# Long-Term Population Projections for Massachusetts Cities and Towns

Vintage 2024 Methodology Statement

May 9, 2024

UMassAmherst | Donahue Institute
Economic and
Public Policy Research

# Long-Term Population Projections for Massachusetts Cities and Towns: Vintage 2024 Methodology Statement

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


**Project Leader**
Susan Strate, Senior Research Manager

**Project Staff**
Denis McAuliffe, Research Analyst II
Meghan Flanagan, Senior Research Analyst

**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.

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 Tables

# List of Figures

# Acknowledgments

The UMass Donahue Institute wishes to acknowledge the Massachusetts Secretary of the Commonwealth for providing funding for the development of and user support for this statewide, public-use projections series.

We also wish to acknowledge the contributions of Dr. Henry Renski, Professor of Landscape Architecture and Regional Planning and Graduate Program Director (MRP) at the University of Massachusetts Department of landscape Architecture and Regional Planning, for his major contributions to the first series of Long-term Population Projections developed by UMDI in 2013 and 2015.

Finally, we wish to thank the Massachusetts Department of Transportation for their periodic support of updates to the UMDI projections series at the regional level.

# Background

From 2021 to 2023, the Massachusetts Department of Transportation (MassDOT) led an effort to update population, household, and employment projections for Massachusetts and its metropolitan planning regions for its 2023 Regional Transportation Plans. Working closely with an advisory committee of regional and state agencies and other interested stakeholders, the University of Massachusetts Donahue Institute (UMDI) and the Metropolitan Area Planning Council (MAPC) contracted with MassDOT to develop, test, and refine a variety of methods and assumptions about the components of socio-economic changes occurring throughout Massachusetts. UMDI's work focused on population and employment trends and projections while MAPC's work focused on housing growth and changes to the labor force. The Central Transportation Planning Staff (CTPS) and regional planning agencies (RPAs) provided input in all areas through regular meetings and consultations. As a result of these efforts, UMDI developed "Vintage 2022" (V2022) population projections by sex and five-year-age cohorts in ten-year increments to 2050 for each of the thirteen Metropolitan Planning Organization (MPO) regions in Massachusetts.

With additional support from the Massachusetts Secretary of the Commonwealth, UMDI used the projections created for the Regional Transportation Plan to develop city and town-level population projections in five-year intervals to 2050 and by age and sex. Consistent with the statewide transportation plan, the V2022 municipal-level projections represented an average spring/fall condition and were consistent with Census residency rules, excluding seasonal-only residents and accounting for group quarters populations --such as students in college towns -- where they reside most of the time.

In 2023 and 2024, with support from the Massachusetts Secretary of the Commonwealth, UMDI leveraged previously unavailable Census 2020 population counts by sex and 5-year age groups – namely, the 2020 Census Demographic Profile and Demographic and Housing Characteristics File (DHC) released on May 25, 2023. UMDI used this more current and detailed data to develop an updated "Vintage 2024" projections series. In the new V2024 series, UMDI also makes updates to various other the data inputs, including more recent vital statistics and updated U.S. projections, refines the methodology for some components, and makes special adjustments for some geographies.

This report details the methods, data sources, and assumptions used to develop the UMass Donahue Institute (UMDI) *Vintage 2024 Long-Term Population Projections for Massachusetts Municipalities. Areas.*

# Limitations

It is important to note that modeled projections cannot and do not purport to predict the future, but rather may serve as points of reference for planners and researchers. Like all forecasts, the UMDI projections rely on assumptions about future trends based on past and present trends that may or may not actually persist into the future. The V2024 series employs a *status-quo* model approach to predict future population change. It assumes that recently observed trends in the components of population change, including birth, death, and migration rates, will persist in future years. It is also a demographically-based model, assuming that population change is driven by births, deaths, and the persistence of historic migration rates into the future.

As suggested by the demographic-accounting framework, the V2024 projections are based on demographic components of change to the exclusion of other factors, such as housing or transportation development initiatives, large-scale institutional changes, cultural shifts, and public policy revisions. To the extent that geographically-specific birth, death, and migration trends from the last ten years reflect the development that occurred in that place over the past ten years, the V2024 projections should serve as reasonable reflections of future development should development continue at the same relative pace in that geography. Should a region's economic development outlook change dramatically, relative to other places in the state or the U.S., then the migration component in the model may no longer reflect the migration that may be anticipated in future years. An important counterpoint to the very likely possibility of future changes in migration, however, is that the strongest predictor of future population in almost all places is the population residing there today.

Factors specific to the timing of this series may also greatly impact the accuracy of the V2024 projections. For one, the projections are based on trends unfolding during what may be described as an off-trend period. The COVID-19 pandemic drastically shifted short-term trends in births and deaths -- two of the main components used as direct inputs in the UMDI population projections method -- not only in Massachusetts but around the U.S. as a whole. Secondly, the pandemic altered typical migration and immigration patterns, with an already declining trend in immigration exacerbated by the global pandemic and with a shift in domestic migration out of urban and into more rural and seasonal areas. While population data from 2020 are incorporated into the launch populations in our projections models, at the time of the V2024 series development, numerous components, including births, deaths, immigration, and migration showed signs of starting to return to pre-pandemic patterns.

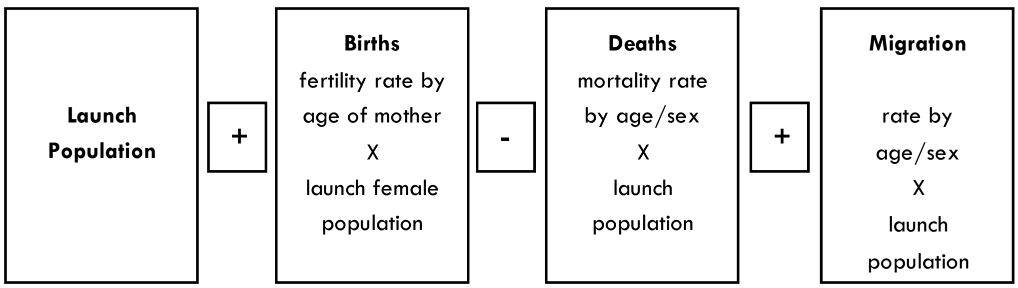
Another major consideration in the evaluation of the V2024 projections is the quality of the 2020 Census itself – a key input in the model. Along with the abrupt shift in trends that occurred around 2020, the pandemic also influenced the quality of the deccenial count, including but not limited to complications brought about by proxy respondents, limited access to some group quarters facility populations, the extension of count activities well into the summer and fall in Census "update/leave" areas including the seasonal Cape and Islands, and child undercounts in many areas. Finally, the intentional addition of "noise" by the U.S. Census Bureau to the 2020 release products create population distortions throughout the household and group quarters population and age structures reported in the DDHC data release. Not only do these distortions affect the base populations, they are also baked into the rates that are perpetuated forward in the model.

For all of these reasons, researchers should use caution when planning initiatives around the V2024 population projections, and be thoughtful about the data sources, methods, and assumptions that underpin the series. This methodology report represents UMDI's efforts to provide transparency and clarity on the inputs, methods, and assumptions used in the series so that potential users may be well informed on the components used to generate the final V2024 results.

# Methodology Overview

The UMDI Vintage 2024 (V2024) population projections are based on a demographic accounting framework for modeling population change, commonly referred to as a *cohort-component* model. The cohort-component method recognizes that there are only four ways that a region's population can change from one time-period to the next. It can add residents through either births or in-migration, or it can lose residents through deaths or out-migration. Figure 1 below displays the basic concept of a cohort-component model.

**Figure 1. Cohort-Component Model Overview**




The cohort-component approach also accounts for population change associated with the aging of the population. The current age profile is a strong predictor of future population levels, and growth and decline and can differ greatly from one region to another based on their profiles, as the likelihood of birth, death, and in- and out-migration all vary by age. For example, because fertility rates are highest among women in their twenties and thirties, a place that is anticipating a large number of women coming into their twenties and thirties in the next decade will likely experience more births. Similarly, mortality rates are notably higher for persons 70-years and older, such that an area with a large concentration of elderly residents will experience more deaths in decades to come.

The V2024 projections methodology may also be described as a "*status-quo"* projections model; it assumes that recent trends in the demographic components of population change, such as fertility, mortality, and migration by age, will persist in future periods. While it is reasonable to expect that these rates will change in future years, predicting the directionality of these trends invites additional assumptions into the model and, with them, additional uncertainty. The recent COVID-19 pandemic is an example of how an unexpected event can reverse an apparently steady component trend, with mortality rates increasing after a long period of gradual decrease in most age groups. Likewise, fertility rates have been slowing over a long period, but economic or social influences could just as readily disrupt that trend, as happened with the unforeseen "baby boom" that kicked off in the late 1940s. Fluctuations in immigration and migration are even less predictable. For example, there was a steep drop off in net immigration to Massachusetts following the 2016 elections. This trend was further exacerbated by a global pandemic in 2020, but substantially reversed again by 2023 under a new administration. For these reasons, the UMDI V2024 series may be defined strictly as "projections" and not as "scenarios" or "forecasts."

In the V2024 population projections series, UMDI uses a cohort-component model based on a combination of trends in fertility, mortality, and migration from 2010 through 2020 and decennial Census data from 2000, 2010, and 2020. The method produces population projections for three different geographic levels: municipalities, counties, and sub-state "migration" regions defined by the Census 2010 migration-PUMA (MIGPUMA) boundaries. These regional levels are controlled to one another using a "top-down" approach by which age/sex projections for smaller geographies are controlled "up" to the larger geography age/sex projections.

The "MIGPUMA" regional-level method makes use of American Community Survey sample data on migration rates by age and uses a gross, multi-regional approach in forecasting future levels of migration.[1] The county and municipal-level estimates both rely on residual net migration rates computed from vital statistics and decennial Census data. Municipal age/sex projections are controlled to the regional or county age/sex projections -- or both, depending on the region.

In the following sections, we discuss in more detail the methods, assumptions, data sources, and research considerations applied in the population projections produced for this report.

[1] PUMAs are the smallest geographic units used by the U.S. Census Bureau for reporting data taken from the detailed (micro) records of the American Community Survey (ACS) – our primary source of migration data. PUMA boundaries are defined so that they include no fewer the 100,000 persons, while Migration PUMAs (MIGPUMAs) must also incorporate the entirety of any county within their borders, leading to the aggregation of PUMAs into much larger MIGPUMAs in some areas of Massachusetts.

