormitories, nursing homes, etc.), all based on the Census 2020 Demographic and Housing Characteristics file, with some exceptions. A few of the group quarters identified for correction were predominately male or predominately female, and for these we adjusted the cohort ratios. These included the Mass. Maritime Academy in Bourne and Wareham, which was 87% male, St. John’s Seminary in Boston, which was 100% male, the Western MA Regional Women’s Correctional Facility in Springfield which was 100% female and Simmons College in Boston which was 100% female. The resulting populations by age, sex, and geography are then added to the Census 2020 DHC data for an adjusted 2020 population.

Table 1 below shows the population added to the 2020 launch for each of the adjusted municipalities, including the Census group quarters population type for each. In the table below “GQ 1” represents population in correctional facilities; “GQ 3” is nursing home population; “GQ 5” is college student dormitory population; and “GQ 7” represents, in this case, population living in group homes. Appendix Table A1 displays the Massachusetts age/sex shares used for each group quarters type, as reported in the Census 2020 DHC-A file.

**Table 1. Population Adjustment by Municipality and Challenge Type**

| MCD         | Challenge Type | Total Population Change           |
|-------------|----------------|-----------------------------------|
| Middlefield | Housing Units  | 41                                |
| Erving      | Housing Units  | 93                                |
| Boston      | Group Quarters | GQ1: 264 GQ5: 6,026 <sup>27</sup> |
| Bourne      | Group Quarters | GQ5: 220                          |
| Chicopee    | Group Quarters | GQ1: 124                          |
| Dartmouth   | Group Quarters | GQ1: 73                           |
| Dedham      | Group Quarters | GQ1: 122                          |
| Franklin    | Group Quarters | GQ5: 74                           |
| Middleton   | Group Quarters | GQ1: 67                           |
| Randolph    | Group Quarters | GQ3: 60 GQ7:87                    |
| Springfield | Group Quarters | GQ1: 70                           |
| Wareham     | Group Quarters | GQ5: 83                           |
| Wenham      | Group Quarters | GQ5: 79                           |

Finally, we adjust population in our model for the Town of Lincoln and for the census tract in Lincoln that represents the Hanscom Air Force Base.<sup>28</sup> We do this because Lincoln was counted in Census 2010 with a significantly reduced population—as much as 21% of the town and 55% of the tract population by our estimates—because a large number of the housing units at the military base were demolished just prior to the 2010 count and then replaced in 2011. Using the temporarily diminished population in 2010 as a base for calculating ratios and projecting future estimates produces very unreasonable results. To correct for this, we substitute the 2000 tract population for 2010, reasoning that population housed at a military base does not necessarily age or migrate at the same rate as the household population, but, instead, functions like a “revolving door” population in terms of age breakdown.<sup>29</sup> Although this substitution introduces the assumption that the count and characteristics of Lincoln population in 2010 were the same as that of 2000, we believe that overall, this processing decision produces a much more realistic estimate for that area than the unadjusted Census 2010 count. For the Lincoln tract-level populations used in the model, 2000 population was expressed in 2010 and 2020 tract geography, as required by the model. For the MCD-level, a 2010 population was modelled using 1990-2000 CCRS applied to the 2000 population by age and sex.

---

<sup>27</sup> This number includes the 1,077 additional GQ5 population of a new dormitory associated with UMass Boston. This additional population is added to the expected Census 2020 value when calculating errors between the projected and expected 2020 populations, however, it is not used in the creation of the CCRs or the initial 2020 launch population. See the section on “Special adjustments: college tract overrides” for more detail on how this population is added into the projection.

<sup>28</sup> Census 2010 tract 3601.00 in Middlesex County

<sup>29</sup> Also assuming that the mission of the base does not significantly change, which would lead to a base change in workforce and population.

---

## 4) Special adjustments: college tract overrides

In addition to housing units and group quarters populations that UMDI identified as undercounted in the 2020 Census, UMDI uncovered a few more particular cases for population adjustment when reviewing the first run of retrospective 2020 estimates against the 2020 Census count. In this comparison, cohorts in a few geographies were generating unusually large percentage differences that warranted more detailed investigation.

The first of these occurred in the Census tract in Amherst that includes Hampshire College.<sup>30</sup> Between the 2010 Census and the 2020 Census, Hampshire College reduced its student population significantly, from 1,157 students housed in its dormitories (or “Group Quarters Type 5, or “GQ5” population, in U.S. Census terms) to 579 in 2020. Because this tract is comprised exclusively of college students in 2020, the reduction greatly affected both the error, when comparing the 2020 estimate to the 2020 Census count, and the CCRs used to project the population to 2030. To correct for this issue, in the 2030 projection UMDI takes the Census 2020 population by age and sex in the tract and carries it forward to 2030, unchanged, before controlling to the county estimate. This adjustment assumes that the population living at Hampshire College may be treated as a “revolving door” population, with students in 2020 replaced by other students of the same age in future years.

A second, similar situation was observed in the Census tract in Newton that includes the Mount Ida college campus.<sup>31</sup> The college closed between 2010 and 2020 and was purchased by UMass as a location for student intern programs, with limited on-campus housing. According to U.S. Census counts, the on-campus student population (GQ5 population) was reduced from 868 in 2010 to just 54 in 2020. Like the Hampshire College case, the drastic decrease in a very specific age cohort was having a large effect on the cohort-change ratios for those cohorts. Unlike the Hampshire College case, however, the Mount Ida tract also included a very large “household” population not residing in dorms and making up 98.9% of the total tract population. For this tract, UMDI holds constant to 2030 only the 20-24 and 25-29 populations from the Census 2020 count.

Finally, the tract location in Boston for UMass Boston experienced a large increase in the on-campus population due to the opening of student housing at the college between 2010 and 2020.<sup>32</sup> In this case, the GQ5 population increased from 0 in 2010 to 1,077, comprising 29.9% of the total tract population. To adjust for this, UMDI starts with the Census 2020 values for a launch population, performs a CCR based on that, and interpolates to 2030. Then, we add 1,077 (the new dorm population) to the tract population for all years from 2020 through 2030, again assuming that the students added in 2020 will be replaced in future years by new students. To distribute the 1,077 students to age/sex cohorts, we use the state-level age/sex distribution of the GQ5 population.

All of the college tract adjustments are built into the model before controlling to the Census Bureau’s annual county-level age/sex and the age/sex/race/ethnicity estimates for the V2025 model, and before distributing 5-year age cohorts to single years of age. Table 2 below displays the population detail for the college override tracts.

---

<sup>30</sup> 2020 Census tract 25015820802

<sup>31</sup> 2020 Census tract 25017374000

<sup>32</sup> 2020 Census tract 25025090901

**Table 2. College Override Tracts and Populations**

| College        | Hampshire College |        | Mt. Ida     |       | UMass Boston |       |                    |
|----------------|-------------------|--------|-------------|-------|--------------|-------|--------------------|
| GEOID          | 25015820802       |        | 25017374000 |       | 25025090901  |       |                    |
| Year           | 2010              | 2020   | 2010        | 2020  | 2010         | 2020  | 2020 <sup>33</sup> |
| Total Pop      | 1,321             | 579    | 5,237       | 4,808 | 3,730        | 3,600 | 3,600              |
| GQ 5 pop       | 1,157             | 579    | 868         | 54    | -            | -     | 1,077              |
| GQ5 % of total | 87.6%             | 100.0% | 16.6%       | 1.1%  | 0.0%         | 0.0%  | 29.9%              |

## 5) Control race/ethnicity/age/sex launch to age/sex launch

After the "Some other race" category has been redistributed and populations have been assigned to 2020 geographies, the age/sex/race/ethnicity launch populations in 2010 and 2020 require an additional control to the age/sex launch populations in order to incorporate the population adjustments described above. In doing this, we assume that the distributions of race and ethnicity within each age/sex/geography cohort remain the same after adjustments to the total population. While this step is taken to ensure relative consistency between the age/sex and age/sex/ethnicity/race projections, due to rounding at the end of this distribution process, summing to age/sex cohorts from age/sex/ethnicity/race cohorts may include differences of +/- 2 people when compared to the age/sex launches.

## 6) Calculating cohort-change ratios and child-to-woman ratios

### *Calculating Ratios*

Once the decennial launch populations have been established, the next steps are to calculate the 2010 to 2020 CCRs for population age 10 and over and child-to-woman ratios (CTWs) for the population aged 0-9. The CCR method accounts for the aging of each individual cohort from one census to the next and creates a ratio between a specific cohort population (by age, sex, geography, and sometimes additional variations such as race) age  $a$  in year  $y$  to its corresponding cohort aged  $a-10$  in year  $y-10$ . For 2010 to 2020 CCRs, this is expressed as:

$$CCR_{age\ a} = \frac{Population\ 2020_{age\ a}}{Population\ 2010_{age\ a-10}}$$

For example, the CCR for a 25-29-year-old cohort would be calculated as:

$$CCR_{age25-29} = \frac{Population\ 2020_{age\ 25-29}}{Population\ 2010_{age\ 15-19}}$$

<sup>33</sup> Including 1,077 missing students identified by UMDI in Census 2020 evaluation as part of the *Post-Census Group Quarters Review Program* and submitted to the U.S. Census Bureau for correction.

Note that for the 85-plus population, the CCR takes the sum of the population aged 75 and over as its denominator.

The child-to-woman ratios are calculated as follows in our model, with children aged 0-4 calculated as a ratio of the female population aged 20-44 and children aged 5-9 calculated as a ratio of the population aged 30-49:

$$CTW_{age0-4} = \frac{Population\ 2020_{age\ 0-4}}{Population\ 2020_{age\ 20-44}}$$

$$CTW_{age5-9} = \frac{Population\ 2020_{age\ 5-9}}{Population\ 2020_{age\ 30-49}}$$

Using the formulas expressed above, we calculate geographically specific CCRs and CTWs for each age/sex cohort at both the town and tract geographic levels. For all age groups, the resulting cohort-specific ratio is then applied to the corresponding base population (the Census 2020 population in our model) in order to estimate the population 10 years later (2030 in this case).

### ***Capping Ratios***

Before we integrate the resulting CCRs into our model, we cap them at “1” for cohort groups including fewer than 25 people and “2” for groups under 100 people. This “capping” process is a modification to the standard Hamilton-Perry method that we developed through performance testing on the V2016 model. The rationale behind this approach is that, in the case of small cohorts, minor changes in population may result in large percentage increases—or large change ratios in the context of a Hamilton-Perry method. For example, a cohort population of 2 persons aging forward and increasing to 12 persons from 2010 to 2020 would yield a change ratio of  $12 \div 2$ , or 6. Applying the CCR of 6 to the 2020 base population of 12 would now suggest that the future cohort population would increase to 72 by 2030. While unusual or erratic population jumps do sometimes occur—and are accounted for in our estimated error ranges—they do not necessarily need to be perpetuated forward by the model. Unreasonable increases that occur solely as the by-product of a ratio method should be limited, which can be achieved by applying caps to the ratios themselves.

To decide where to cap the CCRs, in the V2017 projections, UMDI examined the frequency of actual 2000-2010 CCR values by cohort size and noted the values at which most CCRs naturally occur and at which level they are outliers. We then conducted sensitivity testing of various CCR cap values applied to different cohort sizes to see which adjustment brought our 2010 estimates closest to the Census 2010 actuals. We determined that a reasonable cap—one that minimizes error in our projections—is the CCR value “1” for cohort groups including fewer than 25 people and “2” for groups under 100 people. In our example of 2 persons increasing to 12, the effect of a CCR cap of “1” would mean that the group of 12 ages forward to a population of 12 again, not 72. For larger groups, a cohort of 30 people that aged up to 90 people over ten years (indicating a CCR of “3”) would be allowed to age up to 60 people using a CCR cap of “2,” but no higher. While the larger CCR cap of “2” still allows for significant growth, both caps mitigate the possibility of “runaway” cohorts that may grow into unreasonably large groups in small places.

Due to the prevalence of very small maternal cohorts in the age/sex/race/ethnicity projection, we adjusted the cap of the child-to-woman ratios for that estimate series. Small maternal cohorts can sometimes generate large child-to-woman ratios simply due to their size. For example, if in 2020 a maternal cohort had 6 women and the 0-4 child cohort had 8 children, the child-to-woman ratio would be 1.33. It is plausible that in this “snapshot” that several women would have had multiple children over the course of 5 years (the time span in which 0-4 year olds would be born) or that families may have moved in with children, but we could not necessarily expect this trend to continue at this rate into the future. For this reason, if the sum of the females used to calculate the child cohort is less than 100, we use a cap of 0.64.

This cap level was selected based on an analysis of the child-to-women (CTW) ratios observed in the Census 2020 data for Massachusetts MCDs. Out of 16,849 observed CTWs by age (0-4 and 5-9), sex, race, ethnicity, and MCD, we first filtered out observations for which the corresponding maternal cohorts summed to fewer than 25 people, leaving 5,366 observations. This filtering was applied to overcome the challenges of small populations resulting in implausible ratios. Next, we analyzed the 90% and 95% CTWs by race. The 90% CTWs, or the CTW ratios under which 90% of all observations fell, by race, included: Asian 0.22, AIAN 0.34, Black 0.35, NHPI 0.47, Two or More Races 0.64, and White 0.34. Because the Two or More races 90% maximum showed the highest fertility of all races in our dataset, we selected its value as the cap. The cap selected was also determined to closely approximate the overall 2023 fertility rate for the Hispanic population of the U.S., which was reported as 65 per 1,000 women by the NCHS CDC.<sup>34</sup> After the NHPI group, this was the second highest fertility rate by reported among the defined major race and ethnicity groups in the U.S. In general, our selection aimed to allow for maximum fertility under capping rules while still minimizing the potentially distorting effects of small maternal populations on the ratios used to project future children.

Because the resulting cohort estimates are ultimately controlled to current age/sex/race/ethnicity estimates at the county level, rapid changes in one age or ethnic group are still permitted in the model. For example, if an unprecedented number of young people aged 20-24 moved into Suffolk County, the UMDI estimates for the 20-24 cohort would be controlled to that new county total, with the sub-county proportions of our estimates model applied to the new and perhaps unanticipated cohort population total. Likewise, if a particular race group increases at an unprecedented rate in a region of the state, the sub-county (town- or tract-level) estimates will sum to meet the new population total or likewise decrease in the case of a declining age or race group. In this way, the caps moderate the level of growth or decline but still ultimately allow for extreme increase within a sub-population should it actually occur.