# Technical Discussion of Methods and Assumptions

This section provides a technical description of the process used to develop 1) sub-state regional, 2) county, and 3) municipal-level population projections. While all levels of projections are prepared using a cohort-component method, the major methodological difference is in the way migration is modeled. Both the *county* and *municipal-level* estimates (also referred to as Minor Civil Divisions, or MCDs) rely on residual net migration rates computed from vital statistics and decennial Census data. The sub-state *regional* projections use gross domestic migration rates based on the American Community Survey Public Use Microdata (ACS PUMS) for some state regions and residual net migration for other regions. MCD-level age/sex projections are controlled to age/sex projections developed for fourteen county regions and some select counties are controlled to higher-level sub-state geographic regions.[2]

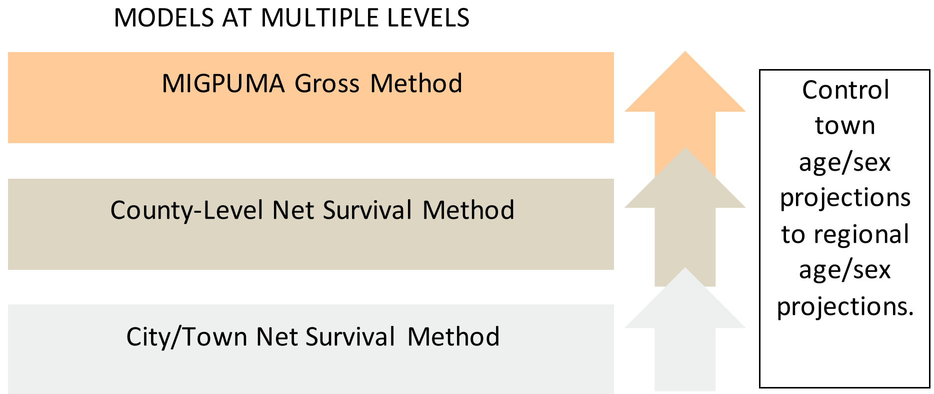
## Defining Regions and Regional Controls

The UMDI V2024 model may be described as a "top-down" projections model, with smaller, or "lower-level" geographies controlling to larger or "higher" geographies. This method is often preferred in projections modeling because it allows the estimates to take advantage of data sources oftentimes only available at the higher level and smooths out irregularities in trends caused by the smaller number of observations in small geographies. For example, birth and death data are readily available by age of mother down to the municipal level in Massachusetts, while direct measures of migration are available only at the county-level or higher. ACS PUMS migration data is preferred in some areas because it provides a direct measure of migration broken out by age and sex and allows for the calculation of gross migration rates, however in some regions the dataset combines too many counties of disparate migration trends to be useful. For this reason, the UMDI V2024 Population Projection model employs different cohort-component methods for different geographies, choosing the ideal method at each geographic level based on that level's available migration data.

Figure 2 below describes the general geographic hierarchy used in the V2024 methodology and the migration method used at each level.

[2] With the exception of Chelsea, Revere, and Winthrop in Suffolk County, which are not controlled to a higher level of geography. See the *Municipal Level Methods and Assumptions* section of this report for more information.

**Figure 2. Model Control Hierarchy and Migration Method**




## *American Community Survey MIGPUMA Regions*

For a previous projections series developed in 2018, UMDI accessed data on gross migration by age, sex, and Public-Use Microdata Sample Area (PUMA) for 52 PUMAs in Massachusetts. This allowed for the aggregation of PUMAs into areas roughly corresponding to Massachusetts planning regions, with some geographic re-controlling to account for the imperfect overlap between MPO regions and PUMA boundaries. Starting with PUMS data released in 2012, the Census Bureau changed the geographic levels for which they release migration data from PUMAs to much larger "Migration PUMAS", or "MIGPUMAs". A key feature of the MIGPUMA development was that any time a PUMA crossed over a county boundary, the total extent of both counties represented in the PUMA had to be aggregated into the same MIGPUMA. The result of this was that instead having access to 52 Massachusetts PUMAs with gross migration data last decade, the most current ACS series includes gross migration data for only five large Massachusetts MIGPUMAs.

The five sub-state MIGPUMA regions for which 2012-2019 ACS PUMS migration data are available include:

- a Berkshire MIGPUMA, which aligns with Berkshire County
- a Western Mass. MIGPUMA, which encompasses Franklin, Hampshire, and Hampden Counties
- a Suffolk MIGPUMA, which aligns with Suffolk County
- a Cape and Islands MIGPUMA, which encompasses Barnstable, Dukes, and Nantucket Counties
- an Eastern Mass. MIGPUMA, which encompasses the remaining six Massachusetts counties, including Bristol, Essex, Middlesex, Norfolk, Plymouth, and Worcester Counties.

Figure 3 below displays the most current ACS PUMA boundaries in Massachusetts as compared to county boundaries; Figure 4 displays how these counties are aggregated to encompass all county-PUMA overlaps; and Figure 5 displays the resulting MIGPUMA geography.

**Figure 3. Massachusetts PUMAs and County Boundaries for 2012-2019 ACS PUMS Data**

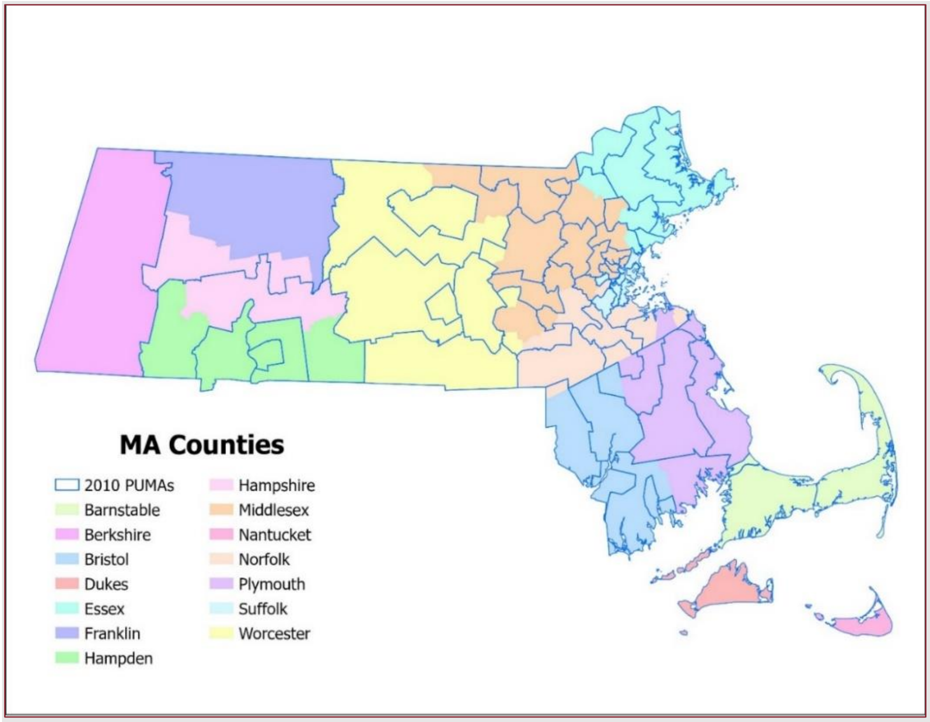



**Figure 4. Massachusetts PUMAs Aggregated to County Boundaries for 2012-2019 ACS PUMS Data**

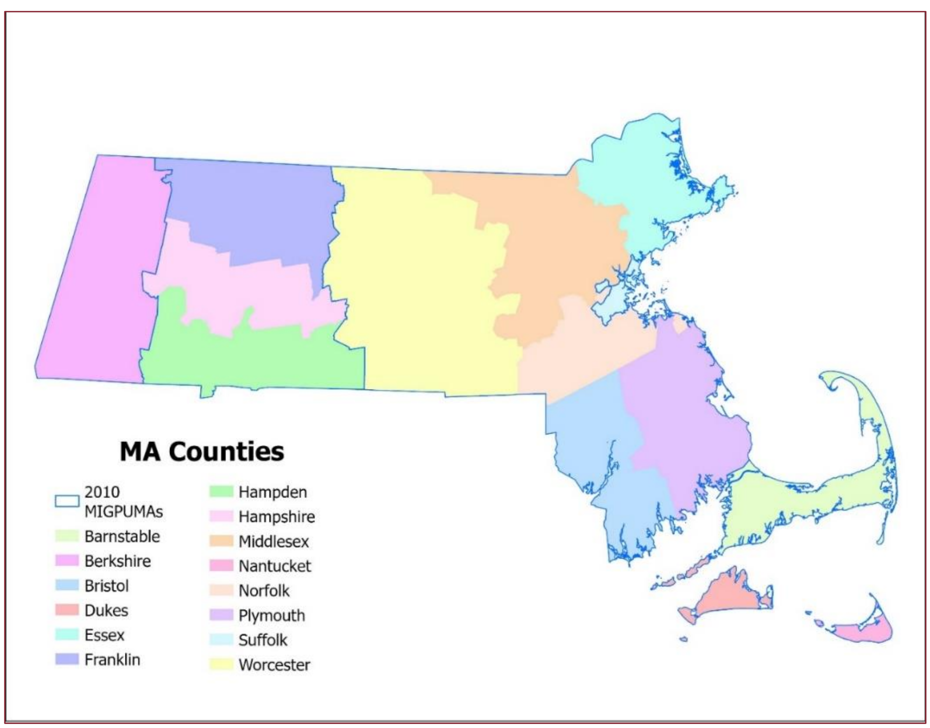
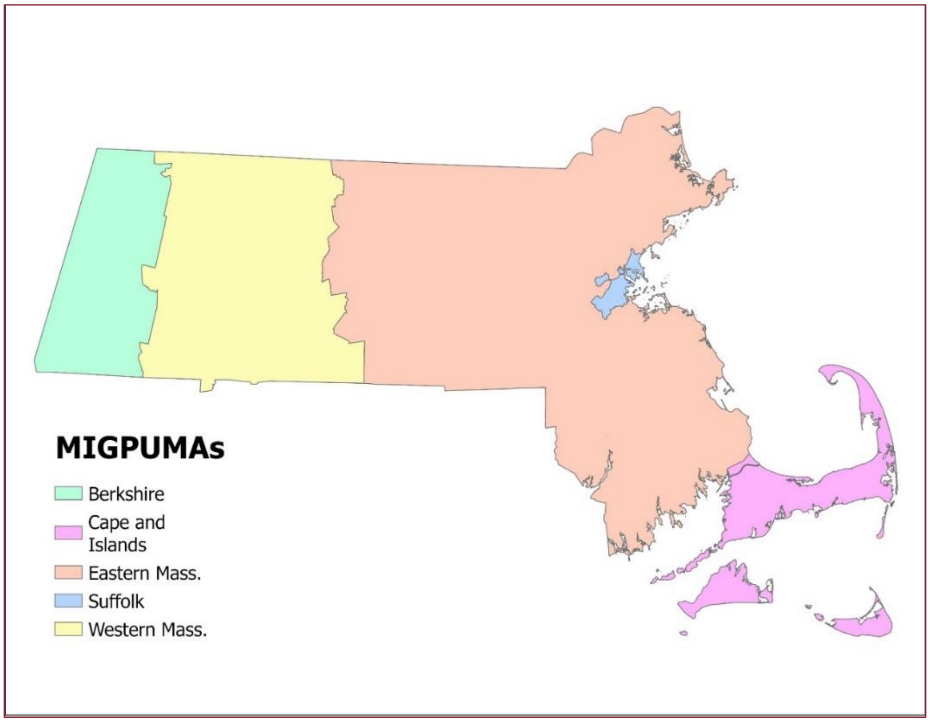


**Figure 5. MIGPUMA Boundaries for 2012-2019 ACS PUMS Data**




As seen in Figure 5 above, by aggregating counties and PUMAs in every instance of a border overlap, the resulting MIGPUMAs in some areas of the state are very large. While this is helpful in the sense that it provides a larger statistical sample for the migration question in the ACS survey, the disadvantage is that it diminishes the geographic precision of the migration trends it captures. For areas where counties and MIGPUMAs align, including Berkshire and Suffolk, the model can take advantage of the gross-migration component in the ACS PUMS data without difficulty. In areas where counties are combined into a single large MIGPUMA with similar migration-by-age patterns, as in the Eastern MIGPUMA, the model is run at both the county and MIGPUMA level, with county results controlling to the MIGPUMA results. This way, the model can leverage gross-migration data while still allowing for county-specific tendencies. For areas where counties are combined into a single large MIGPUMA but have significantly varying migration-by-age patterns, however, a projection is developed at the county level only, using a net-residual method to calculate migration rates.

## *Migration Methodology Variations, Overview*

As a broad overview, the migration methods used in the V2024 estimates for each geographic control level are as follows:

The **MIGPUMA Gross-Migration model** calculates in and out-migration separately, using in-migrants and out-migrants by age and sex estimated in the American Community Survey (ACS) Public-Use-Microdata Sample (PUMS) file for each Migration PUMA (MIGPUMA) and for averaged years 2012 through 2019. MIGPUMA regions are Census statistical regions defined by the U.S. Census Bureau and aligning with county boundaries. In Massachusetts, there are five MIGPUMA regions including: Berkshire (Berkshire County), Pioneer Valley (combining Franklin, Hampden, and Hampshire counties), Cape and Islands (Barnstable, Dukes, and Nantucket Counties) and Eastern (combining all remaining Massachusetts Counties, from Worcester County and east, excluding Cape and Islands). The benefit of this method is that it is sensitive to the interplay between a region's population and the fluctuations of the greater U.S. population. For example, if a region typically attracts in-migrants of a certain age group from other states, as that "pool" of potential migrants changes over time, migration levels into the region will also change in response. Because MIGPUMA migration estimates are derived from a sample survey (ACS), this method tends to work better in large regions, where the sample size is adequate and margins of error are reduced.

The **MIGPUMA model with a College Fix** acknowledges that college-aged populations are notoriously difficult to capture in "direct" measures of migration, including IRS tax-filing statistics as well as the American Community Survey. This may be due in part to survey response rates, confusion over where to report "usual residence", as well as generally increased mobility in this age group. The U.S. Census Bureau applies a "college-fix" in their annual county-level population estimates, which "fixes" some portion of the college-aged population in place and time. Rather than aging forward, college students are treated like a "revolving-door" population, continuously refreshed in each new interval in the same age groups, while the rest of the population migrates and ages forward. The *UMDI MIGPUMA College Fix model* applies a similar approach, with college and non-college populations modeled separately. *College-enrolled populations* are replaced at each interval by a constant "enrolled" share of the U.S. cohort that is projected for each corresponding interval while the *non-college population* is subject to aging and migration, according to gross-migration rates calculated for the non-college population by age and sex. Finally, some share of the college-enrolled population is allowed to age forward and stay in the region, with this share determined by historic cohort-change ratios.

The **County Net-Migration model** is a widely used cohort-component model that estimates net migration by age and sex in a region by using a cohort-survival-residual calculation. In this method, the 2010 Census population by age, sex and county is used as the base. Actual births and deaths experienced in the region over the 10-year period from 2010 to 2019 are added and subtracted from the base population, and the resulting, "surviving" population is aged forward 10 years to calculate an expected, or "natural increase" population. The natural increase population is then compared to the "actual" population counted at the next Census – in this case the 2020 Census – and the difference between the actual population and the natural increase population is attributed to net migration. The number of net migrants by age and sex are then used to calculate a net-migration rate for the corresponding cohort, using the regional cohort population as its denominator, and this rate is then applied to the base populations projected for each subsequent interval.

The **County Net-Migration model controlled to MIGPUMA Gross-Migration model** is a combination of the County Net Migration and MIGPUMA Gross-Migration models described above. While the Berkshire, Pioneer Valley, Suffolk, and Cape and Islands MIGPUMAs in Massachusetts are modestly sized, ranging from about 125,000 to 800,000 population in each, the Eastern Massachusetts MIGPUMA is tremendously large in comparison, with a population of almost 5 million. While the *gross-migration model* based on PUMAs is valuable for establishing connections between migrants and the rest of the U.S., the local characteristics of one sub-region over another can be washed out by controlling MCD-level projections straight up to the MIGPUMA results. For each Massachusetts county in the Eastern MIGPUMA, including Bristol, Essex, Middlesex, Norfolk, Plymouth, and Worcester, UMDI produces *county-net-migration* projections as well as a single Eastern *MIGPUMA gross-migration projection*. The age/sex/county cohort results are controlled to the Eastern MIGPUMA age/sex results at each interval. These "controlled" county-level projections are then used as the regional age/sex control totals for the MCDs in the Eastern Massachusetts region. This method effectively preserves the local migration characteristics of each county while still leveraging the benefits of the gross-migration model.

Finally, for the three municipalities that make up the balance of Suffolk County after Boston -- Chelsea, Revere, and Winthrop – UMDI uses a **net residual survival method** at the municipal level that is **not controlled** to a county or regional projection.[3]

## *Migration Methodology by Region*

UMDI produced and evaluated the above model variations for all counties before choosing which county to assign to which model. The primary determinant for choosing one model over another for any given region was the plausibility of the resulting future age profile progression compared to the actual age-over-time progression observed in previous Census counts from 2000 forward for the region. For example, if a region had tended to lose young people from Census to Census as they aged forward in time but showed a sudden reversal of that trend in the model, it indicated that the model version was less appropriate for the region. The most common example of this is seen when non-homogenous counties share a MIGPUMA. In the Cape and Islands region, for example, Nantucket County has a history of attracting large numbers of young families while Barnstable County does not, but instead attracts early retirees in large numbers. The individual characteristics of each of these counties are canceled out when both control to the same MIGPUMA model. In another example, both Franklin and Hampden Counties showed implausible boosts in the 15–19-year-old cohorts – compared to their past age profile progressions - when they were controlled to the same MIGPUMA region as Hampshire County, which maintains a perennial, large college-aged population in the region.

Two counties, Suffolk and Berkshire, correspond to MIGPUMAs that are their geographic and statistical equivalents. For these, using the MIGPUMA Gross migration model produced plausible future populations and age profiles. Other counties, including those in the Pioneer Valley and the Cape and Island regions share MIGPUMAs with other counties that have very dissimilar migration-by-age patterns. For these, the most plausible future population and age-profiles were observed in the *county-net-migration model* results. Finally, while the remaining Eastern Massachusetts counties showed differing migration by-age trends, the differences were based in the degree of migration rather than a divergent directionality of migrants as was seen in the Cape and Islands and Pioneer Valley. For these, plausible results were seen in the *county-net model controlled to the MIGPUMA-gross model*.

Table 1 below displays net migration by age rates by county from the V2022 series, using the net survival residual method, and the deviation of these county rates within their corresponding MIGPUMAs. It illustrates that migration rates by age are much more homogenous among Eastern MIGPUMA counties as compared to Cape & Islands and Pioneer Valley MIGPUMA counties. The average standard deviation of migration rates/age by county in the Cape and Island counties (0.051) was twice

[3] See the *Municipal Level Methods and Assumptions* section of this report for more information.

that of the Eastern MA counties (0.025), while the average standard deviation among Pioneer Valley Counties (0.077) was three times that of Eastern counties.

**Table 1. Migration Rates by Age and County and Deviation in County Rates within MIGPUMAs**

| Cape and Islands Net Migration Rate | | | | |
|---|---|---|---|---|
| Age | Barnstable | Dukes | Nantucket | STD DEV P |
| 1 | 0.051 | -0.019 | -0.015 | 0.032 |
| 2 | 0.036 | 0.039 | -0.057 | 0.044 |
| 3 | 0.010 | 0.025 | 0.023 | 0.006 |
| 4 | -0.020 | -0.087 | -0.073 | 0.029 |
| 5 | -0.021 | -0.025 | 0.159 | 0.086 |
| 6 | -0.059 | 0.162 | 0.484 | 0.223 |
| 7 | 0.022 | 0.087 | 0.197 | 0.072 |
| 8 | 0.041 | 0.029 | -0.047 | 0.039 |
| 9 | 0.001 | -0.025 | -0.003 | 0.011 |
| 10 | 0.049 | 0.015 | 0.054 | 0.017 |
| 11 | 0.052 | 0.003 | -0.019 | 0.030 |
| 12 | 0.075 | 0.035 | -0.016 | 0.037 |
| 13 | 0.100 | 0.032 | -0.010 | 0.045 |
| 14 | 0.116 | 0.080 | -0.002 | 0.049 |
| 15 | 0.043 | 0.014 | -0.091 | 0.057 |
| 16 | -0.008 | 0.012 | -0.087 | 0.043 |
| 17 | 0.000 | 0.026 | -0.088 | 0.049 |
| 18 | -0.106 | -0.099 | -0.007 | 0.045 |
| | **avg ST DEV by age group:** | | | **0.051** |

| Eastern MA Net Migration Rate | | | | | | | |
|---|---|---|---|---|---|---|---|
| Age | Bristol | Essex | Middlesex | Norfolk | Plymouth | Worcester | STD DEV P |
| 1 | -0.017 | 0.043 | -0.007 | 0.052 | 0.079 | 0.032 | 0.033 |
| 2 | 0.046 | 0.035 | 0.006 | 0.060 | 0.066 | 0.043 | 0.019 |
| 3 | 0.031 | 0.031 | 0.017 | 0.034 | 0.037 | 0.019 | 0.007 |
| 4 | 0.030 | 0.049 | 0.134 | 0.026 | -0.015 | 0.072 | 0.046 |
| 5 | -0.007 | -0.006 | 0.108 | -0.015 | -0.078 | -0.035 | 0.056 |
| 6 | -0.010 | 0.010 | 0.192 | 0.101 | -0.034 | -0.018 | 0.081 |
| 7 | 0.051 | 0.081 | 0.040 | 0.100 | 0.116 | 0.090 | 0.027 |
| 8 | 0.036 | 0.093 | -0.010 | 0.084 | 0.141 | 0.046 | 0.048 |
| 9 | 0.006 | 0.024 | -0.021 | 0.034 | 0.033 | 0.002 | 0.019 |
| 10 | 0.018 | 0.030 | -0.001 | 0.011 | 0.040 | 0.021 | 0.013 |
| 11 | 0.014 | 0.014 | -0.011 | 0.000 | 0.012 | 0.000 | 0.009 |
| 12 | 0.011 | 0.008 | -0.016 | -0.017 | 0.018 | 0.000 | 0.013 |
| 13 | -0.004 | -0.011 | -0.033 | -0.039 | 0.012 | -0.015 | 0.017 |
| 14 | -0.007 | -0.013 | -0.045 | -0.042 | 0.003 | -0.025 | 0.018 |
| 15 | 0.001 | -0.016 | -0.037 | -0.037 | 0.009 | -0.015 | 0.017 |
| 16 | -0.008 | 0.003 | -0.014 | -0.002 | 0.023 | 0.008 | 0.012 |
| 17 | 0.032 | 0.038 | 0.011 | 0.040 | 0.042 | 0.049 | 0.012 |
| 18 | -0.065 | -0.057 | -0.070 | -0.059 | -0.050 | -0.054 | 0.007 |
| | | | | | | | **0.025** |