Two last adjustments are made in the CCR model application. First, populations are rounded to the nearest person prior to the calculation of the CCR and before caps are introduced. This is to prevent unrealistic population growth due to very small fractions in the cohort values. Second, cohorts that have zero population are adjusted to a value of “0.1”. This adjustment serves two purposes. When replacing the denominator in the CCR calculation, the “0.1” replacement corrects division errors caused by a zero denominator. When used as a numerator in the CCR calculation, it ensures that no particular age/sex group will zero out for the full time series with no chance of recuperation. For example, a particular

---

<sup>34</sup> Hamilton BE, Martin JA, Osterman MJ., Births: provisional data for 2023. Vital Statistics Rapid Release; no 35. April 2024. DOI: <https://dx.doi.org/10.15620/cdc/151797>.



town or tract may have zero population in the male 75–79-year-old cohort, meaning that no 65–69-year-olds in the geography survived and stayed to age into the 75-79 cohort between 2010 and 2020, yielding a CCR of “0”. However, if the county experiences a large wave of retiree migrants, we still want the town or tract to be able to pick up some small share of the influx. By first converting the zero populations to a very small value, the CCR value is calculated at >0 and will produce a negligible population in geography without cancelling it out in perpetuity.

### ***Alternative Race Ratios Testing***

While we ultimately decided to use the CCR method outlined above, we did perform testing at an MCD level to determine if allowing the race and ethnicity proportions of each cohort in the base year to age up into the projection year (i.e. CCR =1) would produce more accurate race and ethnicity estimates than producing individual CCRs, due to differences in data collection between 2010 and 2020 and processing choices in 2020.

During the production of the Census 2020 published data and the development of the 2020 MARC file, the Census Bureau used various techniques to code or re-code races depending on what information was available and while adhering to the U.S. Census standards for determining race from write-in nationalities or origins. There were differences in protocol for both the Census and race bridging files between 2010 and 2020. The 2020 Census survey included more characters for survey takers to print their origins, thus allowing more detailed responses. Also, language in the “Some other race” was changed from “Print race” in 2010 to “Print race or origin” in 2020, which, again, allowed for more detailed responses<sup>35</sup>. Coding of origins to main race categories also changed from 2010 to 2020. Using an example in a Census blog post explaining changes to the 2020 Census, someone who identified their origins as “Cuban, Thai, Filipino” would be coded as “Asian” in 2010 (as “Thai” and “Filipino” are detailed race groups), but would be coded as “Some other race and Asian” in 2020, in order to preserve the detail of “Cuban”.<sup>36</sup> Changes such as these may have contributed to a proportionally higher selection of “Some other race” than in previous decades and required the Census to make modifications to the processing of the bridging file to accommodate them.

In 2010 and in 2020, when race was coded as “Some other race” in the Census survey, and there was also another specified race associated with the record, the “Some other race” would be dropped and the specified race would be preserved (e.g. “Some other race and White alone” would become “White alone”). When race was not available from a Census record in 2010, data from the household would be used to impute the race of that individual, and barring that, a hot-deck matrix would be used to assign race.<sup>37</sup> In the 2020 MARC file, first household information would be used to impute race as in 2010. Unlike in 2010, if that was not successful, an additional step of using other administrative data would be

---

<sup>35</sup> Improvements to the 2020 Census Race and Hispanic Origin Question Designs, Data Processing, and Coding Procedures. Marks, Rache and Merarys Rios-Vargas, U.S. Census Bureau, Population Division. August 3, 2021. <https://www.census.gov/newsroom/blogs/random-samplings/2021/08/improvements-to-2020-census-race-hispanic-origin-question-designs.html>

<sup>36</sup> Ibid.

<sup>37</sup> *Modified Race Summary File Methodology*. U.S. Census Bureau, Population Division. July 5, 2012. <https://www2.census.gov/programs-surveys/popest/technical-documentation/methodology/modified-race-summary-file-method/mrsf2010.pdf>



used to impute race. “Hot-decking” was the last method used to assign race if other methods were not successful.<sup>38</sup>

Before proceeding with CCR-based race projections, UMDI wanted to explore whether changes in the Census form (and the resulting Census coding of responses) from 2010 to 2020 caused implausible race projections. We built a visual display dashboard that looks at race by ethnicity by MCD for years 2000, 2010, 2020, and projected 2030 using both our standard method (which takes change from 2010 to 2020 to project to 2030) and a revised method which holds race constant to 2020 distributions as the cohorts age forward. Because each race group is controlled back to the age groups, both 2030 age/sex populations would be consistent, but the race and ethnicity would be different using different CCRs. We discovered that in many places, it does not make much of a difference which ratio we use, or that results from both versions of the 2030 projections looked equally plausible depending on future conditions. However, for MCDs with large college populations, such as Amherst or Boston, Hispanic and some race trends looked extremely implausible when we hold race and ethnicity to 2020 proportions. We suspect two factors may be at play here. The first is that the student population may have a very different demographic makeup than the surrounding city or town. The second is the effect that small numbers of people of a certain race or ethnicity in an age group with a small population can have if the model uses that ratio to age the population forward. A few children in a cohort of 100 can translate into very large numbers when its relative proportion is multiplied by a college population in the thousands—a population which, in actuality, draws its students from other places and is not dependent on the size and make-up of local younger cohorts in the way that most other age groups are. There may be a few places with diverse demographics that may improve slightly from keeping a CCR that includes race and ethnicity equal to 1, but this requires more exploration and analysis to determine which places would benefit from this modification and how it should be applied.

## 7) Applying ratios and caps to the modified Census 2020 base population to create “uncontrolled” 2030 estimates

The next step in the estimation model is to apply the 2010-2020 CCR and CTW ratios to the 2020 Census base population to create 2030 estimates. At both the town and tract levels, we apply the geographically specific CCRs and CTWs for each age/sex cohort to their corresponding 2020 base population. For example, if in Town A there were ten 10-14 year old males in Census 2010 and eight 20-24 year old males in 2020, the CCR would be calculated as  $8 \div 10$ , or 0.80. In this step, the CCR value of 0.80 is now applied to the population of 10-14 year olds in Town A counted in 2020 to project the population of 20-24 year olds in the town by 2030. Race/ethnicity/age/sex projections are produced in a similar fashion, with each CCR and CTW corresponding to a specific race/ethnicity/age/sex/geography cohort. The 2030 race/ethnicity/age/sex projections are controlled to the 2030 age/sex projections prior to interpolating race/ethnicity/age/sex populations for the years 2021-2029.

---

<sup>38</sup> 2020 Modified Age & Race Census File Methodology Statement. U.S. Census Bureau. Sept 3, 2024. <https://www2.census.gov/programs-surveys/popest/technical-documentation/methodology/modified-race-summary-file-method/marc2020-imprace-us.pdf>

## 8) Distribute age/sex estimates to single years 2021-2030

The application of the calculated 2010-2020 CCRs to the Census 2020 launch populations produces CCR estimates for the year 2030. The next step is to distribute the resulting estimates to the single years from 2021 through 2030. Population change is not necessarily linear; it can accelerate or decelerate over time depending on numerous factors, including changes in localized migration due to economic or other development at the local level. However, because our model cannot account for the factors that may cause uneven change in one place or cohort from year to year, we simply create a linear interpolation for each age/sex /geography cohort from 2020 to 2030. The formula for this, as applied to each cohort, is:

$$\text{Population year } y = \left( \frac{\text{Population 2030} - \text{Population 2020}}{10} \right) \times (y - 2020)$$

While this method assumes that change is evenly distributed through the projection period, toggling our estimates results to the year-by-year county-level estimates, as described below, may help to ameliorate the effect of this assumption on our annual estimates.<sup>39</sup>

## 9) Control the 2021 to 2030 estimates by tract and by town to the county-level annual post-census estimates

As described in the *Method Overview* section of this report, we control the sub-county (tract- and town-level) age/sex estimates and age/sex/race/ethnicity estimates to the county-level estimates of equivalent demographic detail produced by the Census Bureau on an annual basis. To do this, we simply sum the individual cohort estimates produced thus far in our model from the sub-county level (town or tract) to county totals and calculate each sub-county cohort's share of the corresponding county cohort. These shares are then applied to the Census Bureau's current county-level estimates to produce a new "controlled" estimate for each cohort for each year. At the time of this report, estimates for 2021 through 2023 have been controlled and estimates from 2024 through 2030 are uncontrolled, as the Bureau has not yet released estimates for these latter years. For this reason, a review of the entire estimates time-series will show a break in series from 2023 to 2024. In general, places that have been growing more quickly this decade than what the last decade would have anticipated will show a drop from 2023 to 2024, while places that have been growing more slowly than anticipated will show the opposite. Once future years are also controlled, the time-series will trend will be smoother. Note that when 2024 and subsequent vintages of the Census Bureau estimates are released and incorporated into the model as controls, all years in the time-series, from 2021 to the current year, should be updated again, as with each new release the Census Bureau makes revisions to previous years in the post-censal series.

---

<sup>39</sup> Because the annual, intercensal estimates produced by the Census Bureau are themselves estimates that use a combination of year-by-year component data adjusted to two census-count endpoints, we have no means of quantifying the effect of the control to the county on error, having no accurate count to measure them against.

## 10) Distributing 5-year age-group estimates and errors to single years of age for ages 0-20

One of the last steps in our method is to calculate estimates for single-year-of-age cohorts for ages 0 through 20. Because we found that the CCRs and CTWs generated for larger cohort groups generally perform better than smaller groups, we controlled our single-year estimates to the estimates we had produced for the 5-year age-group cohorts 0-4 through 20-24.<sup>40</sup>

We distributed the 5-year groups into single-years-of-age using a constant-share ratio method. For each geography (tract or town), we took the share of each single-year-of-age/sex cohort within its 5-year age group in 2020 and applied it to the 5-year age/sex estimates for 2021-2030. This constant share method ensures that if a particular geography has a distinct single-year distribution within a 5-year cohort, such as a large share of 18 and 19 year-olds within the 15-19 group within a college town, the proportion is preserved moving forward.

Other than these types of localized single-year shares, which are preserved in our method, we have no theoretical basis for assuming how, or even that, shares of a single age within its 5-year group will shift from decade to decade. These shifts can and do occur, but their exact causes are not predictable based on historic trends. To support this assumption, we reviewed the time series of single-year age distributions within 5-year groups in Massachusetts from 1990 through 2020. Table 3 below shows the percentage of each single-year within its 5-year cohort, by decade, along with the 10-year percentage point shift. Percentages remain fairly even, especially from 2000 to 2020. However, even when they do shift more significantly, the direction of the shift is not consistent from decade to decade; it may increase over one decade and decrease over the next. This directional change underscores the unpredictability of the shift-in-share, and so we assume a constant share based on the most recent census in 2020.

**Table 3. Single year of age as percent of 5-year age group and 10-year percentage point shift, 1990-2020**

| Age      | Single year of age as percent of 5-year age group |      |      |      | 10-Year Percentage Point Shift |           |           |
|----------|---|------|------|------|--------------------------------|-----------|-----------|
|          | 1990  | 2000 | 2010 | 2020 | 1990-2000                      | 2000-2010 | 2010-2020 |
| < 1 year | 18%   | 20%  | 19%  | 19%  | 2%                             | 0%        | 0%        |
| 1 year   | 22%   | 20%  | 20%  | 19%  | -2%                            | 0%        | 0%        |
| 2 years  | 21%   | 20%  | 20%  | 20%  | -1%                            | 0%        | 0%        |
| 3 years  | 20%   | 20%  | 20%  | 20%  | 0%                             | 0%        | 0%        |
| 4 years  | 20%   | 21%  | 20%  | 21%  | 1%                             | 0%        | 1%        |
| 5 years  | 21%   | 19%  | 20%  | 20%  | -2%                            | 0%        | 0%        |
| 6 years  | 20%   | 19%  | 20%  | 20%  | -1%                            | 0%        | 0%        |
| 7 years  | 20%   | 20%  | 20%  | 20%  | 0%                             | 0%        | 0%        |
| 8 years  | 19%   | 20%  | 20%  | 20%  | 1%                             | 0%        | 0%        |

<sup>40</sup> This assessment is described in greater detail in the V2016 method report: *Small Area Population Estimates for 2011 through 2020. Prepared for the Massachusetts Department of Public Health's Bureau of Environmental Health*. UMass Donahue Institute, October 2016.

|          |     |     |     |     |     |     |     |
|----------|-----|-----|-----|-----|-----|-----|-----|
| 9 years  | 20% | 21% | 20% | 20% | 1%  | 0%  | 0%  |
| 10 years | 21% | 21% | 20% | 19% | 0%  | -1% | 0%  |
| 11 years | 20% | 20% | 20% | 20% | 0%  | 0%  | 0%  |
| 12 years | 20% | 20% | 20% | 20% | 0%  | 0%  | 1%  |
| 13 years | 19% | 19% | 20% | 20% | 0%  | 1%  | 0%  |
| 14 years | 19% | 19% | 21% | 20% | 0%  | 1%  | 0%  |
| 15 years | 17% | 20% | 18% | 18% | 3%  | -1% | 0%  |
| 16 years | 17% | 19% | 19% | 19% | 2%  | 0%  | 0%  |
| 17 years | 18% | 19% | 19% | 19% | 1%  | 0%  | 0%  |
| 18 years | 22% | 21% | 21% | 20% | -1% | 1%  | -1% |
| 19 years | 26% | 22% | 22% | 23% | -4% | 1%  | 1%  |

Five-year age cohorts by race and ethnicity were distributed to single-year cohorts by the same method, using the assumption that, in any year, the race and ethnicity within any single-year- of-age/sex cohort would have the same distribution as the 5-year-age/sex cohort to which it belonged. For example, if in 2021 the 5-year 10-14 female cohort in a particular geography was proportionally 30% white/Hispanic, then 30% of the 10 year old females in 2021 would be assigned white/Hispanic, 30% of the 11 year old females in 2021 would be assigned white/Hispanic, and so on.