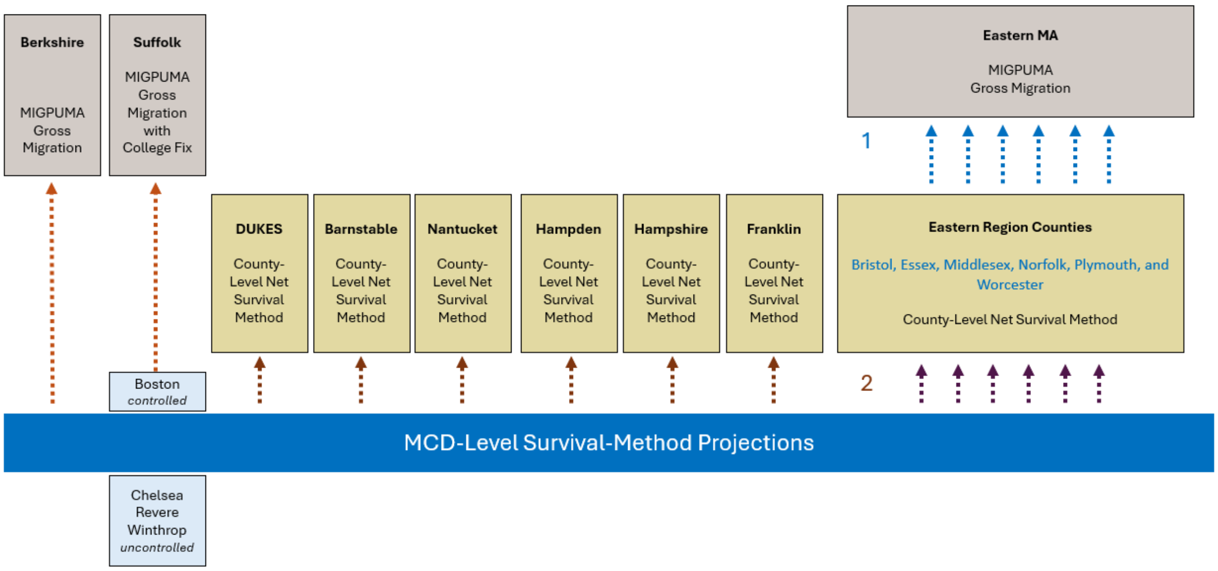
| Pioneer Valley Net Migration Rate | | | | |
|---|---|---|---|---|
| Age | Franklin | Hampden | Hampshire | STD DEV P |
| 1 | 0.006 | -0.003 | 0.040 | 0.019 |
| 2 | 0.008 | 0.010 | 0.061 | 0.025 |
| 3 | 0.005 | 0.012 | 0.040 | 0.015 |
| 4 | -0.103 | 0.071 | 1.324 | 0.636 |
| 5 | -0.077 | -0.024 | 0.410 | 0.218 |
| 6 | 0.006 | -0.063 | -0.618 | 0.279 |
| 7 | 0.041 | 0.018 | -0.056 | 0.041 |
| 8 | 0.070 | 0.011 | 0.025 | 0.025 |
| 9 | 0.009 | -0.008 | -0.023 | 0.013 |
| 10 | 0.007 | 0.002 | -0.001 | 0.003 |
| 11 | -0.031 | -0.003 | -0.010 | 0.012 |
| 12 | 0.022 | 0.001 | -0.014 | 0.015 |
| 13 | 0.016 | -0.007 | -0.016 | 0.013 |
| 14 | 0.026 | -0.027 | -0.004 | 0.022 |
| 15 | -0.018 | -0.013 | -0.004 | 0.006 |
| 16 | -0.012 | -0.004 | 0.005 | 0.007 |
| 17 | 0.007 | 0.032 | 0.030 | 0.011 |
| 18 | -0.071 | -0.079 | -0.041 | 0.016 |
| | | | | **0.077** |

After analysis and testing of alternative control schemes for each geography, UMDI developed the following scheme for the controls and migration methods used for each of the regions, also depicted in Figure 6 below:

- Berkshire County: *MIGPUMA Gross-Migration model*
- Suffolk County, Boston: *MIGPUMA Gross-Migration model with a College Fix*
- Suffolk County, balance (Chelsea, Revere, and Winthrop): *MCD-Net Migration Model, Uncontrolled*
- Pioneer Valley and Cape and Islands, including Hampshire, Hampden, and Franklin (PV) and Barnstable, Dukes, and Nantucket County (Cape and Islands): *County Net-Migration model* [4]
- Remaining Eastern MA Counties (Bristol, Essex, Middlesex, Norfolk, Plymouth, and Worcester): *County Net-Migration model controlled to MIGPUMA Gross-Migration model.*

[4] with special adjustments to Nantucket County.

**Figure 6. Regional Cohort Controls by Region**




The following sections of this report describe in more detail how population projections are modeled at the MIGPUMA, county, and municipal levels.

# Regional and County-Level Methods and Assumptions

## *Summary*

This section describes the process and data used to develop the regional population projections at both the MIGPUMA and county levels. A description of the methodology used for municipal-level projections follows in a subsequent section.

While the *Defining Regions and Regional Controls* section of this report describes the differences in how migration is modeled in each region or county, all regional models in the UMDI V2024 series share common features and basic assumptions. All regional models can be described as cohort-component models, meaning that each cohort - in this case each 5-year age group by sex by geography – is subject to cohort-specific trends in the *demographic components of population change:* fertility, mortality, and migration. All regional models are also based on the "*status quo*" assumption that recent trends in births, deaths and migration by age, sex, and region will persist in future periods. All models start with a launch population by age/sex/region in 2020 and 5-year fertility, mortality, and migration rates are applied to each launch cohort in 5-year intervals to 2050, with each 5-year age/sex projection serving as the new launch population for the subsequent interval. In addition to a 2020 age/sex/region launch population, the models require 2010 and 2015 base populations in order to calculate the migration, fertility, and mortality rates applied in the model. Sources and assumptions for all of these components are described in the following sections, including sections explaining the launch populations, survival and fertility methods common to all region types, and migration methodologies by regional type.

## *Determining the launch population and cohort classes*

**LAUNCH POPULATIONS**

The first step in the cohort-component model is to classify the composition of resident population into discrete cohorts by age and sex. Following standard practice, in the 2024 vintage series, we used five-year-age cohorts (e.g., 0-4 years old, 5-9… 80-84, and 85-and older) and developed separate profiles for males and females. These values are taken from the 2020 Decennial Census count, namely the 2020 Census Demographic Profile and Demographic and Housing Characteristics File (DHC) released on May 25, 2023. For some geographies, modifications were made to accommodate population challenges and child undercounts, the calculations of which are detailed in the MCD section of this report.

**BASE POPULATIONS FOR RATES AND RATIOS**

In addition to a launch population, population projection models also require population data to be used in rate or ratio calculations which, in this context, we are calling the base or "endpoint" populations. For the 2010 endpoint population, we take the population counts by age and sex from the decennial Census 2010 Summary File 1 (SF1) file. For the 2020 age/sex endpoint populations, we use the DHC derived launch population.

Two different data sources are used for population estimates for the year 2015, which are needed to develop some of the five-year rates used in the model. For Barnstable, Dukes, and Nantucket, the 2015 values of the Census V2020 county-level population estimates are used.[5] For the rest of the regions and counties, the average of the Census 2010 and 2020 endpoints are used. We use an alternative data source for counties in the Cape and Islands region to compensate for a disruption of typical migration patterns in 2020 due to the COVID-19 pandemic. This region experienced a large influx of people in 2020 around the time of the Census 2020 count such that an interpolation of gteh 2010 and 2020 Census counts results in an inflated 2015 population. While we accept that pandemic-related disruptions in place of residence affected the 2020 starting population in the projections series, we observed that using this off-trend 2020 population point to create a rate -- which is then applied for the next thirty years -- produces unreasonable projections, either inflating or decreasing the expected future populations to levels out of alignment with recent historic trends. Therefore, the Census estimate for 2015 was used for counties in this region to provide more reasonable denominators.

## *Deaths and Survival Rates*

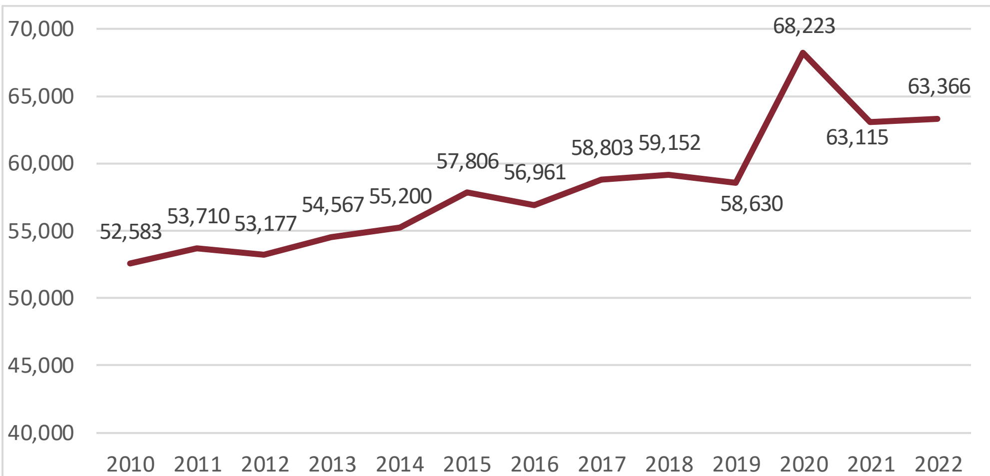
The first component of change in our regional model is survival. Our projections require an estimate of the number of people in the current population who are expected to live an additional five years into the future. Estimating the survival rate of each cohort is fairly straightforward. For the UMDI V2024 series, we develop survival rates using vital statistics data from the CDC. The CDC WONDER data includes all known deaths in the Commonwealth that occurred between 2010 to the end of calendar year 2021. This database includes information on the sex, age, and place of residence of the deceased, which we aggregate into our study regions by age/sex cohort and then develop into rates.

[5] Annual Resident Population Estimates, Estimated Components of Resident Population Change, and Rates of the Components of Resident Population Change for States and Counties: April 1, 2010 to July 1, 2020 (CO-EST2020-ALLDATA). U.S. Census Bureau Population Division. Release date: May 2021

Deaths in the first 5-year interval of the projection (2020-2025) are treated differently from later years in the projection due to off-trend deaths occuring during the COVID-19 pandemic. For the years 2020 through 2024, we estimate the number of deaths by age and sex in each study area as the sum of 2020 deaths plus four times the deaths in 2021 to create a 5-year total. In this way, this estimate includes the increased mortality seen in 2020, plus four years of more typical numbers of post-pandemic deaths from 2021.

Figure 7 below, shows the trends in deaths for Massachusetts from 2010 through 2022 from the CDC Wonder dataset.[6]

**Figure 7. Massachusetts Total Deaths 2010-2022**




For intervals 2025 through 2050, in the regional and county models, we estimate the five-year survival rate for each cohort (*j*) in study region (*i*) as one minus the average number of deaths over the past ten years (2010 to 2019) divided by the base population in 2015 and then raised to the fifth power, or:

$$(1) \quad Survival\ Rate_{i,j} = \left[1 - \left(\frac{Deaths_{i,j}}{Population_{i,j}}\right)\right]^5$$

Following the recommendations of Isserman (1993), we calculate an operational survival rate as the average of the five-year survival rates across successive age cohorts. The operational rate recognizes

[6] Source 2010-2017: Centers for Disease Control and Prevention, National Center for Health Statistics. National Vital Statistics System, Mortality 1999-2020 on CDC WONDER Online Database, released in 2021. Data are from the Multiple Cause of Death Files, 1999-2020, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program. Source 2018-2022: Centers for Disease Control and Prevention, National Center for Health Statistics. National Vital Statistics System, Mortality 2018-2022 on CDC WONDER Online Database, released in 2024. Data are from the Multiple Cause of Death Files, 2018-2022, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program.

that, over the next five years, the average person will spend half their time in their current age cohort and half their time in the next cohort. Finally, we estimate the number of eventual survivors in each cohort by multiplying the operational survival rate against the interval's launch population (for example, the 2025, 2030, 2035 launch populations) and repeat this process for each successive period. In the model, survival rates are calculated separately for each age group, sex, and county or region.

## *Migration*

Migration is the most dynamic component of change, the most difficult to estimate, and the most likely source of uncertainty and error in population projections. Whereas fertility and mortality follow fairly regular age-related patterns, the migration behavior of similar age groups is influenced by regional and national differences in socio-economic conditions. Furthermore, the data needed to estimate migration is often restricted or limited, especially for many small areas. Even when it is available, it is based on statistical samples and not actual population counts, and thus is prone to sampling error – which will be larger for smaller regions. Due to data limitations and the other methodological challenges, applied demographers have developed a variety of alternate models and methods to estimate migration rates. No single method works best in all circumstances, and we evaluated numerous approaches in the development of our projections.

**MIGPUMA GROSS-MIGRATION MODEL**

The migration approach used in the viable MIGPUMA regions (including Berkshire, Suffolk, and Eastern Massachusetts) are based on a somewhat novel approach known as a multi-region gross migration model as discussed by Isserman (1993); Smith, Tayman and Swanson (2001); and Renski and Strate (2013). Most analysts use a net migration approach, where a single net migration rate is calculated as the number of net new migrants per cohort (in-migrants minus out-migrants) divided by the baseline cohort population of the study region. Although common, the net migration approach suffers from several conceptual and empirical flaws. A major problem is that denominator of the net migration rate is based purely on the number of residents in the study region. However, none of the existing residents are at risk of migrating into the region – they already live there. While this may seem trivial, it has been shown to lead to erroneous and biased projections, especially for fast growing and declining regions.

A gross-migration approach calculates separate rates for in- and out-migrants. Beyond generating more accurate forecasts in most cases, it has an added benefit in that it connects regional population change to broader regional and national forces rather than simply treating any one region as an isolated area. This type of model is made possible by utilizing the rich detail of information available through the Public Use Micro-Samples (PUMS) of the American Community Survey (ACS). The ACS is a data product of the U.S. Census Bureau that replaced the detailed information collected on the long form of the decennial census (STF 3) in censuses prior to 2010. It asks residents questions about where they lived one year prior, which can be used to estimate the number of domestic in- and out-migrants. Unfortunately, the ACS does not report enough detail to estimate migration rates by detailed age-sex cohorts in its standard products. This information can, however, be tabulated from the ACS PUMS –

which is 5% random sample of individual records drawn from the ACS surveys[7]. In our model, we develop migration rates using averaged data from the 2012 to 2019 ACS PUMS, the most recent years available for post-2010 PUMA geographies. [8]

It is very important to realize that the PUMS records are based on small, although representative, samples, and that the smaller the sample the greater the margin of error[9]. Sample sizes can be particularly small when distributed by age and sex cohorts for different types of migrants, especially in small regions. For this reason, the Berkshire region results may be treated with more skepticism in our projections results and are subject to greater cross-examination by alternative methods[10]. The Berkshire region population averaged just 127,648 per year over the 8-year sample period, compared to an average population of 779,927 in Suffolk and 4,942,310 in the Eastern Massachusetts MIGPUMA.

To develop out-migration rates for each cohort, we take the 2012-2019 average of the cohort population living outside of the MIGPUMA region and reporting residence one-year-ago within the MIGPUMA region divided by the average 2012-2019 MIGPUMA region population. Because current residents of the study region (*i*) are those who are 'at risk' of moving out, the appropriate cohort (*j*) migration rate is:

$$(2) \quad Out\ Migration\ Rate_{i,j} = \left(\frac{OutMigrants_{i,j}}{Population_{i,j}}\right).$$

Because migration in the ACS is based on place of residence one-year prior, the out-migration rate reported in equation (2) is the equivalent of a single-year rate. We multiply this average single-year rate by five to estimate the five-year equivalent rate, and, as we did with survival rates, average the five-year rates across succeeding cohorts to craft an operational five-year rate.[11] The operational rate for each cohort is then multiplied against the number of eventual survivors in 2020 to estimate the number of likely out-migrants from the surviving population, and the process is repeated for each successive interval.

[7] To account for small or missing samples in some cohorts in some regions, we make some limited adjustments to the ACS PUMS data before calculating migration rates based on the data. In the Berkshire region, male and female migrants under the age of 15 are assigned the male/female average number of migrants before a rate is calculated to smooth out male/female ratios resulting from small sample sizes. In other regions, cohorts under age 75 with a sample size of zero in the ACS data are assigned values from the opposite gender when it is available to reduce instances of rates calculated from a null value.

[8] Due to operational issues with survey responses related to the COVID-19 in the year 2020, the U.S. Census Bureau published ACS data for 2020 on an "experimental" basis only. For more information, see: https://www.census.gov/programs-surveys/acs/data/experimental-data.html

[9] While we are aware of the potential for sampling error in using ACS PUMS data for these small regions, it is the only direct source of gross migration by age available to us currently. IRS data on migration does include gross migration data for tax-filers at the county level; however, the released data does not include age detail. The Current Population Survey, another sample survey product from the U.S. Census Bureau, provides migration data by age, but only down to the U.S. regional level of geography. Other methods commonly used to estimate migration do so using an indirect method of calculating net migration by age as a residual of a cohort-survival method

[10] For information on alternative projections methods and results for the Berkshire/Franklin regions, researchers may contact the Population Estimates Program of the UMass Donahue Institute.

[11] This differs from calculating the five-year survival rate, where the one-year rate was taken to the fifth power. Survival is modeled as a non-recurring probability since a person can only die once. However, we assume that any individual migrant could move more than once during the study period and multiply the single year rate by five to estimate a five-year equivalent.

In-migration is more challenging. The candidate pool of potential domestic in-migrants is not those currently living in the region, but people living elsewhere in the U.S. Modeling in-migration thus requires collecting data on the age-sex profile of not only the study region, but for other regions as well. We model two separate regions as possible sources of incoming migrants in the multi-regional framework - those originating in neighboring regions and states (New York, Connecticut, Rhode Island, New Hampshire, and other Massachusetts regions) and those coming from elsewhere in the U.S. By doing so, we recognize that most inter-regional migration is fairly local, and that the migration behavior of the Northeast is likely to differ considerably from that of the rest of the nation – in part due to our older and less racially diverse demographic profile.

Thus, the in-migration rates characterizing migration behavior from neighboring regions (*NE*) to study region (*i*) and from the rest of the United States (U.S.) are calculated as:

$$(3) \qquad In\ Migration\ Rate_{NE\ to\ i,j} = \left(\frac{InMigrants_{NE\ to\ i,j}}{Population_{NE,j} - Population_{i,j}}\right)$$

$$(4) \qquad In\ Migration\ Rate_{US\ to\ i,j} = \left(\frac{InMigrants_{US\ to\ i,j} - InMigrants_{NE\ to\ i,j}}{Population_{US,j} - Population_{NE,j}}\right).$$

As with the out-migration, each single-year in-migration rate is converted into a five-year operational migration rate. Unlike out-migration, these in-migration rates are not multiplied against the surviving regional population for the study region but instead the cohort population for the region of origin (neighboring regions for equation 3 or the rest of the U.S. for equation 4) to reflect the true population at risk of in-migration.

To establish the Northeast and U.S. potential in-migrant populations used in the multi-region gross migration model, UMDI uses a combination of age/sex population projections from the University of Virginia Weldon Cooper Center and the U.S. Census Bureau. The New England regional projections for New York, Connecticut, Rhode Island, New Hampshire are taken from the Weldon Cooper Center's Vintage 2018 population projections, developed for all 50 U.S. States and released in ten-year increments from 2020 through 2040. [12] UMDI modifies Weldon-Cooper's 2020 projection by controlling the age/sex results for each state to the 2020 population counts released in the Census 2020 PL-94 dataset. To project beyond 2040 for these four states, UMDI develops 2050 projections by applying a CCR method to the Weldon-Cooper 2030 and 2040 projections by age and sex, developing 2030-to-2040 change ratios for each cohort and applying these to the corresponding 2040 launch population. Child-to-Women ratios were based on 5-year-age groups from 20-24 through 35-39.[13] For the "rest-of-the-nation" population, UMDI subtracts the aggregate New England states from the U.S. Census Bureau's 2020 DHC file for 2020, and for later years, we subtract the New England states from the U.S. Census

[12] University of Virginia Weldon Cooper Center, Demographics Research Group. (2018). National Population Projections. Retrieved from https://demographics.coopercenter.org/national-population-projections

[13] 91% of U.S. fertility is accounted for in these age groups according to National Center of Health Care Statistics fertility data for 2018.

Bureau's V2023 middle-series ("Main Series") U.S. population projections to 2050 by age and sex.[14] Lastly, populations for years ending in "5" are interpolated from each ten-year projection.

**MIGPUMA MODEL COLLEGE FIX**

Tracking the migration of college students is often problematic for researchers, as neither the ACS nor conventional tax-return migration data – the two "direct measures" of migration - capture their movement comprehensively or accurately. For this reason, the U.S. Census Bureau applies a "college fix" in their annual county-level population estimates to areas that meet their criteria for percent of population enrolled in college and other population thresholds[15]. In the basic application of the "college fix", the college-enrolled population in a region is held back from aging and migration experienced by the non-college population over the specified time-period and is then restored to the region at the end of the period. In this way, the college-enrolled population remains more or less fixed for a region while other cohorts migrate and age over time.

While measuring the movement of college students is less prone to error when using *indirect* methods of calculating migration, like the net-residual survival approach used in the county-level migration model, the Suffolk MIGPUMA region in our model relies upon a *direct* measure of migration (ACS survey data in the PUMS dataset) and also includes a high share of population enrolled in college, and thus merits a college fix in its migration rates.[16]

The UMDI college fix method, like the Census Bureau's, removes the college-enrolled portion of the 15-19, 20-24, and 25-29 age cohorts from aging and migration calculations and then adds it back into its original cohort five years later. We use 2012-2019 ACS data to determine the share of population enrolled in college or graduate school in each of the age cohorts. The share is based on the region's enrolled cohort as a percent of the total U.S. cohort. We apply this share to the base-year cohort populations to estimate the regional college population and then subtract this from the total regional population. The difference is the estimated "non-college" population. This non-college population is subject to the same migration method described in the domestic migration section above, except that the migration rates are based solely on the non-college population and migrants in the ACS data. The resulting net number of non-college domestic migrants is added to each non-college cohort, which is then aged forward by five years. Finally, the enrollment share for each cohort is applied to the U.S. cohort total population in the next projected time interval to determine a new estimate of the college-enrolled population for the region. This updated college estimate is added to the projected population. Below is an example of the method applied to the 2020-to-2025 period (Figure 8).