## 11) Controlling race and ethnicity projections to age and sex projections

We control all years of race/ethnicity/age/sex projections to the age/sex projections to ensure consistency between both sets. This includes controlling both the 5-year and single-year age groups. Higher variability is expected in the single year age groups because of the very small size of each race/ethnicity/age/sex/geography cohort and the effects that rounding will have during the control process. Summing to age/sex from age/sex/race/ethnicity for single year of age may produce differences of +/- 4 people between corresponding cohorts in the race/ethnicity projection and the age/sex projection due to rounding of each cohort. For this reason, if race and ethnicity are not necessary for a use case, we recommend the use of the V2024 age/sex projections.

## 12) Assigning error and confidence intervals

A key component of the UMDI estimates series for small geographies is the range of error associated with each estimate. Because detailed estimates at fine levels of geography are prone to large estimation error, we need a measure of how accurate, or inaccurate, a specific estimate is to better understand its utility. To produce error ranges for the estimates, we run a “historic” or “retrospective” version of the model, using 2000 and 2010 data to produce age/sex estimates for 2020 at the MCD and tract level and then compare these to the actual Census 2020 counts after we have adjusted for missed housing units and group quarters populations. The difference between the model-generated projections by age, sex, and geography and the adjusted Census counts are considered “errors.”

For the purposes of this project, these errors are tallied in three different ways for three distinct purposes. *Mean absolute percent errors* are used to describe overall model performance; *mean percent*

*errors* describe both the model performance and the directionality of the errors – whether the model over- or under-predicted the Census count value; and *error calculations* are applied to each cohort estimate from the model to compensate for the degree and directionality of the difference between the Census count and the estimate observed in the retrospective test. Each of these error types are described in the following three report sections.

### **Mean Absolute Percent Errors**

One output of this historic comparison is the mean absolute percent error (MAPE). This is simply the absolute percent difference between the age/sex/geography model estimate for 2020 and the adjusted Census 2020 count. Because previous testing in Massachusetts revealed that age and cohort size are the most significant determinants of error in our cohort-change model, we group errors together by age and cohort sizes and combine male and female errors into the same age/size groups.<sup>41</sup> Table A2 in the Appendix of this report displays the age by size groups used for the MCD and tract-level errors. The MAPEs may be best interpreted as measures of the overall historic accuracy of the model by age group by cohort size. Because the Cape and Islands region of the state, including Barnstable, Dukes, and Nantucket counties, all experienced unusual influxes of population in 2020, for the purpose of model evaluation we provide error information for 1) all Massachusetts geographies; 2) Cape and Island geographies only, and; 3) the balance of Massachusetts geographies excluding the Cape and Island counties.

Tables A3 through A5 in the Appendix to this report display the resulting MAPEs and number of observations by age group and cohort size for the MCD-level projections for all Massachusetts MCDs, for the Cape and Island MCDs, and for Massachusetts MCDs excluding the Cape and Islands, respectively. Tables A6 through A8 in the Appendix to this report display the resulting MAPEs and number of observations by age group and cohort size for the tract-level projections for all Massachusetts tracts, for the Cape and Island tracts, and for Massachusetts tracts excluding the Cape and Islands, respectively.<sup>42</sup>

Tables A3 through A5 in the Appendix to this report display the resulting MAPEs and number of observations by age group and cohort size for the MCD-level projections for all Massachusetts MCDs, for the Cape and Island MCDs, and for Massachusetts MCDs excluding the Cape and Islands, respectively. Tables A6 through A8 in the Appendix to this report display the resulting MAPEs and number of observations by age group and cohort size for the tract-level projections for all Massachusetts tracts, for the Cape and Island tracts, and for Massachusetts tracts excluding the Cape and Islands, respectively.<sup>43</sup>

### **Mean Percent Errors**

The mean percent error (MPE) indicates both the magnitude and the directional bias of the error associated with the historic model, such that if our model tended to over- or under-predict particular age/size cohorts, this directionality is also captured in our MPEs. We take the mean of the individual

---

<sup>41</sup> For testing results, refer to our report: *Small Area Population Estimates for 2011 through 2020, Prepared for the Massachusetts Department of Public Health's Bureau of Environmental Health*, October 2016.

<sup>42</sup> No errors are generated for cohorts with population 0 in the base year, and these are not included in the observation counts.

<sup>43</sup> No errors are generated for cohorts with population 0 in the base year, and these are not included in the observation counts.

---

percent errors of the 18 age groups by 10 size categories at the tract level and by 9 size categories at the town level, which are calculated as follows.<sup>44</sup>

$$PE = \frac{\text{projected value} - \text{Census value}}{\text{Census value}}$$

Like the MAPES, the MPEs are grouped together by age by cohort size, and Cape and Islands MPEs are reported separately from the rest of the state.

Tables A6 through A8 in the Appendix to this report display the MPEs and number of observations by age group and cohort size for the MCD-level projections for all Massachusetts MCDs, for the Cape and Island MCDs, and for Massachusetts MCDs excluding the Cape and Islands, respectively. Tables A9 through A14 in the Appendix to this report display the MPEs and number of observations by age group and cohort size for the tract-level projections for all Massachusetts tracts, for the Cape and Island tracts, and for Massachusetts tracts excluding the Cape and Islands, respectively.<sup>45</sup>

## Error Calculations

The model output for the V2024 series includes both “CCR estimates” and “adjusted estimates”. The “CCR estimate” value represents the direct model output value that has not been adjusted by historical error, while the “adjusted estimate” incorporates corrections to the model associated with the MPE. While the MPE is a measure of how close our projection gets to the “true value” of the Census, to adjust our model, we need to develop a factor that will bring our model closer to the “true value,” essentially an “applied” MPE. This requires a different calculation of the percent error, using the projected value as the denominator, as follows:

$$\text{applied PE} = \frac{\text{Census value} - \text{projected value}}{\text{projected value}}$$

In our error assignment, we take the applied mean percent errors and associated standard deviations for the 18 age groups by 10 size categories at the tract level and by 9 size categories at the town level.<sup>46</sup> Having determined little variation between male and female error in earlier model testing, for error assignment we group male and female error together in same age/size groups.<sup>47</sup> For each estimate value, we assign the corresponding historic age/size applied MPE and standard deviation. Next, using the following formula for sample standard deviation, we calculate 95% confidence intervals of percent error for each age/size category (using the number of cohorts in each category to determine the t-value)

---

<sup>44</sup>While errors are calculated for 18 five-year age categories, single year age cohorts are assigned the same errors as the 5-year cohort groups to which they belong. See Table A2 in the Appendix of this report for size categories and age groups used in error assignment.

<sup>45</sup>No errors are generated for cohorts with population 0 in the base year, and these are not included in the observation counts.

<sup>46</sup>While errors are calculated for 18 five-year age categories, single year age cohorts are assigned the same errors as the 5-year cohort groups to which they belong. See Table A2 in the Appendix of this report for size categories and age groups used in error assignment.

<sup>47</sup>For testing results, refer to our report: *Small Area Population Estimates for 2011 through 2020, Prepared for the Massachusetts Department of Public Health's Bureau of Environmental Health*, October 2016.

and output the values as lower and upper confidence intervals around the estimate, where  $\bar{x}$  = applied MPE,  $x$  = individual observation,  $t$  = t-value, and  $n$  = sample size. For age/size groups for which we have zero observations or only 1 observation, we assign the standard deviation of the next youngest cohort in the same size category and the largest t-value possible (degrees of freedom = 1) to calculate the error.

$$\left( \bar{x} - t^* \sqrt{\frac{\sum(x - \bar{x})^2}{(n - 1)}} , \quad \bar{x} + t^* \sqrt{\frac{\sum(x - \bar{x})^2}{(n - 1)}} \right)$$

The assumption in our error assignment method—and one of the major assumptions in our model—is that the model performance will be the same from 2020 to 2030 as it was from 2010 to 2020. In actuality, a number of factors may affect the model’s performance from one period to another. These include both changes in migration, fertility, and death-by-age trends from one decade to the next as well as differences in the accuracy of the actual census counts from 2000 to 2010 to 2020. Because our model incorporates a control back to current county estimates we believe that our current estimates will perform better than the historic estimates we use to assign error. However, the extent of this improvement is unquantified and is reliant also upon the strength of the county estimates themselves.

It is critical to note that the application of the directional bias indicated by the MPE assumes that the model will perform the same this decade as it did in the last. So, if our method under-predicted the population of 0-4 year olds in 2020, we assume that it will under-predict them again in 2030, and the “adjusted estimate” is calculated accordingly, based on our historic MPE. In our output, we refer to the estimates that account for the historic MPEs and standard deviations as our “adjusted estimates,” and our upper and lower bound estimates—our “Lower Pop” and “Upper Pop” values are likewise a product of the historically-based applied MPEs and standard deviations. Like the MAPES, the MPEs and applied MPEs are grouped together by age by cohort size, and Cape and Islands MPEs are reported separately from the rest of the state. In some rare cases, our un-adjusted CCR estimate may fall outside the limits of the “Lower Pop” or “Upper Pop”. This is due to assumptions made in the binning of individual cohorts based on size, age, and geography. While in the majority of cases, the assumption that cohorts with the same size, age and geographical characteristics will behave similarly in the model as each other-- as well as similarly in 2010-to-2020 and 2020-to-2030 projections—there may be occasional outliers where these assumptions do not hold. These cases would probably be identifiable only through the error assignment process and only further research on an individual basis would show whether additional corrections to that specific cohort projection are necessary. Additionally, our upper and lower population bounds represent confidence 95% intervals, and we would expect approximately 5% of the projections to fall outside of that range.

In cases where a researcher does not want to assume that the model will perform similarly in this decade as compared to last, in terms of under- or overestimating a particular group, the unadjusted “CCR estimate” provides a better denominator for rates calculation. For example, in the Cape and Islands and Berkshires, the seasonal or “recreational” regions of the state, a sudden influx of migrants in 2020 due to the COVID-19 pandemic, coupled with delayed decennial census operations that took place well into the summer months, meant that the 2020 Census count came in much higher than what 2010-2019 estimates predicted. Because of this, the “retrospective” model estimates for 2020 indicate that many of the cohorts in this region were under-predicted by the model. Given the special circumstances,



however, one would not want to assume that the model will underpredict again in these areas when projecting 2030 population. In fact, the model may actually over-predict in these areas, as the short-term migration experienced in 2020 -- and captured in the 2010-to-2020 change ratio -- subsides. In these situations, researchers are advised to use the “CCR estimate” instead of the “adjusted estimate,” because the CCR estimate will not include the *directional* (over-count or undercount) adjustment.

### 13) Distributing error and confidence intervals over single years 2021-2030

Our historic model estimates provide us with the historic error associated with a 10-year CCR estimate (in this case a 2020 estimate based on 2010) that we then apply to our 2030 output. Since we create estimates by single year in the entire 2020-2030 series, we also need to distribute this 10-year error across the single year estimates. To do this, we assume that the 10-year error will be evenly distributed across the 10-year period. By the 9<sup>th</sup> year in the time series, for example, the estimates will have reached 90% of their full 10-year errors. In reality, growth or decline in a particular geography may start slowly at the start of a decade and then accelerate, or it may start off strong and then taper off, or even change sporadically. In the absence of definitive information, we assume even change over the ten-year period.

To do this, we start with zero error in our base year 2020, assign the full 10-year error by age/size to the 2030 estimate, and interpolate the error over the interim years. In this method, the first year estimate (2021) will entail 1/10<sup>th</sup> of the full error, the second year (2022) will entail 2/10<sup>th</sup>, and so on until the full 10-year error is reached in 2030. In all years, the age/size assignment is based on the cohort size in the base year 2020. The assumption that every 5-year age/sex and age/sex/race/ethnicity cohort’s error will increase or decrease linearly over the 10-year period also causes the standard deviation to be evenly distributed over the 10-year period. As a result, the confidence interval bounds will linearly expand by fractions of 10. Formulaically, this works out to the following:

$$\text{percent Error year } y = \left( \frac{\text{Error 2030}}{10} \right) \times (y - 2020)$$

Finally, the interpolated mean percent error (“MPE” in the output file) for each year is multiplied by the population estimate for that year (“CCR\_PROJECTION”) and added to the annual population estimate to calculate the adjusted population estimate (“ADJ\_ESTIMATE”). Because a population of zero would necessarily result in an error of zero using this formula, for projections with a value of 0 after 2020, we make the adjusted population estimate equal to the MPE. This is essentially the same approach used to calculate the adjusted population estimate, but using a value of “1” to multiply the MPE. The same calculation is performed using the upper and lower confidence intervals, instead of the MPE, to determine the upper and lower population estimates between which the adjusted estimate is 95% likely to fall. In our final output, these upper and lower bounds are labelled as “LOWER\_POP” and “UPPER\_POP” estimates and are equidistant in value from the “ADJ\_ESTIMATE” value for that year.<sup>48</sup> Error for single year of age is applied by performing these calculations using the MPE of the 5-year

---

<sup>48</sup> Equidistance is approximate due to rounding and may vary by +/- 2 people.



cohort to which the single age cohort belongs (e.g. 10 year olds in a location where there are 40 10-14 years olds are given the MPE of the 10-14 cohort of bin size 25-49).

## 14) Tract-level “college flag”

Tracts that contain colleges or universities pose unique projection challenges for cohort change ratio models. These types of models are dependent on the assumption that the future population of any given cohort is based on the present-day population aging forward with some change due to mortality and migration. In our model, for example, this means that the population of 15–19-year-olds in 2030 is based on the population of 5–9-year-olds in 2020. However, university and college populations (generally associated with cohorts 15-19, 20-24, and, in some cases 25-29), are not, in reality, dependent on the child population in the estimated geography. In some extreme cases, tracts that are solely or primarily comprised of student dormitories may have zero children under 15, but hundreds or thousands of 15–24-year-olds. Using a CCR model to predict college student populations based on child populations will necessarily have a large amount of uncertainty, because the college aged population does not primarily come from the children within the tract, aged forward.

Furthermore, universities may build or remove dormitories, or even close completely during a given decade, resulting in large population changes. These unpredictable “once in a lifetime” events greatly affect the cohort change ratios if they occur within the 2000 – 2010 period and strongly contribute to error associated with the model if they occur during the 2010 – 2020 period.