[14] Projected Population by Single Year of Age, Sex, Race, and Hispanic Origin for the United States: 2022 to 2100 File: 2023 National Population Projections Source: U.S. Census Bureau, Population Division Release date: November 2023

[15] The "College Fix": Overcoming Issues in the Age Distribution of Population in College Counties. Ortman, Sink, King. Population Division, U.S. Census Bureau. October 2014.

[16] 32.4% of the cohorts aged 15-29 in the Suffolk County region are enrolled in college or graduate school according to ACS PUMS data for 2012-2019, averaged.

**Figure 8. College Fix Method Example**

| 2020 | | 2025 |
|---|---|---|
| non college pop 10-14 | *age 5 years and add net migrants 2020-2025→* | non-college pop 15-19 |
| college pop 15-19 | *not aged; apply % enrolled to 2025 U.S. population 15-19→* | college pop 15-19 |
| non college pop 15-19 | *age 5 years and add net migrants 2020-2025→* | non-college pop 20-24 |
| college pop 20-24 | *not aged; apply % enrolled to 2025 U.S. population 20-24→* | college pop 20-24 |
| non college pop 20-24 | *age 5 years and add net migrants 2020-2025→* | non college pop 25-29 |
| college pop 25-29 | *not aged; apply % enrolled to 2025 U.S. population 25-29→* | college pop 25-29 |
| non college pop 25-29 | *age 5 years and add net migrants 2020-2025→* | non college pop 30-34 |

Because the college population is held out of the aging process, and because migration is only captured for the non-college population, we make two additional adjustments to our model. First, we allow portions of the college-enrolled cohorts aged 15-19, 20-24, and 25-29 to age forward into the non-college population. This accounts for the college-enrolled population that ages in place into the non-college population (i.e., those that come for college and stay after graduating or un-enrolling). The share of "aging stayers" is determined in our model by calculating the historic change ratios between the non-college cohort in a particular age group to the summed college and non-college cohort populations five years younger and five years earlier.[17] Five-year ratios are calculated for the three historic time periods – 2000 to 2005; 2005 to 2010; and 2010 to 2015 – which are then averaged together and applied to future college-aged cohorts in the region to determine the age-in-place populations. Finally, we account for some portion of the region's non-college population joining the college population elsewhere upon migrating out of the region (i.e., those who leave their homes in Massachusetts to attend college elsewhere in the U.S.) by accounting for the college-enrolled out-migrants captured in the 2012-2019 PUMS data[18].

### MIGPUMA IMMIGRATION

While the ACS PUMS data provides information on gross *domestic* migration – allowing us to calculate in- and out-migrants distinctly, it cannot be used in the same way for international migration. While it captures the characteristics of recent immigrants -- defined in the survey as having a place of residence outside of the U.S. one year ago – it cannot capture *emigration* in the same way, as people who have moved out of the U.S. are no longer part of the U.S. Census survey frame. For this reason, in our regional model we instead estimate international migration as a single, net component.

17 The historic shares of college and non-college populations are determined by applying the 2012-2019 average share of population enrolled in college to each historic age/sex cohort in the 15-29 age cohorts. This method assumes that the regional college enrollment rates of by age and sex in the populations 2000 through 2015 are the same as in the 2012-2019 averages.

18 Out-migrants that are enrolled in college in regions outside of the study area with residence one year ago in the study region, as captured in the 2012-2019 ACS PUMS datasets.

Net international migration in our regional model is based on the average annual number of net international migrants estimated for each region in the U.S. Census Bureau's annual county-level population estimates series over the years 2010-2019.[19] Because the annual county-level components-of-change estimates released by the Bureau do not break the components into age/sex cohorts, we take the age/sex shares of *immigrants* reported in the averaged 2012-2019 ACS PUMS data, and apply these to the net international migrant totals for each corresponding region. This method assumes that emigrants in each region – persons leaving the U.S. for other countries – have the same age distribution as immigrants coming into each region.

Another major assumption in this method is that the number of annual net international migrants for each region will persist for the entire forecast horizon and, unlike domestic migration in our model, the estimates of net international migrants are not converted to rates. With domestic migration, we can more comfortably assume that there is a relationship between the number of migrants (our numerator) and another region (our denominator) that might be expected to remain relatively constant over time - for example the number of out-migrants relative to the region's population or the number of in-migrants relative to the U.S. population. In the case of international migration, it is harder to assume that, for example, as the world population by age increases, the region's immigrants will increase at the same rate. In reality, a great number of factors not related to any particular region's current population will influence future immigration levels, including federal immigration policy change, college recruitment policies, and political disruptions in other parts of the world -- to name just a few. Instead of trying to guess at which way these changes will affect immigration in each region, we assume that the levels experienced in recent history, in this case the 2010-2019 period, will be sustained over the full projection period. We operationalize the number of immigrants by taking the average number of immigrants of the current and next older cohort, and adding that to the survived population. The half-cohort of 0-4 migrants that would not be included in the 5-9 year old projected population due to this operationalization is added along with projected births into the projected 0-4 cohort.

**SURVIVING STAYERS**

The final step of the MIGPUMA regional migration model adds the estimated net number of domestic migrations (in-migrants minus out-migrants) and the estimated international migrants to the expected surviving population to estimate the expected number of "surviving stayers." This is an estimate of the number of current residents who neither die nor move out of the region in the coming five years, plus any new migrants to the region.

## *County-Level Net Migration Model*

As described in the *Defining Regions and Regional Controls* section of this report, some state regions are modeled using ACS PUMS data at the MIGPUMA-level (Suffolk and Berkshire), some are modeled at the county level (Cape and Island and Pioneer Valley Counties), and others are modeled at the county-level before controlling to the MIGPUMA region results (Eastern and Central Massachusetts counties in the Eastern MIGPUMA). While the model for MIGPUMAs incorporates gross-migration data from the ACS PUMS, there exists no direct source of gross migration by age at the county level or below. In the

[19] Annual Resident Population Estimates, Estimated Components of Resident Population Change, and Rates of the Components of Resident Population Change for States and Counties: April 1, 2010 to July 1, 2020 (CO-EST2020-ALLDATA). U.S. Census Bureau Population Division. Release date: May 2021.

county-level model, migration by age, sex, and county is instead estimated using a residual net migration method that relies on vital statistics and decennial Census data.

**RESIDUAL NET MIGRATION FROM VITAL STATISTICS**

The residual net migration method is used in the county model to account for the migration component of population change. "Residual" refers to the fact that migration is assumed to be responsible for past population change after accounting for births and deaths. This residual net migration is then used to estimate past migration rates. The procedure applies the resulting net migration rates by age/sex estimated for each county to the county's survived population by age/sex to project net migration by age/sex for the population ages five and older. For the population ages 0-4, child-survival ratios are calculated based on historical births-to-population trends, and are inclusive of population changes due to both migration and mortality.

**DETERMINATION OF NET-MIGRATION RATES**

Vital statistics are used to infer total net migration totals for 2010 to 2020 in five-year increments. For all counties, except for Nantucket, migrants are then converted to five-year rates and averaged together. To calculate five-year net migration by age, sex, and county, natural increase (births minus deaths) by age/sex for 2010 to 2015 is added to the 2010 population by age/sex for each county to calculate an "expected" population based on natural increase. The results are then subtracted from the interpolated 2015 population (or the "actual" population) by age/sex for each county to estimate net migration by age/sex and county from 2010 to 2015. This number of net migrants is then divided by the 2010 base population to calculate a five-year-migration rate for each cohort. This same process is used to calculate migration between 2015 and 2020, dividing the number of calculated, residual migrants over the 2015-2020 period by the 2015 base pop to create a 2015-2020 net migration rate. The two five-year net migration estimates are then averaged together and applied to the corresponding base populations by age and sex to project five-year net migration. The five-year net migration rates are held constant throughout the projection period.

In Nantucket, due to its small size and large differences in population between some cohorts, the method above leads to extreme migration rates. On the other hand, by *counts of migrants* Nantucket shows more consistent and regular trends over time. For these reasons, instead of net migration rates we use a constant net number of migrants for Nantucket as observed from 2010 to 2020 by age and sex.

The sources for the net residual migration calculations include deaths by age, sex, and county from the CDC Wonder dataset, births from the Massachusetts Department of Public Health, and Census 2010 Summary File data for the 2010 base populations by age and sex. The 2015 base population for all counties excluding Barnstable, Dukes, and Nantucket are interpolated from Census 2010 SF1 file and the 2020 DHC-derived launch population. The 2015 population from the 2020 Evaluation Estimates[20] was used for the 2015 base population for Dukes and Barnstable.

[20] Annual Resident Population Estimates, Estimated Components of Resident Population Change, and Rates of the Components of Resident Population Change for States and Counties: April 1, 2010 to July 1, 2020 (CO-EST2020-ALLDATA). U.S. Census Bureau Population Division. Release date: May 2021.

Because the residual-net-migration method accounts for all migrants over a select period – including domestic in- and out-migrants, net international migrants, and college and non-college population combined -- all of these are modeled in a single net migration rate and without applying a college-fix.

**KEY ASSUMPTIONS**

The use of a net migration rate relies on a base for migration that includes only current residents – in other words, only those at risk of out-migration. Nonresidents who are at risk of in-migration are not explicitly accounted for in the county method, and this results in some inaccuracy which is minimized by the process of controlling to regional total projections that are based on a gross migration model in those regions where control to a MIGPUMA region are reasonable. We also assume that age, sex, and county are the key factors by which migration rates vary. Other non-demographic factors, such as macroeconomic factors or local policy changes, are not explicitly included in this model. To the extent that recent, historic trends in development and economic activity are captured in the regional migration that occurred between 2010 and 2020, these factors are indirectly accounted for. Finally, this model assumes that the rates of net migration by age, sex, and region that occurred between 2010 and 2020 will persist in future years.

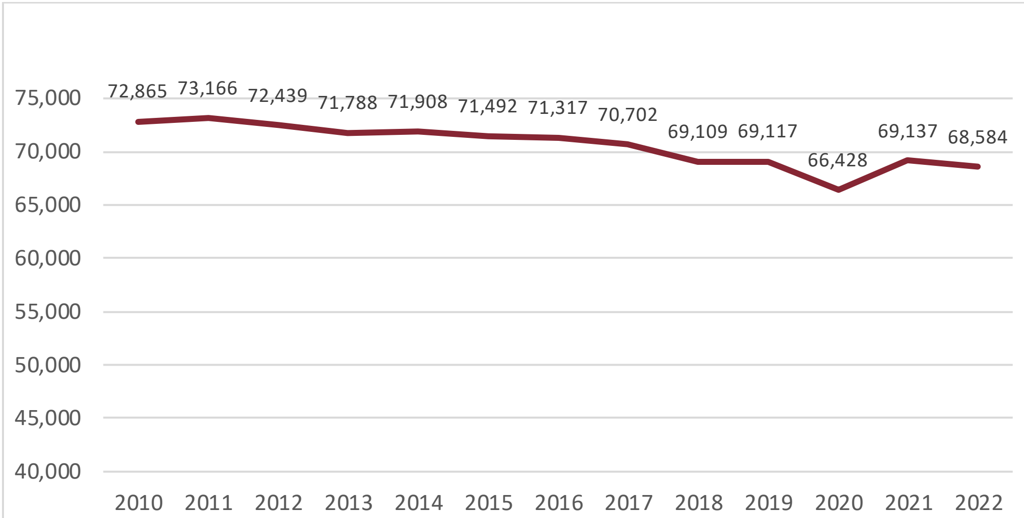
## *Births and Fertility*

The last component, in both our regional and county-level cohort-component models, involves estimating children born during the projection years. This is done in two ways, both using births by mothers' age and city or town of residence for years 2017 through 2022 and for age cohorts 15-19 through 45-49[21] provided by the Massachusetts Department of Public Health (DPH). For the 2020 to 2025 projection period, we estimate births using actual DPH births from 2020 through 2022. We take the sum of (1 x 2020 births) plus (2 x 2021 births) plus (2 x 2022 births) to synthesize five years of data. Births are further allocated into male and female shares using a multiplier of 0.512 for male births and 0.488 for female births. We use the most recent years of births in this case to account for the dramatic reduction in births in 2020 in the first-year period of projections. In later intervals, we instead use rates that exclude the exceptional 2020 year. Figure 9 below, shows the trends in births for Massachusetts from 2010 through 2022 from the CDC Wonder dataset[22].