Because of these known limitations of our model when dealing with college populations, we have created the variable “col\_flag” and flagged as “TRUE” cohorts with ages 15-19, 20-24, and 25-29 in tracts which are known to contain GQ Type 5 populations (dormitories) according to the 2020 Census. Many of these “college tracts” have extremely large errors due to the factors described above, showing that there is a great deal of uncertainty around our projection for these cohorts. Additionally, many tracts which do not contain colleges or university have 15-19-, 20-24-, or 25–29-year-old populations that are similar in size to tracts which do contain colleges and universities. Because we assign errors on the basis of region (Cape or non-Cape), age, and population size, these non-college tracts will have large errors associated with them if they are similar in size to college tracts. Cohorts 15-19, 20-24, and 25-29 of these tracts are also coded as “TRUE” in the “col\_flag” column.

## 15) Summing to total race, ethnicity, and sex

In the series output, we include sums to more generalized demographics constructed from our age/sex and age/sex/race/ethnicity projections. The age/sex projections are summed to create the following totals:

- Total population by age (e.g., all 0-4 year-olds)
- Total population by sex (e.g., all females)

The age/sex/race/ethnicity projections are summed to create the following totals:

- Total races by age, sex, ethnicity (e.g., Hispanic, 0-4, female, of any race)
- Total ethnicities by age, sex, race (e.g., Black, 0-4, female, of any ethnicity)
- Total sexes by age, race, and ethnicity (e.g., 0-4, Black, Hispanic, of any sex)

- Total sexes by age and race (e.g., 0-4, Black, of any sex or ethnicity)
- Total sexes by age and ethnicity (e.g., 0-4, Hispanic, of any sex or race)

CCR Projections, adjusted projections, and upper and lower confidence intervals are simply summed from these same outputs from the individual cohorts which comprise them. This method creates conservative error ranges, particularly for the age/sex/race/ethnicity projections. Because of our controls to age/sex projections for consistency and improved overall accuracy in our age/sex/race/ethnicity model, covariance exists between different ethnicities and different races within a geography. For example, as a person can only be categorized in this model as Hispanic or Not Hispanic, the more Hispanic people there are within a geography (whose combined Hispanic/Not Hispanic population is dictated by the results of our age/sex model), the fewer non-Hispanic people there must necessarily be, and vice versa. If we improve our projection estimates with one demographic, this will likely improve our projection of another demographic because of this interplay. Unfortunately, these effects are difficult to quantify within the model, so, while error effects are not strictly additive, we have chosen to treat them as such, knowing that this will produce a conservative estimate. Additionally, rounding effects become more visible as we sum to broader characteristics. During our quality control process, we found that totals can be up to +/- 4 people due to differences in processing between R, Stata, and Excel, and dependent on how many variables were being summed. At a state level, we found a 345 person state-wide difference between R and Excel/Stata outputs at the tract level, summing all characteristics, which amounted to a 0.005% difference in total population.

## Projections Output

---

Provided to MA DPH with this report are the population estimates files and a geographic crosswalk file. Estimate files are output by year and by MCD and tract for years 2020-2030, for a total of 22 files: 11 at the MCD level and 11 at the tract level. Each of the estimate files includes the variables described in Table 4 below: Geocode, Sex, Age\_Group\_Code, Age\_Group, Race, Ethnicity, CCR\_ESTIMATE, LOWER\_POP, ADJ\_ESTIMATE, UPPER\_POP, EST\_YEAR, and Series. The tract file also includes the “col\_flag” field.

**Table 4. Population Estimates Output Variable Names and Descriptions**

| <i>Variable Name</i>                                    | <i>Variable Description</i>  |
|---|--|
| Geocode   | Census 2020 Tract or MCD code as provided by NHGIS   |
| Sex   | "m" for male, "f" for female, "t" for all sexes  |
| Age_Group_Code  | Coded age cohorts. 5-year age groups coded 101-118 for 0-4 through 85+; single-year age cohorts coded 201-221 for year 00 through 20; all ages coded 301.  |
| Age_Group   | Age of cohort expressed as 00_04 through (top-coded) 85_99. Note that 85_99 includes persons over 99. 00_99 in the age/sex projections include all ages.   |
| Race<br>(age/sex/race/<br>ethnicity model<br>only)      | Census defined race groups: American Indian or Alaska Native (“AIAN”), Asian (“A”), Black or African American (“B”), Native Hawaiian or Other Pacific Islander (“NHPI”), White (“W”), Two or More Races (“TOM”). Also includes any race group (“ALL”),   |
| Ethnicity<br>(age/sex/race/<br>ethnicity model<br>only) | Census defined ethnicity groups: Hispanic (“H”) and Not Hispanic (“NH”). Also includes any ethnicity (“ALL”).  |
| CCR_ESTIMATE  | Estimated population without accounting for historic model error. Use for most cases.  |
| LOWER_POP   | The CCR estimate with the <i>mean percent error</i> and the <i>standard deviation</i> applied to it to create a lower population bound. The <b><i>mean percent error</i></b> (MPE) and the <b><i>standard deviation</i></b> (STD-DEV) are both associated with each age/size cohort according to retrospective 2020 model estimates compared to the (adjusted) Census 2020 count. They capture the directionality and extent of the over/underestimate in the 2020 test. Use in cases which assume |

|                       |   |
|-----------------------|---|
|                       | that the model will over or underpredict the 2030 cohort population with the same directional bias it had when predicting 2020 population.  |
| ADJ_ESTIMATE          | CCR estimate with the MPE applied to it. Population estimate that is adjusted to reflect the over- or under- estimate according to the historic "error" generated by the retrospective 2020 model estimates compared to the (adjusted) Census 2020 count.   |
| UPPER_POP             | The CCR estimate with the <i>mean percent error</i> and the <i>standard deviation</i> applied to it to create an upper population bound. Based on the 2020 retrospective model performance compared to the adjusted Census 2020 count. See "LOWER_POP" note for more detail. Use in cases which assume that the model will over or underpredict the 2030 cohort population with the same directional bias it had when predicting 2020 population. |
| EST_YEAR              | The year of the population estimate.  |
| Series                | Refers to the vintage of the estimate value, in this case: 2025.01 for the January 2025 series.   |
| col_flag (tract only) | College flag. "TRUE" indicates this tract and cohort combination is associated with MPE values derived from tracts that contain college student dormitories and their associated populations (ages 15-29). LOWER_POP, ADJ_POP, and UPPER_POP values may greatly from the projected value due to large uncertainties in projecting college populations using this method.  |

Finally, the series output includes a geographic crosswalk file that displays the geographic code provided by NHGIS along with U.S. Census Bureau Census 2020 geographic codes so that the dataset may be joined to other GIS-based applications.

As described in *the Detailed Methodology* section of this report, estimates for years 2021 through 2023 in the V2025 output are controlled to current U.S. Census Bureau annual county-level population estimates by age, sex, race, and ethnicity. Estimates for the years 2024 through 2030 are "uncontrolled" because county-level estimates for these years had not yet been released at the time of the V2025 series production. Before use, estimates for the years 2024 through 2030 should be controlled to Census county estimates as they are released. In the meantime, users will note a break in the time-series, showing abrupt changes in cohort populations, when comparing years 2023 and 2024.

## Appendix A

**Table A1. Age/Sex Share of Census 2020 Massachusetts Group Quarters Population by Type**

| Sex   | Age Group | GQ1: Correctional Facilities | GQ3: Nursing Homes | GQ5: College Housing | GQ7: Other Non-Institutional |
|-------|-----------|------------------------------|--------------------|----------------------|------------------------------|
| Male  | Under     |                              |                    |                      | 0.9%                         |
| Male  | 5 to 9    |                              |                    |                      | 0.4%                         |
| Male  | 10 to     |                              |                    |                      | 0.4%                         |
| Male  | 15 to     | 1.1%                         |                    | 15.9%                | 2.3%                         |
| Male  | 20 to     | 8.1%                         | 0.1%               | 25.6%                | 3.1%                         |
| Male  | 25 to     | 14.8%                        | 0.2%               | 2.3%                 | 4.0%                         |
| Male  | 30 to     | 16.1%                        | 0.1%               | 0.8%                 | 4.6%                         |
| Male  | 35 to     | 13.8%                        | 0.2%               | 0.3%                 | 4.4%                         |
| Male  | 40 to     | 10.9%                        | 0.3%               | 0.2%                 | 4.2%                         |
| Male  | 45 to     | 8.4%                         | 0.4%               | 0.1%                 | 4.4%                         |
| Male  | 50 to     | 7.7%                         | 0.9%               | 0.1%                 | 5.5%                         |
| Male  | 55 to     | 6.1%                         | 1.8%               | 0.1%                 | 6.5%                         |
| Male  | 60 to     | 3.7%                         | 3.1%               | 0.1%                 | 5.7%                         |
| Male  | 65 to     | 2.4%                         | 3.7%               | 0.0%                 | 3.7%                         |
| Male  | 70 to     | 1.3%                         | 4.5%               |                      | 2.5%                         |
| Male  | 75 to     | 0.6%                         | 4.7%               |                      | 1.5%                         |
| Male  | 80 to     | 0.1%                         | 4.6%               |                      | 1.0%                         |
| Male  | 85 plus   | 0.1%                         | 10.1%              |                      | 2.1%                         |
| Femal | Under     |                              |                    |                      | 0.7%                         |
| Femal | 5 to 9    |                              |                    |                      | 0.3%                         |
| Femal | 10 to     |                              |                    |                      | 0.3%                         |
| Femal | 15 to     | 0.0%                         |                    | 21.7%                | 1.5%                         |
| Femal | 20 to     | 0.4%                         | 0.1%               | 30.1%                | 2.5%                         |
| Femal | 25 to     | 0.7%                         | 0.1%               | 1.8%                 | 3.0%                         |
| Femal | 30 to     | 0.9%                         | 0.1%               | 0.5%                 | 3.6%                         |
| Femal | 35 to     | 0.9%                         | 0.1%               | 0.2%                 | 2.8%                         |
| Femal | 40 to     | 0.5%                         | 0.2%               | 0.1%                 | 2.5%                         |
| Femal | 45 to     | 0.5%                         | 0.3%               | 0.0%                 | 2.3%                         |
| Femal | 50 to     | 0.3%                         | 0.7%               | 0.1%                 | 2.7%                         |
| Femal | 55 to     | 0.2%                         | 1.3%               | 0.0%                 | 3.4%                         |
| Femal | 60 to     | 0.1%                         | 2.3%               | 0.0%                 | 3.1%                         |
| Femal | 65 to     | 0.1%                         | 3.5%               | 0.0%                 | 2.4%                         |
| Femal | 70 to     | 0.0%                         | 5.1%               |                      | 2.3%                         |
| Femal | 75 to     | 0.0%                         | 7.5%               |                      | 1.9%                         |
| Femal | 80 to     | 0.0%                         | 9.6%               |                      | 2.1%                         |
| Femal | 85 plus   | 0.0%                         | 34.2%              |                      | 5.4%                         |

Source: Census 2020 DHC-A. GQ1 "under 20" population reported in Census DHC-A is assigned to "15-19" for this purpose; GQ3 "under 25" is assigned to "20-24"; GQ5 "under 20" is assigned to "15-19" and "65 and above" is assigned to "65-69".

**Table A2. Cohort-size categories used in error assignment**

| Tract Level | Town Level  |
|-------------|-------------|
| 0-29        | 0-24        |
| 30-49       | 25-49       |
| 50-74       | 50-99       |
| 75-99       | 100-199     |
| 100-124     | 200-499     |
| 125-149     | 500-999     |
| 150-174     | 1,000-1,999 |
| 175-199     | 2,000-9,999 |
| 200-299     | 10,000+     |
| 300+        |             |

**Table A3. Mean Absolute Percent Errors by Age Group and Cohort Size, MCD-Level, All Massachusetts Geographies**

Mean absolute percent error by age group and cohort size (in 2010) Estimates produced for 2020 at age-group/sex/town level and measured against Census 2020 values. Male and female errors combined in MAPEs summary.

| MAPEs by Age and Cohort Size                  |          |     |     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |
|---|----------|-----|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-----|
| Cohort Size                                   | Avg. All | 0-4 | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+ |
| 0_24  | 40%      | 40% | 45% | 46%   | 37%   | 53%   | 37%   | 39%   | 37%   | 45%   | 60%   | 38%   | 34%   | 28%   | 21%   | 24%   | 37%   | 40%   | 48% |
| 25_49   | 22%      | 29% | 25% | 24%   | 27%   | 26%   | 25%   | 20%   | 24%   | 30%   | 23%   | 24%   | 22%   | 18%   | 16%   | 16%   | 15%   | 20%   | 19% |
| 50_99   | 18%      | 20% | 22% | 17%   | 19%   | 26%   | 21%   | 24%   | 21%   | 25%   | 24%   | 17%   | 15%   | 12%   | 13%   | 13%   | 13%   | 16%   | 20% |
| 100_199                                       | 14%      | 18% | 21% | 15%   | 16%   | 21%   | 14%   | 15%   | 15%   | 19%   | 13%   | 13%   | 11%   | 10%   | 10%   | 10%   | 12%   | 13%   | 18% |
| 200_499                                       | 12%      | 14% | 13% | 12%   | 11%   | 19%   | 15%   | 16%   | 15%   | 16%   | 12%   | 9%    | 7%    | 8%    | 8%    | 8%    | 9%    | 11%   | 13% |
| 500_999                                       | 10%      | 12% | 12% | 12%   | 10%   | 15%   | 11%   | 13%   | 13%   | 14%   | 12%   | 7%    | 5%    | 6%    | 8%    | 7%    | 9%    | 10%   | 12% |
| 1000_1999                                     | 9%       | 13% | 10% | 12%   | 11%   | 12%   | 8%    | 10%   | 10%   | 13%   | 9%    | 6%    | 5%    | 7%    | 9%    | 8%    | 10%   | 9%    | 9%  |
| 2000_9999                                     | 10%      | 9%  | 9%  | 11%   | 13%   | 16%   | 11%   | 10%   | 9%    | 8%    | 7%    | 7%    | 7%    | 7%    | 8%    | 9%    | 10%   | 16%   | 15% |
| 10000+  | 7%       | 10% | 9%  | 3%    | 9%    | 22%   | 9%    | 6%    | 4%    | 2%    | 2%    | 4%    | 7%    | 7%    | 9%    |       |       |       |     |
| Number of Observations by Age and Cohort Size |          |     |     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |
|   | Total n  | 0-4 | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+ |
| 0_24  | 1,131    | 70  | 56  | 43    | 44    | 63    | 75    | 74    | 52    | 31    | 22    | 21    | 17    | 20    | 45    | 82    | 114   | 137   | 165 |
| 25_49   | 1,081    | 61  | 59  | 59    | 58    | 69    | 60    | 66    | 54    | 59    | 37    | 32    | 35    | 37    | 66    | 75    | 77    | 87    | 90  |
| 50_99   | 1,465    | 71  | 62  | 59    | 58    | 83    | 91    | 82    | 74    | 52    | 69    | 70    | 74    | 83    | 77    | 103   | 112   | 135   | 110 |
| 100_199                                       | 1,836    | 98  | 75  | 70    | 72    | 98    | 118   | 108   | 90    | 85    | 73    | 76    | 81    | 98    | 139   | 152   | 152   | 129   | 122 |
| 200_499                                       | 3,397    | 199 | 204 | 197   | 213   | 209   | 176   | 184   | 195   | 180   | 172   | 176   | 214   | 226   | 210   | 190   | 164   | 148   | 140 |
| 500_999                                       | 2,167    | 114 | 143 | 163   | 148   | 87    | 88    | 98    | 133   | 162   | 168   | 179   | 165   | 151   | 114   | 75    | 66    | 54    | 59  |
| 1000_1999                                     | 1,121    | 61  | 82  | 91    | 79    | 50    | 55    | 53    | 71    | 96    | 122   | 107   | 88    | 66    | 43    | 21    | 14    | 9     | 13  |
| 2000_9999                                     | 411      | 26  | 19  | 18    | 28    | 41    | 37    | 35    | 31    | 35    | 37    | 39    | 26    | 19    | 7     | 4     | 3     | 3     | 3   |
| 10000+  | 27       | 2   | 2   | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 1     |       |       |       |     |

**Table A4. Mean Absolute Percent Errors by Age Group and Cohort Size, MCD-Level, Cape and Island Geographies**

Mean absolute percent error by age group and cohort size (in 2010) Estimates produced for 2020 at age-group/sex/town level and measured against Census 2020 values. Male and female errors combined in MAPEs summary. For MCDs in Barnstable, Dukes, and Nantucket Counties.

| MAPEs by Age and Cohort Size                  |          |     |     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |
|---|----------|-----|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-----|
| Cohort Size                                   | Avg. All | 0-4 | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+ |
| 0_24  | 52%      | 30% | 99% | 48%   | 57%   | 58%   | 52%   | 47%   | 41%   | 59%   | 97%   | 65%   | 24%   | 32%   | 13%   | 46%   | 68%   | 43%   | 50% |
| 25_49   | 27%      | 30% | 26% | 25%   | 30%   | 45%   | 25%   | 19%   | 19%   | 33%   | 36%   | 20%   | 33%   | 40%   | 38%   | 18%   | 25%   | 27%   | 20% |
| 50_99   | 26%      | 18% | 23% | 17%   | 25%   | 35%   | 37%   | 30%   | 27%   | 27%   | 46%   | 36%   | 25%   | 5%    | 23%   | 15%   | 30%   | 19%   | 31% |
| 100_199                                       | 18%      | 21% | 18% | 10%   | 19%   | 27%   | 19%   | 11%   | 17%   | 21%   | 13%   | 22%   | 18%   | 17%   | 17%   | 11%   | 26%   | 14%   | 21% |
| 200_499                                       | 16%      | 14% | 16% | 10%   | 15%   | 24%   | 19%   | 23%   | 17%   | 20%   | 19%   | 21%   | 16%   | 12%   | 10%   | 8%    | 16%   | 12%   | 12% |
| 500_999                                       | 10%      | 15% | 9%  | 5%    | 11%   | 25%   | 13%   | 5%    | 7%    | 6%    | 8%    | 9%    | 8%    | 8%    | 12%   | 13%   | 15%   | 12%   | 7%  |
| 1000_1999                                     | 12%      | 7%  | 22% | 18%   | 21%   | 27%   | 5%    | 10%   | 10%   | 19%   | 10%   | 8%    | 9%    | 9%    | 9%    | 17%   | 14%   |       |     |
| 2000_9999                                     | 12%      |     |     |       |       |       |       |       |       |       |       | 12%   |       |       |       |       |       |       |     |
| Number of Observations by Age and Cohort Size |          |     |     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |
|   | Total n  | 0-4 | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+ |
| 0_24  | 98       | 7   | 5   | 4     | 6     | 6     | 6     | 6     | 6     | 4     | 4     | 4     | 4     | 4     | 4     | 6     | 6     | 7     | 9   |
| 25_49   | 71       | 5   | 6   | 7     | 5     | 5     | 2     | 3     | 2     | 2     | 2     | 2     | 2     | 1     | 2     | 2     | 5     | 10    | 8   |
| 50_99   | 126      | 9   | 11  | 7     | 8     | 13    | 11    | 11    | 9     | 7     | 2     | 1     | 1     | 1     | 6     | 11    | 9     | 5     | 4   |
| 100_199                                       | 166      | 7   | 4   | 8     | 7     | 4     | 10    | 8     | 11    | 13    | 16    | 14    | 14    | 14    | 10    | 5     | 5     | 6     | 10  |
| 200_499                                       | 209      | 13  | 13  | 10    | 11    | 12    | 10    | 12    | 10    | 9     | 11    | 10    | 11    | 11    | 12    | 15    | 15    | 13    | 11  |
| 500_999                                       | 113      | 3   | 5   | 8     | 7     | 3     | 5     | 4     | 6     | 9     | 6     | 10    | 10    | 10    | 7     | 6     | 5     | 5     | 4   |
| 1000_1999                                     | 44       | 2   | 2   | 2     | 2     | 3     | 2     | 2     | 2     | 2     | 5     | 4     | 4     | 5     | 5     | 1     | 1     |       |     |
| 2000_9999                                     | 1        |     |     |       |       |       |       |       |       |       |       | 1     |       |       |       |       |       |       |     |



**Table A5. Mean Absolute Percent Errors by Age Group and Cohort Size, MCD-Level, Balance of State**

Mean absolute percent error by age group and cohort size (in 2010) Estimates produced for 2020 at age-group/sex/town level and measured against adjusted Census 2020 values. Male and female errors combined in MAPEs summary. For MCDs in Massachusetts excluding Barnstable, Dukes, and Nantucket Counties.

| MAPEs by Age and Cohort Size                  |          |     |     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |
|---|----------|-----|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-----|
| Cohort Size                                   | Avg. All | 0-4 | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+ |
| 0_24  | 39%      | 41% | 40% | 46%   | 33%   | 53%   | 36%   | 38%   | 37%   | 43%   | 52%   | 32%   | 37%   | 27%   | 22%   | 22%   | 35%   | 40%   | 48% |
| 25_49   | 22%      | 29% | 24% | 24%   | 27%   | 24%   | 25%   | 20%   | 24%   | 30%   | 22%   | 24%   | 21%   | 18%   | 15%   | 16%   | 15%   | 19%   | 19% |
| 50_99   | 18%      | 21% | 21% | 17%   | 18%   | 25%   | 19%   | 24%   | 20%   | 25%   | 23%   | 17%   | 15%   | 12%   | 13%   | 13%   | 12%   | 16%   | 20% |
| 100_199                                       | 14%      | 17% | 22% | 16%   | 16%   | 21%   | 14%   | 15%   | 15%   | 19%   | 13%   | 11%   | 9%    | 9%    | 9%    | 10%   | 11%   | 13%   | 18% |
| 200_499                                       | 12%      | 14% | 13% | 13%   | 11%   | 19%   | 15%   | 16%   | 15%   | 16%   | 12%   | 8%    | 7%    | 8%    | 7%    | 8%    | 8%    | 11%   | 13% |
| 500_999                                       | 10%      | 12% | 12% | 12%   | 10%   | 14%   | 10%   | 13%   | 13%   | 15%   | 12%   | 7%    | 5%    | 6%    | 8%    | 7%    | 8%    | 10%   | 12% |
| 1000_1999                                     | 9%       | 14% | 9%  | 12%   | 10%   | 12%   | 8%    | 10%   | 10%   | 13%   | 9%    | 6%    | 5%    | 7%    | 8%    | 8%    | 10%   | 9%    | 9%  |
| 2000_9999                                     | 10%      | 9%  | 9%  | 11%   | 13%   | 16%   | 11%   | 10%   | 9%    | 8%    | 7%    | 7%    | 7%    | 7%    | 8%    | 9%    | 10%   | 16%   | 15% |
| 10000+  | 7%       | 10% | 9%  | 3%    | 9%    | 22%   | 9%    | 6%    | 4%    | 2%    | 2%    | 4%    | 7%    | 7%    | 9%    |       |       |       |     |
| Number of Observations by Age and Cohort Size |          |     |     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |
|   | Total n  | 0-4 | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+ |
| 0_24  | 1,033    | 63  | 51  | 39    | 38    | 57    | 69    | 68    | 46    | 27    | 18    | 17    | 13    | 16    | 41    | 76    | 108   | 130   | 156 |
| 25_49   | 1,010    | 56  | 53  | 52    | 53    | 64    | 58    | 63    | 52    | 57    | 35    | 30    | 33    | 36    | 64    | 73    | 72    | 77    | 82  |
| 50_99   | 1,339    | 62  | 51  | 52    | 50    | 70    | 80    | 71    | 65    | 45    | 67    | 69    | 73    | 82    | 71    | 92    | 103   | 130   | 106 |
| 100_199                                       | 1,670    | 91  | 71  | 62    | 65    | 94    | 108   | 100   | 79    | 72    | 57    | 62    | 67    | 84    | 129   | 147   | 147   | 123   | 112 |
| 200_499                                       | 3,188    | 186 | 191 | 187   | 202   | 197   | 166   | 172   | 185   | 171   | 161   | 166   | 203   | 215   | 198   | 175   | 149   | 135   | 129 |
| 500_999                                       | 2,054    | 111 | 138 | 155   | 141   | 84    | 83    | 94    | 127   | 153   | 162   | 169   | 155   | 141   | 107   | 69    | 61    | 49    | 55  |
| 1000_1999                                     | 1,077    | 59  | 80  | 89    | 77    | 47    | 53    | 51    | 69    | 94    | 117   | 103   | 84    | 61    | 38    | 20    | 13    | 9     | 13  |
| 2000_9999                                     | 410      | 26  | 19  | 18    | 28    | 41    | 37    | 35    | 31    | 35    | 37    | 38    | 26    | 19    | 7     | 4     | 3     | 3     | 3   |
| 10000+  | 27       | 2   | 2   | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 1     |       |       |       |     |

**Table A6. Mean Absolute Percent Errors by Age Group and Cohort Size, Tract Level, All Massachusetts Geographies**

Mean absolute percent error by age group and cohort size (in 2010) Estimates produced for 2020 at age-group/sex/tract level and measured against adjusted Census 2020 values. Male and female errors combined in MAPEs summary.

| MAPEs by Age and Cohort Size                  |          |     |     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
|---|----------|-----|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|
| Cohort Size                                   | Avg. All | 0-4 | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+  |
| 0-29  | 41%      | 62% | 57% | 51%   | 57%   | 58%   | 37%   | 68%   | 59%   | 79%   | 75%   | 70%   | 71%   | 60%   | 42%   | 30%   | 32%   | 34%   | 39%  |
| 30-49   | 24%      | 36% | 26% | 29%   | 32%   | 34%   | 41%   | 31%   | 29%   | 38%   | 39%   | 29%   | 28%   | 25%   | 20%   | 19%   | 20%   | 20%   | 26%  |
| 50-74   | 21%      | 26% | 26% | 23%   | 23%   | 28%   | 25%   | 31%   | 28%   | 24%   | 24%   | 21%   | 18%   | 18%   | 16%   | 15%   | 15%   | 16%   | 22%  |
| 75-99   | 19%      | 23% | 21% | 18%   | 20%   | 24%   | 22%   | 25%   | 20%   | 21%   | 20%   | 18%   | 16%   | 14%   | 13%   | 11%   | 13%   | 15%   | 28%  |
| 100-124                                       | 17%      | 22% | 16% | 18%   | 17%   | 22%   | 21%   | 21%   | 20%   | 20%   | 18%   | 15%   | 14%   | 12%   | 11%   | 12%   | 13%   | 14%   | 22%  |
| 125-149                                       | 16%      | 19% | 17% | 16%   | 16%   | 21%   | 19%   | 21%   | 19%   | 19%   | 14%   | 14%   | 11%   | 11%   | 10%   | 10%   | 12%   | 13%   | 30%  |
| 150-174                                       | 16%      | 19% | 16% | 16%   | 15%   | 19%   | 19%   | 19%   | 19%   | 19%   | 15%   | 11%   | 10%   | 10%   | 10%   | 11%   | 12%   | 25%   | 36%  |
| 175-199                                       | 16%      | 20% | 17% | 16%   | 14%   | 20%   | 25%   | 20%   | 18%   | 18%   | 15%   | 11%   | 10%   | 9%    | 9%    | 24%   | 8%    | 14%   | 35%  |
| 200-299                                       | 16%      | 22% | 16% | 15%   | 15%   | 25%   | 29%   | 24%   | 20%   | 18%   | 13%   | 9%    | 8%    | 9%    | 9%    | 31%   | 25%   | 25%   | 39%  |
| 300+  | 28%      | 33% | 17% | 17%   | 42%   | 61%   | 49%   | 23%   | 14%   | 15%   | 12%   | 9%    | 6%    | 20%   | 61%   |       | 17%   | 35%   | 72%  |
| Number of Observations by Age and Cohort Size |          |     |     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
|   | Total n  | 0-4 | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+  |
| 0-29  | 4,953    | 130 | 208 | 206   | 118   | 70    | 90    | 97    | 77    | 78    | 82    | 91    | 90    | 113   | 211   | 439   | 675   | 980   | 1198 |
| 30-49   | 4,861    | 141 | 159 | 159   | 132   | 128   | 163   | 126   | 83    | 61    | 56    | 59    | 92    | 153   | 382   | 630   | 776   | 830   | 731  |
| 50-74   | 6,956    | 332 | 276 | 254   | 260   | 341   | 342   | 325   | 222   | 213   | 202   | 198   | 265   | 383   | 602   | 829   | 807   | 627   | 478  |
| 75-99   | 7,186    | 476 | 394 | 385   | 350   | 487   | 402   | 452   | 426   | 308   | 254   | 246   | 360   | 456   | 644   | 585   | 398   | 337   | 226  |
| 100-124                                       | 6,610    | 486 | 431 | 412   | 418   | 427   | 419   | 446   | 470   | 390   | 321   | 354   | 396   | 439   | 459   | 309   | 199   | 112   | 122  |
| 125-149                                       | 5,825    | 428 | 391 | 358   | 397   | 437   | 379   | 414   | 459   | 413   | 325   | 357   | 412   | 415   | 331   | 111   | 79    | 43    | 76   |
| 150-174                                       | 4,937    | 371 | 372 | 335   | 364   | 307   | 292   | 345   | 366   | 409   | 368   | 370   | 349   | 364   | 208   | 43    | 11    | 12    | 51   |
| 175-199                                       | 3,624    | 268 | 246 | 248   | 275   | 217   | 214   | 248   | 315   | 275   | 317   | 304   | 315   | 256   | 82    | 4     | 5     | 6     | 29   |
| 200-299                                       | 6,660    | 308 | 424 | 496   | 486   | 335   | 446   | 409   | 492   | 669   | 763   | 765   | 626   | 367   | 34    | 4     | 3     | 4     | 29   |
| 300+  | 1,560    | 14  | 53  | 101   | 154   | 205   | 207   | 92    | 44    | 138   | 266   | 210   | 49    | 8     | 1     |       | 1     | 3     | 14   |