[21] Only data for 2021 and 2022 are available for the 45-49 cohort, and rates and birth estimates for this cohort are based on those data years alone. 2022 births data are provisional.

[22] Centers for Disease Control and Prevention, National Center for Health Statistics (NCHS). National Vital Statistics System, Natality on CDC WONDER Online Database. Data are from the Natality Records 1995-2022, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program.

**Figure 9. Massachusetts Total Births 2010 - 2022**




For the remaining projection periods (2025-2050), we estimate fertility rates using data on the number of live births by age and residence of the mother for years 2017 through 2022 and age cohorts 15-19 through 45-45 provided by DPH. We exclude births in the year 2020 due to anomalies associated with the pandemic, and aggregate the births in other years for each maternal cohort by county or region. Next, an average annual number of births for each maternal cohort is calculated and multiplied by five to determine the average number of births per five-year period for each age/region cohort. The five-year average births over the 2017-2022 period are then divided by the corresponding 2020 female population by age and region to transform average births into five-year fertility rates. Each fertility rate is further allocated into male and female shares of births using the multipliers described above. Finally, the estimated fertility rates are multiplied against the average of the launch and projected female population in each of the child-bearing age cohorts. For example, the average of the 20-24 female cohort in 2020 and the projected 20-24 female cohort in 2025 would be the population to which the 20-24 fertility rate is applied. This provides an estimate of the number of infants that are anticipated within the next five years, and this number is summed across all maternal age cohorts.

## *Child-Survival Ratios*

In the V2024 series, we take an additional step to translate births into aged 0-4 child cohorts. To do this, we develop Child-Survival Ratios (CSRs) that include population changes due to the combined effects of mortality and migration for births during the course of the projection period. In prior models, we assumed that the migration of children born would be included in the mothers' migration patterns, but we found this assumption to be inaccurate in some places with large populations and high out migration among familes.

A CSR for a given period is the ratio of the decennial 0-4 single-sex cohort and the sum of the five years of single sex births prior to that decennial year.[23] In the regional and county models, CSRs were developed for each sex for periods 2005-2010 and 2015-2020, and averaged. To estimate the number of children that will "survive" in the projected 0-4 cohort after mortality and migration, the sex-specific CSR is multiplied by the total number of projected births for each sex.[24]

**AGING THE POPULATION AND GENERATING PROJECTIONS FOR LATER YEARS**

The last step in generating our first five-year forecast (for year 2025) is to age the surviving stayers in all cohorts by five years. For example, the population aged 10-14 in 2020 is "survived" according to mortality rates or deaths and migrants are added or subtracted as the cohort becomes 15-19 years old in 2025. The first (0-4) and final (85+) cohorts are treated differently. In the net county model, the number of survived births estimated in the previous step becomes the number of 0-4-year-olds in 2025. In the MIGPUMA regional model, the process is the same, plus half the 0-4 immigrants.[25] The number of persons in the 85+ cohort in 2025 is the sum of surviving stayers from the 80-84 age cohort (in 2020) plus the number of surviving stayers in the 85 and older cohort in 2020. This process is repeated for all future year projections; the 2025 projection becomes the launch population for estimating the 2030 population, which in turn is used to launch the 2035 population and so-forth.

## Municipal-Level Methods and Assumptions

### *MCD-Level Model Overview*

As described in the regional-level methods section of this report, separate projections are produced for 351 Massachusetts MCDs and for sub-state regions made up of counties or MIGPUMAs. With the exception of Suffolk County, the MCD results are then controlled to the corresponding projected regional cohorts to help smooth any inconsistences in the MCD-level results and to reflect migration trends that may be more accurately captured by the regional projection methodology.[26] While both the regional and MCD-level projections are prepared using a cohort-component method, the MCD estimates, like the county-level projections, rely on residual net-migration rates computed from vital statistics and decennial Census data.

The cohort-component method is used to account for the effects of mortality, migration, and fertility on population change. The population aged five and over is projected by the mortality and migration methods, while the population age 0-4 is projected by the fertility method and incorporation of child-survival ratios. The initial launch year is 2020, with projections made in five-year intervals from 2025 to 2050 using the previous projection as the new launch population. Projections for eighteen five-year age groups (0-4, 5-9 …80-84, and 85–and older) are reported for males and females.

---

[23] For the regional model, due to the way in which international migration is operationalized and how CSRs are applied, the "births" also include half the 0-4 cohorts of international migrants.

[24] Sex-specific CSRs are used at the county and regional levels since these geographies have populations large enough to support more specific rates that may account for small differences in male and female mortality.

[25] The other half of the 0-4 immigrants are added to the 5-9 population as they are "operationalized".

[26] In Suffolk County, Boston is controlled to the regional model, but the remaining cities of Winthrop, Revere, and Chelsea are left uncontrolled to prevent Boston from excessively influencing their population projections.

Population projections for each age and sex cohort for each five-year period are created by applying a survival rate to the base population, adding net migration for each age/sex/MCD cohort, and finally adding births by sex and mother's age, as shown in Table 2 below.

**Table 2. Projection Method by Component**

| Component | Projection |
|---|---|
| Mortality | Survived population by age/sex |
| Migration | Net migration by age/sex |
| Fertility | Births by sex and mother's age |
| Launch | 2020 DHC, with modifications for child undercounts and/or population challenges where needed; five-year projection thereafter |

## Data Sources

Data sources used in the MCD-Level population projections include Census 2020 DHC data for launch populations and to derive base populations, Census 2010 SF1 data to derive child population adjustments and base populations, and births and deaths by place of residence from 2012-2022 provided by the Massachusetts Department of Public Health.

**MCD PROJECTIONS LAUNCH POPULATION**

The initial launch population for the 2020 projection is the 2020 DHC Census population, with some modifications. We adjust the 2020 launch populations for ten MCDs that we identified as having the highest liklihood of child undercounts in the 2020 Census and also for muncipalities that submitted challenges to the 2020 Census count. The calculations used to devlop this "adjusted 2020 DHC" population are discussed below. Each projection thereafter uses the previous projection as the launch population (i.e., the 2030 projection uses the 2025 projection, and so on.)

**CHILD UNDERCOUNT ADJUSTMENTS**

In some municipalities, the 0-4 population appeared to be much lower than expected in the 2020 DHC. We reviewed various indicators, such as the size of maternal cohort, school enrollment patterns, and Census "hard to count" index, and determined that the 0-4 cohort in Boston, Brockton, Chelsea, Everett, Fall River, Holyoke, Lawrence, Lynn, New Bedford, and Springfield was likely undercounted. To attempt to rectify this, we back-calculated a new 0-4 population by applying a child-survival ratio from 2010 (based on 2005-2009 births and Census 2010 0-4 population) to births from 2016 through 2020 by MCD. Sex was estimated by applying the standard ratio for male and female births.

**POPULATION CHALLENGE CORRECTIONS**

After the 2020 Census count, UMDI assisted cities and towns with population challenges and related supporting research to address missing housing units or "group quarters" populations. These included Middlefield, Erving, Chester, Boston, Bourne, Chicopee, Dartmouth, Dedham, Franklin, Middleton, Randolph, Springfield, Wareham, and Wenham. Challenges prepared for or submitted to the Census

Bureau (regardless of acceptance status) are incorporated into the 2020 launch population. The resulting 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 quarter "type" (prisons, student dormitories, nursing homes, etc.), all based on the Census 2020 Demographic and Housing Characteristics file. The resulting population by age, sex, and geography are then added to the Census 2020 DHC data for an adjusted 2020 population. Table 3 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 8" represents population living in group homes.

**Table 3. 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 |
| 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 GQ8:87 |
| Springfield | Group Quarters | GQ1: 70 |
| Wareham | Group Quarters | GQ5: 83 |
| Wenham | Group Quarters | GQ5: 79 |

**MCD PROJECTIONS BASE POPULATION**

Where it was necessary to have a 2020 population for the calculation of model rates, the adjusted 2020 population (including population challenge and child undercount adjustments) was used. Finally, where a “midpoint” 2015 or 2017 population was needed for a rate calculation, the population was interpolated between the Census 2010[27] and the adjusted 2020 values.

## *MCD Projections: Mortality*

**ESTIMATE OF DEATHS BETWEEN 2020 AND 2025**

For the projection period of 2020 to 2025, deaths were estimated by summing one year of 2020 deaths and four years of 2021 deaths, as tabulated in the Massachusetts DPH data. For periods starting in 2025, a forward cohort suvival method is used, as described below.

**FORWARD COHORT SURVIVAL METHOD**

The forward cohort survival method is used to account for the mortality component of population change. This procedure applies five-year survival rates by age/sex to the launch population by age/sex for MCDs to survive their populations out five years, resulting in the expected population age five and over before accounting for migration. This method is used to predict populations from 2030 onward.

**FIVE-YEAR SURVIVAL RATES BY AGE/SEX**

UMDI calculates five-year survival rates by age, sex, and MCD using deaths by age, sex and MCD from 2012 to 2022 (January 1, 2012 through December 31, 2022) from the Massachusetts Department of Public Health, excluding deaths from 2020. 2020 deaths are not used for rate calculations because they were off-trend due to the COVID-19 pandemic. The formula used to develop each age/sex municipal rate for places with populations over 10,000 is the same as that used at the county and regional levels. We estimate the five-year survival rate for each cohort in study region as one minus the average number of deaths over the past ten years divided by the base population in 2017 and then raised to the fifth power. These survival rates by age, sex, and MCD are assumed to be constant for the duration of the projection horizon at the MCD level. Survival rates for each age cohort up to 80-84 are averaged with the next-older cohort to account for the fact that roughly half of each cohort would age into the next cohort over the course of each five-year period. The 85-and older cohort’s survival rate is used as-is, since there is no older cohort to average.

MCDs with smaller populations demonstrated a degree of variability in survival rates that we considered too broad for optimal results. Therefore, for MCDs with populations lower than 10,000 as of the 2010 Census, we use regional survival rates by age and sex instead of MCD-specific rates to smooth the results. [28]

[27] UMDI adjusts the 2010 population in our model for the Town of Lincoln. Lincoln was counted in Census 2010 with a greatly reduced population due to many housing units at the Hanscom Air Force Base being demolished just before the 2010 count, which were then replaced in 2011. Previous research showed using the temporarily diminished population in 2010 as a base for calculation ratios and projecting future estimates produces very unreasonable results. We substitute the 2010 cohort populations with cohort-change-ratio estimates developed using 1990 and 2000 Decennial Census data.