**Table A7. Mean Absolute Percent Errors by Age Group and Cohort Size, Tract-Level, Cape and Island Geographies**

Mean absolute percent error by age group and cohort size (in 2010) Estimates produced for 2020 at age-group/sex/tract level and measured against adjusted Census 2020 values. Male and female errors combined in MAPEs summary. For tracts in Barnstable, Dukes, and Nantucket Counties.

| MAPEs by Age and Cohort Size                  |          |     |     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |
|---|----------|-----|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-----|
| Cohort Size                                   | Avg. All | 0-4 | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+ |
| 0-29  | 55%      | 37% | 72% | 33%   | 46%   | 67%   | 49%   | 44%   | 47%   | 57%   | 39%   | 65%   | 80%   | 130%  | 94%   | 75%   | 75%   | 51%   | 39% |
| 30-49   | 29%      | 28% | 36% | 30%   | 23%   | 37%   | 32%   | 33%   | 32%   | 32%   | 53%   | 28%   |       | 30%   | 7%    | 24%   | 24%   | 28%   | 21% |
| 50-74   | 25%      | 36% | 30% | 20%   | 32%   | 35%   | 26%   | 25%   | 26%   | 30%   | 30%   | 34%   | 25%   | 23%   | 27%   | 20%   | 23%   | 19%   | 17% |
| 75-99   | 19%      | 19% | 20% | 18%   | 15%   | 24%   | 21%   | 21%   | 16%   | 20%   | 21%   | 25%   | 22%   | 15%   | 20%   | 13%   | 17%   | 16%   | 26% |
| 100-124                                       | 18%      | 25% | 18% | 14%   | 17%   | 20%   | 25%   | 23%   | 17%   | 14%   | 19%   | 19%   | 20%   | 12%   | 17%   | 13%   | 16%   | 11%   | 23% |
| 125-149                                       | 16%      | 26% | 14% | 11%   | 14%   | 18%   | 15%   | 26%   | 23%   | 12%   | 15%   | 17%   | 14%   | 14%   | 13%   | 12%   | 11%   | 16%   | 38% |
| 150-174                                       | 14%      | 10% | 17% | 14%   | 12%   | 25%   | 9%    | 9%    | 20%   | 15%   | 14%   | 13%   | 12%   | 17%   | 11%   | 10%   | 19%   | 5%    | 17% |
| 175-199                                       | 14%      | 25% |     | 3%    | 27%   | 22%   | 12%   | 22%   | 10%   | 16%   | 12%   | 12%   | 11%   | 15%   | 11%   |       |       | 24%   | 17% |
| 200-299                                       | 12%      | 34% | 17% | 8%    | 6%    |       | 11%   | 28%   | 16%   | 12%   | 15%   | 12%   | 10%   | 9%    | 7%    | 11%   | 16%   |       | 14% |
| 300+  | 11%      |     |     | 13%   | 26%   | 67%   |       |       |       | 6%    | 4%    | 6%    | 12%   |       |       |       |       |       | 6%  |
| Number of Observations by Age and Cohort Size |          |     |     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |
|   | Total n  | 0-4 | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+ |
| 0-29  | 265      | 21  | 16  | 16    | 15    | 15    | 16    | 20    | 12    | 9     | 9     | 10    | 10    | 10    | 12    | 12    | 15    | 21    | 26  |
| 30-49   | 248      | 24  | 21  | 18    | 18    | 26    | 17    | 17    | 16    | 8     | 2     | 2     |       | 3     | 3     | 6     | 11    | 23    | 33  |
| 50-74   | 396      | 19  | 24  | 28    | 24    | 25    | 27    | 33    | 32    | 21    | 15    | 4     | 6     | 5     | 7     | 24    | 30    | 36    | 36  |
| 75-99   | 433      | 35  | 28  | 13    | 20    | 39    | 33    | 31    | 27    | 25    | 17    | 14    | 12    | 5     | 19    | 31    | 34    | 35    | 15  |
| 100-124                                       | 373      | 19  | 27  | 28    | 22    | 13    | 23    | 14    | 19    | 21    | 20    | 23    | 21    | 18    | 26    | 28    | 28    | 13    | 10  |
| 125-149                                       | 280      | 9   | 8   | 17    | 12    | 9     | 13    | 15    | 16    | 24    | 14    | 17    | 26    | 22    | 28    | 23    | 15    | 6     | 6   |
| 150-174                                       | 201      | 3   | 6   | 7     | 14    | 6     | 3     | 2     | 9     | 13    | 24    | 21    | 17    | 34    | 23    | 11    | 2     | 1     | 5   |
| 175-199                                       | 108      | 5   |     | 2     | 3     | 2     | 2     | 3     | 3     | 7     | 14    | 18    | 19    | 15    | 13    |       |       | 1     | 1   |
| 200-299                                       | 124      | 1   | 6   | 5     | 7     |       | 2     | 1     | 2     | 6     | 16    | 21    | 23    | 24    | 5     | 1     | 1     |       | 3   |
| 300+  | 20       |     |     | 2     | 1     | 1     |       |       |       | 2     | 5     | 6     | 2     |       |       |       |       |       | 1   |

**Table A8. Mean Absolute Percent Errors by Age Group and Cohort Size, Tract-Level, Balance of State**

Mean absolute percent error by age group and cohort size (in 2010) Estimates produced for 2020 at age-group/sex/tract level and measured against Census 2020 values. Male and female errors combined in MAPEs summary. For tracts in Massachusetts excluding Barnstable, Dukes, and Nantucket Counties.

| MAPEs by Age and Cohort Size                  |          |     |     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
|---|----------|-----|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|
| Cohort Size                                   | Avg. All | 0-4 | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+  |
| 0-29  | 40%      | 66% | 56% | 53%   | 58%   | 55%   | 34%   | 76%   | 61%   | 81%   | 77%   | 71%   | 71%   | 56%   | 40%   | 30%   | 31%   | 33%   | 39%  |
| 30-49   | 24%      | 37% | 25% | 29%   | 34%   | 34%   | 42%   | 31%   | 29%   | 38%   | 38%   | 29%   | 28%   | 25%   | 20%   | 18%   | 20%   | 20%   | 27%  |
| 50-74   | 20%      | 25% | 25% | 23%   | 22%   | 28%   | 25%   | 31%   | 28%   | 23%   | 23%   | 21%   | 18%   | 18%   | 16%   | 15%   | 15%   | 16%   | 23%  |
| 75-99   | 19%      | 23% | 21% | 18%   | 21%   | 24%   | 22%   | 25%   | 21%   | 21%   | 20%   | 18%   | 16%   | 14%   | 13%   | 11%   | 13%   | 15%   | 28%  |
| 100-124                                       | 17%      | 22% | 16% | 18%   | 17%   | 22%   | 21%   | 21%   | 20%   | 20%   | 17%   | 15%   | 14%   | 12%   | 11%   | 12%   | 12%   | 14%   | 22%  |
| 125-149                                       | 16%      | 19% | 18% | 16%   | 16%   | 21%   | 19%   | 21%   | 19%   | 20%   | 14%   | 13%   | 11%   | 11%   | 10%   | 10%   | 13%   | 12%   | 29%  |
| 150-174                                       | 16%      | 19% | 16% | 16%   | 15%   | 19%   | 19%   | 19%   | 19%   | 19%   | 15%   | 11%   | 10%   | 9%    | 10%   | 11%   | 11%   | 26%   | 38%  |
| 175-199                                       | 16%      | 20% | 17% | 16%   | 14%   | 20%   | 25%   | 20%   | 18%   | 18%   | 15%   | 10%   | 10%   | 9%    | 8%    | 24%   | 8%    | 12%   | 36%  |
| 200-299                                       | 16%      | 22% | 16% | 15%   | 15%   | 25%   | 29%   | 24%   | 20%   | 18%   | 13%   | 9%    | 8%    | 9%    | 10%   | 38%   | 29%   | 25%   | 42%  |
| 300+  | 28%      | 33% | 17% | 17%   | 42%   | 61%   | 49%   | 23%   | 14%   | 15%   | 12%   | 9%    | 6%    | 20%   | 61%   |       | 17%   | 35%   | 77%  |
| Number of Observations by Age and Cohort Size |          |     |     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
|   | Total n  | 0-4 | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+  |
| 0-29  | 4,688    | 109 | 192 | 190   | 103   | 55    | 74    | 77    | 65    | 69    | 73    | 81    | 80    | 103   | 199   | 427   | 660   | 959   | 1172 |
| 30-49   | 4,613    | 117 | 138 | 141   | 114   | 102   | 146   | 109   | 67    | 53    | 54    | 57    | 92    | 150   | 379   | 624   | 765   | 807   | 698  |
| 50-74   | 6,560    | 313 | 252 | 226   | 236   | 316   | 315   | 292   | 190   | 192   | 187   | 194   | 259   | 378   | 595   | 805   | 777   | 591   | 442  |
| 75-99   | 6,753    | 441 | 366 | 372   | 330   | 448   | 369   | 421   | 399   | 283   | 237   | 232   | 348   | 451   | 625   | 554   | 364   | 302   | 211  |
| 100-124                                       | 6,237    | 467 | 404 | 384   | 396   | 414   | 396   | 432   | 451   | 369   | 301   | 331   | 375   | 421   | 433   | 281   | 171   | 99    | 112  |
| 125-149                                       | 5,545    | 419 | 383 | 341   | 385   | 428   | 366   | 399   | 443   | 389   | 311   | 340   | 386   | 393   | 303   | 88    | 64    | 37    | 70   |
| 150-174                                       | 4,736    | 368 | 366 | 328   | 350   | 301   | 289   | 343   | 357   | 396   | 344   | 349   | 332   | 330   | 185   | 32    | 9     | 11    | 46   |
| 175-199                                       | 3,516    | 263 | 246 | 246   | 272   | 215   | 212   | 245   | 312   | 268   | 303   | 286   | 296   | 241   | 69    | 4     | 5     | 5     | 28   |
| 200-299                                       | 6,536    | 307 | 418 | 491   | 479   | 335   | 444   | 408   | 490   | 663   | 747   | 744   | 603   | 343   | 29    | 3     | 2     | 4     | 26   |
| 300+  | 1,540    | 14  | 53  | 99    | 153   | 204   | 207   | 92    | 44    | 136   | 261   | 204   | 47    | 8     | 1     |       | 1     | 3     | 13   |

**Table A9. Mean Percent Errors by Age Group and Cohort Size, MCD-Level, All Massachusetts Geographies**

Mean percent error by age group and cohort size (in 2010) Estimates produced for 2020 at age-group/sex/town level and measured against Census 2020 values. Male and female errors combined in MPEs summary.

| MPEs by Age and Cohort Size                   |          |      |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
|---|----------|------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|
| Cohort Size                                   | Avg. All | 0-4  | 5-9  | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+  |
| 0_24  | -6%      | -18% | -17% | -3%   | -23%  | -5%   | -10%  | -14%  | -28%  | -28%  | 6%    | -13%  | 4%    | -18%  | -10%  | -5%   | -3%   | 13%   | 3%   |
| 25_49   | -4%      | 10%  | -9%  | -5%   | 0%    | -19%  | -8%   | -8%   | -12%  | -15%  | 3%    | 4%    | 10%   | 1%    | -1%   | -1%   | 0%    | -4%   | -3%  |
| 50_99   | -2%      | -3%  | 0%   | 5%    | -1%   | -21%  | -6%   | -2%   | -2%   | 8%    | 10%   | 5%    | 6%    | 1%    | -3%   | -5%   | -5%   | -5%   | -1%  |
| 100_199                                       | -4%      | 0%   | -4%  | -1%   | -4%   | -19%  | -3%   | 0%    | -3%   | -9%   | -4%   | -1%   | -1%   | -3%   | -3%   | -3%   | -4%   | -6%   | 3%   |
| 200_499                                       | -4%      | 2%   | -6%  | -7%   | -5%   | -16%  | 2%    | 8%    | -5%   | -12%  | -5%   | -3%   | 0%    | -1%   | -5%   | -4%   | -7%   | -10%  | -1%  |
| 500_999                                       | -6%      | 5%   | -6%  | -9%   | -8%   | -14%  | 0%    | -1%   | -6%   | -12%  | -8%   | -4%   | -2%   | -4%   | -6%   | -7%   | -7%   | -9%   | 1%   |
| 1000_1999                                     | -6%      | 12%  | -5%  | -9%   | -9%   | -11%  | 3%    | -4%   | -7%   | -12%  | -8%   | -5%   | -4%   | -6%   | -9%   | -8%   | -10%  | -9%   | -3%  |
| 2000_9999                                     | -7%      | 9%   | -8%  | -11%  | -9%   | -14%  | -3%   | -7%   | -9%   | -5%   | -5%   | -6%   | -7%   | -7%   | -8%   | -9%   | -10%  | -16%  | -15% |
| 10000+  | -4%      | 10%  | 9%   | -3%   | -9%   | -22%  | -9%   | -6%   | -4%   | 1%    | 0%    | -4%   | -7%   | -7%   | -9%   |       |       |       |      |
| Number of Observations by Age and Cohort Size |          |      |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
|   | Total n  | 0-4  | 5-9  | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+  |
| 0_24  | 1,131    | 70   | 56   | 43    | 44    | 63    | 75    | 74    | 52    | 31    | 22    | 21    | 17    | 20    | 45    | 82    | 114   | 137   | 165  |
| 25_49   | 1,081    | 61   | 59   | 59    | 58    | 69    | 60    | 66    | 54    | 59    | 37    | 32    | 35    | 37    | 66    | 75    | 77    | 87    | 90   |
| 50_99   | 1,465    | 71   | 62   | 59    | 58    | 83    | 91    | 82    | 74    | 52    | 69    | 70    | 74    | 83    | 77    | 103   | 112   | 135   | 110  |
| 100_199                                       | 1,836    | 98   | 75   | 70    | 72    | 98    | 118   | 108   | 90    | 85    | 73    | 76    | 81    | 98    | 139   | 152   | 152   | 129   | 122  |
| 200_499                                       | 3,397    | 199  | 204  | 197   | 213   | 209   | 176   | 184   | 195   | 180   | 172   | 176   | 214   | 226   | 210   | 190   | 164   | 148   | 140  |
| 500_999                                       | 2,167    | 114  | 143  | 163   | 148   | 87    | 88    | 98    | 133   | 162   | 168   | 179   | 165   | 151   | 114   | 75    | 66    | 54    | 59   |
| 1000_1999                                     | 1,121    | 61   | 82   | 91    | 79    | 50    | 55    | 53    | 71    | 96    | 122   | 107   | 88    | 66    | 43    | 21    | 14    | 9     | 13   |
| 2000_9999                                     | 411      | 26   | 19   | 18    | 28    | 41    | 37    | 35    | 31    | 35    | 37    | 39    | 26    | 19    | 7     | 4     | 3     | 3     | 3    |
| 10000+  | 27       | 2    | 2    | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 1     |       |       |       |      |

**Table A10. Mean Percent Errors by Age Group and Cohort Size, MCD-Level, Cape and Island Geographies**

Mean percent error by age group and cohort size (in 2010) Estimates produced for 2020 at age-group/sex/town level and measured against Census 2020 values. Male and female errors combined in MPEs summary. For MCDs in Barnstable, Dukes, and Nantucket Counties.

| MPEs by Age and Cohort Size                   |          |      |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
|---|----------|------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|
| Cohort Size                                   | Avg. All | 0-4  | 5-9  | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+  |
| 0_24  | -32%     | -25% | -19% | -41%  | -57%  | -58%  | -52%  | -14%  | -24%  | -49%  | 53%   | -58%  | -24%  | -15%  | -9%   | -46%  | -68%  | -18%  | -26% |
| 25_49   | -24%     | 2%   | -26% | -18%  | -30%  | -45%  | -25%  | -19%  | -11%  | -33%  | -36%  | -20%  | -33%  | -40%  | -38%  | -18%  | -25%  | -23%  | -18% |
| 50_99   | -23%     | -7%  | -23% | -14%  | -25%  | -35%  | -37%  | -22%  | -21%  | -27%  | -46%  | -36%  | -25%  | -5%   | -20%  | -8%   | -30%  | -19%  | -31% |
| 100_199                                       | -11%     | 21%  | -2%  | 1%    | -19%  | -27%  | -17%  | 0%    | 5%    | -15%  | -13%  | -20%  | -17%  | -17%  | -15%  | -11%  | -26%  | -14%  | 3%   |
| 200_499                                       | -10%     | 7%   | -9%  | -9%   | -14%  | -24%  | 0%    | 4%    | -9%   | -20%  | -17%  | -21%  | -16%  | -12%  | -6%   | -6%   | -15%  | -12%  | -2%  |
| 500_999                                       | -5%      | 15%  | 8%   | -4%   | -10%  | -25%  | 2%    | 1%    | -1%   | -2%   | 3%    | -7%   | -6%   | -2%   | -12%  | -13%  | -15%  | -12%  | -7%  |
| 1000_1999                                     | -12%     | -2%  | -22% | -18%  | -21%  | -27%  | -5%   | -10%  | -10%  | -19%  | -10%  | -8%   | -9%   | -9%   | -9%   | -17%  | -14%  |       |      |
| 2000_9999                                     | -12%     |      |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
| -12%  |          |      |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
| Number of Observations by Age and Cohort Size |          |      |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
|   | Total n  | 0-4  | 5-9  | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+  |
| 0_24  | 98       | 7    | 5    | 4     | 6     | 6     | 6     | 6     | 6     | 4     | 4     | 4     | 4     | 4     | 4     | 6     | 6     | 7     | 9    |
| 25_49   | 71       | 5    | 6    | 7     | 5     | 5     | 2     | 3     | 2     | 2     | 2     | 2     | 2     | 1     | 2     | 2     | 5     | 10    | 8    |
| 50_99   | 126      | 9    | 11   | 7     | 8     | 13    | 11    | 11    | 9     | 7     | 2     | 1     | 1     | 1     | 6     | 11    | 9     | 5     | 4    |
| 100_199                                       | 166      | 7    | 4    | 8     | 7     | 4     | 10    | 8     | 11    | 13    | 16    | 14    | 14    | 14    | 10    | 5     | 5     | 6     | 10   |
| 200_499                                       | 209      | 13   | 13   | 10    | 11    | 12    | 10    | 12    | 10    | 9     | 11    | 10    | 11    | 11    | 12    | 15    | 15    | 13    | 11   |
| 500_999                                       | 113      | 3    | 5    | 8     | 7     | 3     | 5     | 4     | 6     | 9     | 6     | 10    | 10    | 10    | 7     | 6     | 5     | 5     | 4    |
| 1000_1999                                     | 44       | 2    | 2    | 2     | 2     | 3     | 2     | 2     | 2     | 2     | 5     | 4     | 4     | 5     | 5     | 1     | 1     |       |      |
| 2000_9999                                     | 1        |      |      |       |       |       |       |       |       |       |       | 1     |       |       |       |       |       |       |      |

**Table A11. Mean Percent Errors by Age Group and Cohort Size, MCD-Level, Balance of State**

Mean percent error by age group and cohort size (in 2010) Estimates produced for 2020 at age-group/sex/town level and measured against adjusted Census 2020 values. Male and female errors combined in MPEs summary. For MCDs in Massachusetts excluding Barnstable, Dukes, and Nantucket Counties.

| MPEs by Age and Cohort Size                   |          |      |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
|---|----------|------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|
| Cohort Size                                   | Avg. All | 0-4  | 5-9  | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+  |
| 0_24  | -4%      | -17% | -17% | 0%    | -18%  | 1%    | -7%   | -14%  | -29%  | -25%  | -4%   | -3%   | 10%   | -18%  | -10%  | -1%   | 1%    | 14%   | 5%   |
| 25_49   | -3%      | 10%  | -7%  | -4%   | 3%    | -17%  | -7%   | -8%   | -12%  | -14%  | 5%    | 6%    | 12%   | 3%    | 0%    | 0%    | 2%    | -2%   | -2%  |
| 50_99   | 0%       | -2%  | 4%   | 7%    | 3%    | -18%  | -2%   | 1%    | 0%    | 13%   | 12%   | 6%    | 7%    | 2%    | -1%   | -4%   | -3%   | -4%   | 0%   |
| 100_199                                       | -3%      | -2%  | -4%  | -1%   | -3%   | -19%  | -1%   | 0%    | -5%   | -8%   | -1%   | 4%    | 3%    | -1%   | -2%   | -3%   | -4%   | -5%   | 3%   |
| 200_499                                       | -4%      | 1%   | -5%  | -7%   | -5%   | -15%  | 2%    | 8%    | -5%   | -12%  | -4%   | -2%   | 1%    | -1%   | -5%   | -4%   | -7%   | -9%   | -1%  |
| 500_999                                       | -6%      | 4%   | -7%  | -9%   | -8%   | -14%  | 0%    | -1%   | -7%   | -13%  | -8%   | -4%   | -2%   | -4%   | -5%   | -6%   | -7%   | -9%   | 1%   |
| 1000_1999                                     | -5%      | 12%  | -4%  | -9%   | -8%   | -10%  | 3%    | -4%   | -7%   | -11%  | -8%   | -5%   | -3%   | -6%   | -8%   | -8%   | -10%  | -9%   | -3%  |
| 2000_9999                                     | -7%      | 9%   | -8%  | -11%  | -9%   | -14%  | -3%   | -7%   | -9%   | -5%   | -5%   | -6%   | -7%   | -7%   | -8%   | -9%   | -10%  | -16%  | -15% |
| 10000+  | -4%      | 10%  | 9%   | -3%   | -9%   | -22%  | -9%   | -6%   | -4%   | 1%    | 0%    | -4%   | -7%   | -7%   | -9%   |       |       |       |      |
| Number of Observations by Age and Cohort Size |          |      |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
|   | Total n  | 0-4  | 5-9  | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+  |
| 0_24  | 1,033    | 63   | 51   | 39    | 38    | 57    | 69    | 68    | 46    | 27    | 18    | 17    | 13    | 16    | 41    | 76    | 108   | 130   | 156  |
| 25_49   | 1,010    | 56   | 53   | 52    | 53    | 64    | 58    | 63    | 52    | 57    | 35    | 30    | 33    | 36    | 64    | 73    | 72    | 77    | 82   |
| 50_99   | 1,339    | 62   | 51   | 52    | 50    | 70    | 80    | 71    | 65    | 45    | 67    | 69    | 73    | 82    | 71    | 92    | 103   | 130   | 106  |
| 100_199                                       | 1,670    | 91   | 71   | 62    | 65    | 94    | 108   | 100   | 79    | 72    | 57    | 62    | 67    | 84    | 129   | 147   | 147   | 123   | 112  |
| 200_499                                       | 3,188    | 186  | 191  | 187   | 202   | 197   | 166   | 172   | 185   | 171   | 161   | 166   | 203   | 215   | 198   | 175   | 149   | 135   | 129  |
| 500_999                                       | 2,054    | 111  | 138  | 155   | 141   | 84    | 83    | 94    | 127   | 153   | 162   | 169   | 155   | 141   | 107   | 69    | 61    | 49    | 55   |
| 1000_1999                                     | 1,077    | 59   | 80   | 89    | 77    | 47    | 53    | 51    | 69    | 94    | 117   | 103   | 84    | 61    | 38    | 20    | 13    | 9     | 13   |
| 2000_9999                                     | 410      | 26   | 19   | 18    | 28    | 41    | 37    | 35    | 31    | 35    | 37    | 38    | 26    | 19    | 7     | 4     | 3     | 3     | 3    |
| 10000+  | 27       | 2    | 2    | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 2     | 1     |       |       |       |      |

**Table A12. Mean Percent Errors by Age Group and Cohort Size, Tract Level, All Massachusetts Geographies**

Mean percent error by age group and cohort size (in 2010) Estimates produced for 2020 at age-group/sex/tract level and measured against adjusted Census 2020 values. Male and female errors combined in MPEs summary.

| MPEs by Age and Cohort Size                   |          |      |     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
|---|----------|------|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|
| Cohort Size                                   | Avg. All | 0-4  | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+  |
| 0-29  | -12%     | -61% | -8% | -11%  | -35%  | -42%  | -29%  | 5%    | -1%   | 28%   | 15%   | 8%    | 1%    | -7%   | -4%   | -14%  | -11%  | -14%  | -10% |
| 30-49   | -6%      | -17% | -7% | -2%   | -16%  | -33%  | 9%    | -7%   | -14%  | -2%   | 0%    | -5%   | -8%   | -10%  | -8%   | -6%   | -7%   | -7%   | 2%   |
| 50-74   | -5%      | -7%  | -3% | -1%   | -9%   | -25%  | -1%   | 3%    | -4%   | -2%   | -3%   | -8%   | -6%   | -5%   | -6%   | -4%   | -5%   | -5%   | 2%   |
| 75-99   | -4%      | 3%   | -1% | -5%   | -8%   | -19%  | 0%    | 2%    | -7%   | -3%   | 1%    | -4%   | -6%   | -6%   | -7%   | -3%   | -3%   | -3%   | 10%  |
| 100-124                                       | -3%      | 7%   | -4% | -4%   | -7%   | -16%  | 0%    | 0%    | -4%   | -5%   | -4%   | -4%   | -4%   | -5%   | -4%   | -1%   | -3%   | -1%   | 10%  |
| 125-149                                       | -3%      | 6%   | -2% | -2%   | -7%   | -14%  | -2%   | 4%    | -4%   | -8%   | -4%   | -5%   | -4%   | -3%   | -4%   | -3%   | 2%    | 0%    | 24%  |
| 150-174                                       | -3%      | 8%   | -4% | -6%   | -7%   | -12%  | -1%   | -2%   | -5%   | -8%   | -2%   | -4%   | -2%   | -4%   | -3%   | -1%   | -1%   | 11%   | 32%  |
| 175-199                                       | -4%      | 12%  | -2% | -6%   | -8%   | -11%  | -10%  | -1%   | -6%   | -9%   | -5%   | -4%   | -1%   | -2%   | -1%   | 21%   | -2%   | 4%    | 22%  |
| 200-299                                       | -4%      | 16%  | -3% | -8%   | -7%   | -16%  | -9%   | -1%   | -3%   | -9%   | -6%   | -3%   | 0%    | -1%   | -1%   | 18%   | 14%   | 25%   | 24%  |
| 300+  | -16%     | 33%  | 1%  | -13%  | -26%  | -52%  | -34%  | -2%   | 3%    | -9%   | -6%   | -1%   | 2%    | 7%    | 61%   |       | 17%   | 35%   | 71%  |
| Number of Observations by Age and Cohort Size |          |      |     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
|   | Total n  | 0-4  | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+  |
| 0-29  | 4,953    | 130  | 208 | 206   | 118   | 70    | 90    | 97    | 77    | 78    | 82    | 91    | 90    | 113   | 211   | 439   | 675   | 980   | 1198 |
| 30-49   | 4,861    | 141  | 159 | 159   | 132   | 128   | 163   | 126   | 83    | 61    | 56    | 59    | 92    | 153   | 382   | 630   | 776   | 830   | 731  |
| 50-74   | 6,956    | 332  | 276 | 254   | 260   | 341   | 342   | 325   | 222   | 213   | 202   | 198   | 265   | 383   | 602   | 829   | 807   | 627   | 478  |
| 75-99   | 7,186    | 476  | 394 | 385   | 350   | 487   | 402   | 452   | 426   | 308   | 254   | 246   | 360   | 456   | 644   | 585   | 398   | 337   | 226  |
| 100-124                                       | 6,610    | 486  | 431 | 412   | 418   | 427   | 419   | 446   | 470   | 390   | 321   | 354   | 396   | 439   | 459   | 309   | 199   | 112   | 122  |
| 125-149                                       | 5,825    | 428  | 391 | 358   | 397   | 437   | 379   | 414   | 459   | 413   | 325   | 357   | 412   | 415   | 331   | 111   | 79    | 43    | 76   |
| 150-174                                       | 4,937    | 371  | 372 | 335   | 364   | 307   | 292   | 345   | 366   | 409   | 368   | 370   | 349   | 364   | 208   | 43    | 11    | 12    | 51   |
| 175-199                                       | 3,624    | 268  | 246 | 248   | 275   | 217   | 214   | 248   | 315   | 275   | 317   | 304   | 315   | 256   | 82    | 4     | 5     | 6     | 29   |
| 200-299                                       | 6,660    | 308  | 424 | 496   | 486   | 335   | 446   | 409   | 492   | 669   | 763   | 765   | 626   | 367   | 34    | 4     | 3     | 4     | 29   |
| 300+  | 1,560    | 14   | 53  | 101   | 154   | 205   | 207   | 92    | 44    | 138   | 266   | 210   | 49    | 8     | 1     |       | 1     | 3     | 14   |