[28] Regions are defined as Berkshires, Cape & Islands (Barnstable, Dukes, and Nantucket counties), Eastern MA (Essex, Worcester, Plymouth, Norfolk, and Middlesex counties), Pioneer Valley (Franklin, Hampden, Hampshire counties), and Suffolk County.

### SURVIVED POPULATION FOR MCDS

The base population by age/sex for MCDs is survived to the next five-year projection by applying the corresponding averaged five-year survival rates by age/sex.

### KEY ASSUMPTIONS

The survival methodology assumes that survival rates vary most significantly by age and sex. To some extent, the use of MCD-specific rates will also indirectly account for varying socioeconomic factors, including race and ethnicity, which vary by MCD and may affect survival rates. The methodology assumes that survival rates by age, sex and MCD will stay constant over the next 30 years. We assume the estimate of deaths between 2020 and 2025 is more accurately represented by summing known deaths than using survival rates, as discussed above.

## *MCD Projections: Migration*

### RESIDUAL NET MIGRATION FROM VITAL STATISTICS

The residual net migration method is used to account for the migration component of population change. "Residual" refers to the fact that migration is assumed to be responsible for past population change after accounting for births and deaths. This residual net migration is then used to estimate past migration rates. The procedure applies the resulting net migration rates by age/sex estimated for each MCD to the MCD's survived population by age/sex to project net migration by age/sex for the population ages five and older. For the population ages 0-4, migration is captured by the child-survival ratios discussed in the fertility section.[29] For MCDs with 2010 Census population below 200, a linear migration assumption (described in the section below) is used to smooth migration.

### DETERMINATION OF NET MIGRATION RATES

Vital statistics are used to infer net migration totals for 2010 to 2019. To calculate five-year net migration by age, sex and MCD, natural increase (births minus deaths) by age/sex for 2010 to 2014 is added to the 2010 population by age/sex for each MCD. The results are then subtracted from the 2015 population by age/sex for each MCD to estimate net migration by age/sex and MCD for 2010 to 2014. A similar process calculates migration between 2015 and 2019.

For MCDs with 2020 population equal to or greater than 2,000, except for Provincetown, the two five-year net migration estimates are averaged, and rates are then calculated for each age, sex and MCD. The resulting rates are applied to the base population to project five-year net migration. The resulting average five-year net migration rates by age/sex are held constant throughout the projection period.

For MCDs with a population greater or equal to 1,000 but less than 2,000 in 2020, migration rates are calculated by the same method as MCDs with populations over 2,000. Then, migration rates for males and females in the 85+ cohorts are averaged to produce a single migration rate for the 85+ cohort for each MCD.

[29] For MCDs with total populations greater than 5,000. See the subsequent section for methodology used in towns <5,000 population.

For MCDs with a population greater or equal to 200 in 2010 but less than 1,000 in 2020, rates are created as with populations over 2,000. Then, male and female rates are averaged for each cohort to produce a single migration rate for each age cohort.

For MCDs with a 2010 population under 200, five-year net migration by age, sex and MCD is held constant, and population cohorts are never allowed to go below zero. This avoids applying unrealistically high migration rates to small populations. For instance, if an MCD starts with four males aged 70-74 and net migration shows four more move in over five years, the result is a migration rate of 2. This results in highly variable and unrealistic results in some cases. In this example, holding migration linear means that in each five-year projection period, four males aged 70-74 will move into the MCD. UMDI conducted sensitivity testing for this method and found that the model with constant migration for small places in most cases resulted in more realistic, gradual population growth or decline, as well as more realistic sex and age profiles for these MCDs.

In Provincetown, we discovered that an unbalanced male/female population and higher male residual migration rates, particularly during the pandemic, were causing unrealistic male populations in projection years. To remedy this, we adjust the migration population for Provincetown. First, migration rates are calculated as for MCDs of populations 2,000 and over. Then, after county controls are applied to all Barnstable MCDs at the end of each projection period, all age cohorts in Provincetown are redistributed to specifc male and female shares. These shares are developed by averaging the female proportion of each age cohort in 2010 and 2020 to calculate the female percentages, and subtracting these from 100% to get the male percentages. Total projected populations by age in Provincetown are then multiplied by these percentage shares to calculate male and female populations.

**KEY ASSUMPTIONS**

The use of a net migration rate relies on a base for migration that includes only current residents – in other words, only those at risk of out-migration. Nonresidents who are at risk of in-migration are not explicitly accounted for in the MCD method, and this results in some inaccuracy which is minimized by the process of controlling to regional total projections that are based on a gross migration model.

We also assume that age, sex and MCD are the key factors by which migration rates vary. Other factors, including non-demographic factors such as macroeconomic factors or local policy changes, are not explicitly included in this model. Finally, we assume that net migration by age and sex experienced in each MCD in the 2010-to-2020-period (using estimated 2020 values) will persist for the next 30 years. For the 85+ cohort in towns between 1,000 and 2,000 people, males and females are assumed to migrate at similar rates. For towns between 200 and 1,000 people, males and females of each cohort are assumed to migrate at similar rates. We assume that the percentage of the population that is male or female will be consistent with its historical values going forward in Provincetown.

## *MCD Projections: Fertility*

**VITAL STATISTICS METHOD**

The V2024 model uses vital statistics on births by MCD for the first projection interval (to 2025) and fertility rates for the remaining intervals (from 2025 to 2050). For the period between 2020 and 2024, total births are estimated from DPH data by summing one year of 2020 births, two years worth of 2021 births and two years of 2022 births to get a 5 year total for eavh MCD.

**FERTILITY RATES BY AGE OF MOTHER**

Starting in 2025 and through 2050, we apply age-specific fertility rates to the migrated and survived female population by age to project the births used to devlop the population aged 0-4. Total births are then derived by summing across all maternal age groups. For all projection years, the distribution of total births to male and female births in each MCD is assumed to be the same as the proportion of male or female births statewide.Average births by age of mother for each MCD are calculated using six maternal age groups (15-19 through 40-44). For each cohort, the sum of births over the period 2012-2022, excluding 2020, divided by two gives average births over a five-year period. This period was chosen to utilize the most recent fertility data, while not incorporating data from the year 2020, which had anomalies in fertility trends due the pandemic, into future projections.

Age-specific fertility rates are then computed by dividing the five-year average number of births by age of mother by the corresponding number of females in that age group in 2017, the midpoint of the 2012-2022 period. The 2017 population is estimated by interpolating 2010 and 2020 Census age/sex/MCD populations. The resulting age-specific fertility rates are held constant throughout the projection period and applied to the projected base population at each interval, that is: the survived, post-migration projected female population by age.

MCDs with smaller populations demonstrated a degree of variability in fertility rates that we considered too broad for optimal results. Therefore, for MCDs with populations lower than 10,000 as of the 2010 Census, we used regional[30] fertility rates by age and sex instead of MCD-specific rates to smooth the results[31].

**CHILD-SURVIVAL RATIOS**

Births generated by the model are also subject to survival rates. Depending on the size of the MCD, child-survival ratios (CSRs) are either soley mortality based, or based on the historic ratio of births to 0-4 year olds in a Census year. The latter calculation enables allowances for not only child mortality but migration as well in MCDs with sufficent population.

For municipalities over 5,000 people, child-survival ratios are developed by calculating the CSR for two periods that terminate at a decennial Census count and averaging them. For the 2005-2010 period, we take DPH births from 2005-2009 and the Census 2010 population counts of 0-4 year olds.[32]. For the 2015-2020 period, we take DPH births from 2015-2019 and the adjusted Census 2020 launch population for the 0-4 cohort. Based on data exploration which showed some unlikely variation in some MCDs

[30] Regions are defined as Berkshires, Cape & Islands (Barnstable, Dukes, and Nantucket counties), Eastern MA (Essex, Worcester, Plymouth, Norfolk, and Middlesex counties), Pioneer Valley (Franklin, Hampden, Hampshire counties), and Suffolk County.

[31] While MCDs with populations less than 10,000 are given the regional rate in this model, we make exception for "college bedroo m" towns. Because fertility rates are generally lower among females enrolled in college compared to the general population of the s ame age group, applying regional fertility rates to small towns with high percentages of college-enrolled population resulted in inflated births. We developed criteria for identifying "college bedroom" towns for an earlier (V2015) population projections se ries, and we apply town-specific fertility rates to these instead of the regional rates normally applied for very small towns. The criteria used to identify "college bedroom" towns included: population under 10,000 in 2010; >20% of 18 and over female popul ation is enrolled in college or graduate school according to 2008-2012 ACS; and the use of regional fertility rate resulted in a ≥25% Increase in the 0-4 age group from 2010 to 2015. The three MCDs subject to the "college bedroom" exception include Wenham, Sunderland, and Williamstown.

[32] Note that we modify the 2010 population for the Town of Lincoln.

between male and female Child-Survival Ratios, we determined it was best if these ratios were not sex specific at the MCD level.

For municipalities under 5,000 people, sex-specific mortality rates are used instead of CSRs. In municpalities of this size, the populations of this cohort are too small to be able to predict the direction and volume of migration. Additionally, populations will be controlled to the county totals in the final step of the model. Because these cohorts are small and individually make up very small percentages of the total county cohort population, the effects of reasonable migration fluctuations would likely be diluted by the redistribution of population after controlling to the county. In general, we would expect that assuming no in or out migration of births during the projection period for this group of towns would be a better representation of the 0-4 cohort numbers in the future.

For the ten municipalities that we determined most likely to have undercounted child populations, we replace the 2020 launch child population based off CSRs from the 2005-2010 time period. These municipalities include Boston, Brockton, Chelsea, Everett, Fall River, Holyoke, Lawrence, Lynn, New Bedford, and Springfield. We continue to use the 2010 CSRs for these municipalities for all projection intervals.

**KEY ASSUMPTIONS**

We assume age, sex and MCD to be adequate indicators of fertility rates for MCDs. We assume that the proportion of male to female births does not vary significantly by geography or maternal age. We assume that fertility rates by maternal age and MCD will not change significantly over time. We assume that child-survial ratios will not change significantly over time, and that they are equivilant for males and females. We assume that future 0-4 populations of smaller geographies are more accurately predicted by mortality rates that are not sex-specific. We assume births in the 2020-2024 period are more accurately represented by summing known births than by using a fertility rate.

## *Controlling to the Regional-level Projections*

With the exception of Suffolk County, the resulting MCD-level projected cohorts are finally controlled to the regional-level projected cohorts. To do this, we assume that each MCD's share of the region's age/sex population is given by the MCD population projections. Those shares are then applied to the regional projections to arrive at adjusted age/sex cohorts for each MCD.

In Suffolk County, Boston's final MCD population is derived from the population controlled to the Suffolk Regional model in order to incorporate patterns associated with the College Fix in that model. The remaining towns of Revere, Chelsea, and Winthrop are left uncontrolled. This is to allow the populations to develop for these cities without the influence of the more populous Boston from the regional model.

# Sources

Isserman, A. M. (1993). "The Right People, The Right Rates - Making Population Estimates with an Interregional Cohort-Component Model." Journal of the American Planning Association 59(1): 45-64.

Renski, H.C. and S. Strate. 2013. “Evaluating the migration component of county-level population estimates.” Journal of Planning Education and Research. 33(3), 325-335.

Smith, S., J. Tayman and D. Swanson. (2001) State and Local Population Projections: Methodology and Analysis. New York: Kluwer Academic. Ch. 3- 7.