**Table A13. Mean Percent Errors by Age Group and Cohort Size, Tract-Level, Cape and Island Geographies**

Mean percent error by age group and cohort size (in 2010) Estimates produced for 2020 at age-group/sex/tract level and measured against adjusted Census 2020 values. Male and female errors combined in MPEs summary. For tracts in Barnstable, Dukes, and Nantucket Counties.

| MPEs by Age and Cohort Size                   |          |      |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
|---|----------|------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|
| Cohort Size                                   | Avg. All | 0-4  | 5-9  | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+  |
| 0-29  | -25%     | -37% | -16% | -27%  | -44%  | -67%  | -49%  | -41%  | -47%  | -19%  | -39%  | -21%  | 10%   | 53%   | 13%   | 15%   | -8%   | -7%   | -30% |
| 30-49   | -23%     | -5%  | -23% | -30%  | -23%  | -37%  | -29%  | -32%  | -21%  | -32%  | -53%  | -20%  |       | -30%  | -2%   | -21%  | -19%  | -28%  | -13% |
| 50-74   | -14%     | 26%  | -17% | -3%   | -17%  | -35%  | -19%  | -12%  | -11%  | -25%  | -14%  | -34%  | -19%  | -23%  | -17%  | -18%  | -23%  | -14%  | -10% |
| 75-99   | -8%      | 10%  | 2%   | -13%  | -11%  | -22%  | -1%   | 13%   | -6%   | -16%  | -19%  | -21%  | -19%  | -15%  | -20%  | -10%  | -17%  | -8%   | 13%  |
| 100-124                                       | -7%      | 16%  | 0%   | -2%   | -12%  | -18%  | 6%    | -2%   | -3%   | -5%   | -17%  | -17%  | -19%  | -10%  | -14%  | -8%   | -15%  | 1%    | 9%   |
| 125-149                                       | -4%      | 22%  | -9%  | -4%   | -8%   | -11%  | -5%   | 20%   | 16%   | -9%   | -2%   | -15%  | -12%  | -12%  | -11%  | -9%   | -8%   | -15%  | 31%  |
| 150-174                                       | -6%      | 10%  | 12%  | -3%   | -10%  | -25%  | -1%   | 1%    | 11%   | -7%   | -4%   | -10%  | -9%   | -11%  | -6%   | -4%   | -19%  | 5%    | 4%   |
| 175-199                                       | -5%      | 25%  |      | 1%    | -2%   | -22%  | 9%    | 6%    | -7%   | -11%  | -3%   | -7%   | -9%   | -8%   | -1%   |       |       | -24%  | -17% |
| 200-299                                       | -3%      | 34%  | 17%  | 0%    | -3%   |       | -1%   | 28%   | 5%    | -9%   | -8%   | -8%   | -4%   | -1%   | -2%   | -11%  | -16%  |       | 3%   |
| 300+  | -5%      |      |      | -13%  | -26%  | -67%  |       |       |       | 1%    | -1%   | 2%    | 7%    |       |       |       |       |       | -6%  |
| Number of Observations by Age and Cohort Size |          |      |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
|   | Total n  | 0-4  | 5-9  | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+  |
| 0-29  | 265      | 21   | 16   | 16    | 15    | 15    | 16    | 20    | 12    | 9     | 9     | 10    | 10    | 10    | 12    | 12    | 15    | 21    | 26   |
| 30-49   | 248      | 24   | 21   | 18    | 18    | 26    | 17    | 17    | 16    | 8     | 2     | 2     |       | 3     | 3     | 6     | 11    | 23    | 33   |
| 50-74   | 396      | 19   | 24   | 28    | 24    | 25    | 27    | 33    | 32    | 21    | 15    | 4     | 6     | 5     | 7     | 24    | 30    | 36    | 36   |
| 75-99   | 433      | 35   | 28   | 13    | 20    | 39    | 33    | 31    | 27    | 25    | 17    | 14    | 12    | 5     | 19    | 31    | 34    | 35    | 15   |
| 100-124                                       | 373      | 19   | 27   | 28    | 22    | 13    | 23    | 14    | 19    | 21    | 20    | 23    | 21    | 18    | 26    | 28    | 28    | 13    | 10   |
| 125-149                                       | 280      | 9    | 8    | 17    | 12    | 9     | 13    | 15    | 16    | 24    | 14    | 17    | 26    | 22    | 28    | 23    | 15    | 6     | 6    |
| 150-174                                       | 201      | 3    | 6    | 7     | 14    | 6     | 3     | 2     | 9     | 13    | 24    | 21    | 17    | 34    | 23    | 11    | 2     | 1     | 5    |
| 175-199                                       | 108      | 5    |      | 2     | 3     | 2     | 2     | 3     | 3     | 7     | 14    | 18    | 19    | 15    | 13    |       |       | 1     | 1    |
| 200-299                                       | 124      | 1    | 6    | 5     | 7     |       | 2     | 1     | 2     | 6     | 16    | 21    | 23    | 24    | 5     | 1     | 1     |       | 3    |
| 300+  | 20       |      |      | 2     | 1     | 1     |       |       |       | 2     | 5     | 6     | 2     |       |       |       |       |       | 1    |

**Table A14. Mean Percent Errors by Age Group and Cohort Size, Tract-Level, Balance of State**

Mean percent error by age group and cohort size (in 2010) Estimates produced for 2020 at age-group/sex/tract level and measured against Census 2020 values. Male and female errors combined in MPEs summary. For tracts in Massachusetts excluding Barnstable, Dukes, and Nantucket Counties.

| MPEs by Age and Cohort Size                   |          |      |     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
|---|----------|------|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|
| Cohort Size                                   | Avg. All | 0-4  | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+  |
| 0-29  | -11%     | -65% | -7% | -10%  | -34%  | -35%  | -25%  | 19%   | 7%    | 31%   | 18%   | 10%   | 0%    | -10%  | -5%   | -14%  | -11%  | -14%  | -10% |
| 30-49   | -5%      | -20% | -5% | 1%    | -16%  | -32%  | 14%   | -4%   | -12%  | 2%    | 2%    | -4%   | -8%   | -9%   | -8%   | -6%   | -7%   | -6%   | 2%   |
| 50-74   | -4%      | -9%  | -1% | -1%   | -8%   | -24%  | 0%    | 5%    | -3%   | 0%    | -3%   | -8%   | -6%   | -5%   | -6%   | -4%   | -4%   | -5%   | 3%   |
| 75-99   | -3%      | 2%   | -1% | -5%   | -8%   | -18%  | 0%    | 1%    | -7%   | -1%   | 3%    | -3%   | -5%   | -6%   | -6%   | -3%   | -2%   | -2%   | 10%  |
| 100-124                                       | -3%      | 7%   | -4% | -4%   | -7%   | -15%  | -1%   | 0%    | -4%   | -5%   | -3%   | -3%   | -3%   | -5%   | -3%   | -1%   | -1%   | -2%   | 10%  |
| 125-149                                       | -3%      | 6%   | -2% | -2%   | -7%   | -14%  | -2%   | 3%    | -5%   | -7%   | -4%   | -4%   | -3%   | -2%   | -3%   | -1%   | 5%    | 2%    | 23%  |
| 150-174                                       | -3%      | 8%   | -4% | -6%   | -7%   | -12%  | -2%   | -2%   | -5%   | -8%   | -2%   | -3%   | -1%   | -3%   | -3%   | 0%    | 3%    | 11%   | 35%  |
| 175-199                                       | -4%      | 12%  | -2% | -6%   | -8%   | -10%  | -10%  | -1%   | -6%   | -9%   | -5%   | -4%   | -1%   | -2%   | -1%   | 21%   | -2%   | 10%   | 23%  |
| 200-299                                       | -4%      | 16%  | -3% | -8%   | -7%   | -16%  | -9%   | -1%   | -3%   | -9%   | -6%   | -3%   | 0%    | -1%   | -1%   | 28%   | 29%   | 25%   | 27%  |
| 300+  | -16%     | 33%  | 1%  | -13%  | -26%  | -52%  | -34%  | -2%   | 3%    | -9%   | -6%   | -1%   | 2%    | 7%    | 61%   |       | 17%   | 35%   | 77%  |
| Number of Observations by Age and Cohort Size |          |      |     |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |      |
|   | Total n  | 0-4  | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 85+  |
| 0-29  | 4,688    | 109  | 192 | 190   | 103   | 55    | 74    | 77    | 65    | 69    | 73    | 81    | 80    | 103   | 199   | 427   | 660   | 959   | 1172 |
| 30-49   | 4,613    | 117  | 138 | 141   | 114   | 102   | 146   | 109   | 67    | 53    | 54    | 57    | 92    | 150   | 379   | 624   | 765   | 807   | 698  |
| 50-74   | 6,560    | 313  | 252 | 226   | 236   | 316   | 315   | 292   | 190   | 192   | 187   | 194   | 259   | 378   | 595   | 805   | 777   | 591   | 442  |
| 75-99   | 6,753    | 441  | 366 | 372   | 330   | 448   | 369   | 421   | 399   | 283   | 237   | 232   | 348   | 451   | 625   | 554   | 364   | 302   | 211  |
| 100-124                                       | 6,237    | 467  | 404 | 384   | 396   | 414   | 396   | 432   | 451   | 369   | 301   | 331   | 375   | 421   | 433   | 281   | 171   | 99    | 112  |
| 125-149                                       | 5,545    | 419  | 383 | 341   | 385   | 428   | 366   | 399   | 443   | 389   | 311   | 340   | 386   | 393   | 303   | 88    | 64    | 37    | 70   |
| 150-174                                       | 4,736    | 368  | 366 | 328   | 350   | 301   | 289   | 343   | 357   | 396   | 344   | 349   | 332   | 330   | 185   | 32    | 9     | 11    | 46   |
| 175-199                                       | 3,516    | 263  | 246 | 246   | 272   | 215   | 212   | 245   | 312   | 268   | 303   | 286   | 296   | 241   | 69    | 4     | 5     | 5     | 28   |
| 200-299                                       | 6,536    | 307  | 418 | 491   | 479   | 335   | 444   | 408   | 490   | 663   | 747   | 744   | 603   | 343   | 29    | 3     | 2     | 4     | 26   |
| 300+  | 1,540    | 14   | 53  | 99    | 153   | 204   | 207   | 92    | 44    | 136   | 261   | 204   | 47    | 8     | 1     |       | 1     | 3     | 13   |

## Works Cited

---

Baker et al. 2013

A Comparative Evaluation of Error and Bias in Census Tract-Level Age/Sex-Specific Population Estimates: Component I (Net-Migration) vs Component III (Hamilton–Perry)

Jack Baker, Adelamar Alcantara, Xiaomin Ruan, Kendra Watkins, Srinivasan

Received: 2 July 2012 / Accepted: 8 July 2013 / Published online: 3 August 2013

*Springer Science+Business Media Dordrecht* 2013

Baker et al. 2012

Baker, J., Alcantara, A., Ruan, X. M., & Watkins, K. The impact of incomplete geocoding on small area population estimates. *Journal of Population Research*, 29, 91–112. 2012.

Hamilton, C., & Perry, J. (1962). A short method for projecting population by age from one decennial census to another. *Social Forces*, 41(2), 163–170.

Schroeder, J. P. (2007).

"Target-density weighting interpolation and uncertainty evaluation for temporal analysis of census data." *Geographical Analysis* 39(3), 311–335. <http://dx.doi.org/10.1111/j.1538-4632.2007.00706.x>

Smith, Tayman, and Swanson

A Practitioner's Guide to State and Local Population Projections, Stanley K. Smith, Jeff Tayman, David A.

Swanson, THE SPRINGER SERIES ON DEMOGRAPHIC METHODS AND POPULATION ANALYSIS. Series

Editor KENNETH C. LAND

Duke University, Springer Science+Business Media, Dordrecht 2013