OFFICE OF ECONOMIC OPPORTUNITY

Arizona State and County Population Projections, 2022-2060: Methodology Report

December 23, 2022

TABLE OF CONTENTS

1. BACKGROUND

Arizona State and County Population Projections (2022 edition) are prepared in accordance with Sections 1, 4 and 5 of Executive Order 2011-04 signed by Governor Janice Brewer:

> Section 1: The Arizona Department of Administration (ADOA) shall be the agency designated to produce the official population estimates and projections for the State of Arizona.
>
> Section 4: ADOA shall produce the official population projections for each year for a minimum of the next 25-year period. The projections shall be dated as of July 1 and shall include projections for the State, its counties, its incorporated jurisdictions, and the unincorporated balance of each county.
>
> Section 5: ADOA shall release the State and county projections as soon as possible following the release of detailed decennial census data by the U.S. Department of Commerce, Bureau of the Census, but no later than December 31 in years ending in 2. These projections shall be updated twice at three-year intervals, prior to the release of the next decennial census data and no later than December 31 in the years ending in 5 and 8.

Executive Order 2011-04 also directs the use of these projections:

> Section 10: Population estimates and projections produced by ADOA in accordance with this Executive Order shall be used by all State agencies for all purposes, including those required by federal law, which necessitates the development of population estimates or population projections.

Executive Order 2011-04 references ADOA because ADOA was the agency that housed the State Demographer's Office at the time the Executive Order was signed. The State Demographer's Office is now part of the Office of Economic Opportunity (OEO). All mentions of ADOA now apply to OEO.

2. METHODOLOGY

The Arizona Population Projections Model is a Cohort-Component model. A component methodology accounts for each aspect of demographic change (fertility, mortality, and migration). These components, each projected separately, are combined to produce population projections by age, sex, race, and ethnic group.

This model was first used in 2012 to project population for 10 race/ethnic groups in 16 geographical areas (the State of Arizona and its 15 counties) over a projection period of 40 years. In 2015, the model was updated to project population in the same 16 geographical areas, but with 6 race/ethnic groups. The groups are retained for the 2022 series. The five non-Hispanic race groups are: White, Black, Native American, Asian (including Native Hawaiian and Pacific

Islander), and Other (including two or more races). The sixth group includes Hispanic persons of all races.

2.1 The Cohort-Component Model

This version of the Cohort-Component Model (CCM) divides the population into sex and age groups or cohorts and 6 race/ethnic groups categorized by race and Hispanic origin. Movement of population from one time period to the next is accomplished by adding births and in-migration and subtracting deaths and out-migration to each cohort. Because in-migration and out-migration are usually not available as separate measures, in actual operation of the model, they are combined into "net migration" (in-migration minus out-migration), which can be positive or negative. The basic projection equations are:

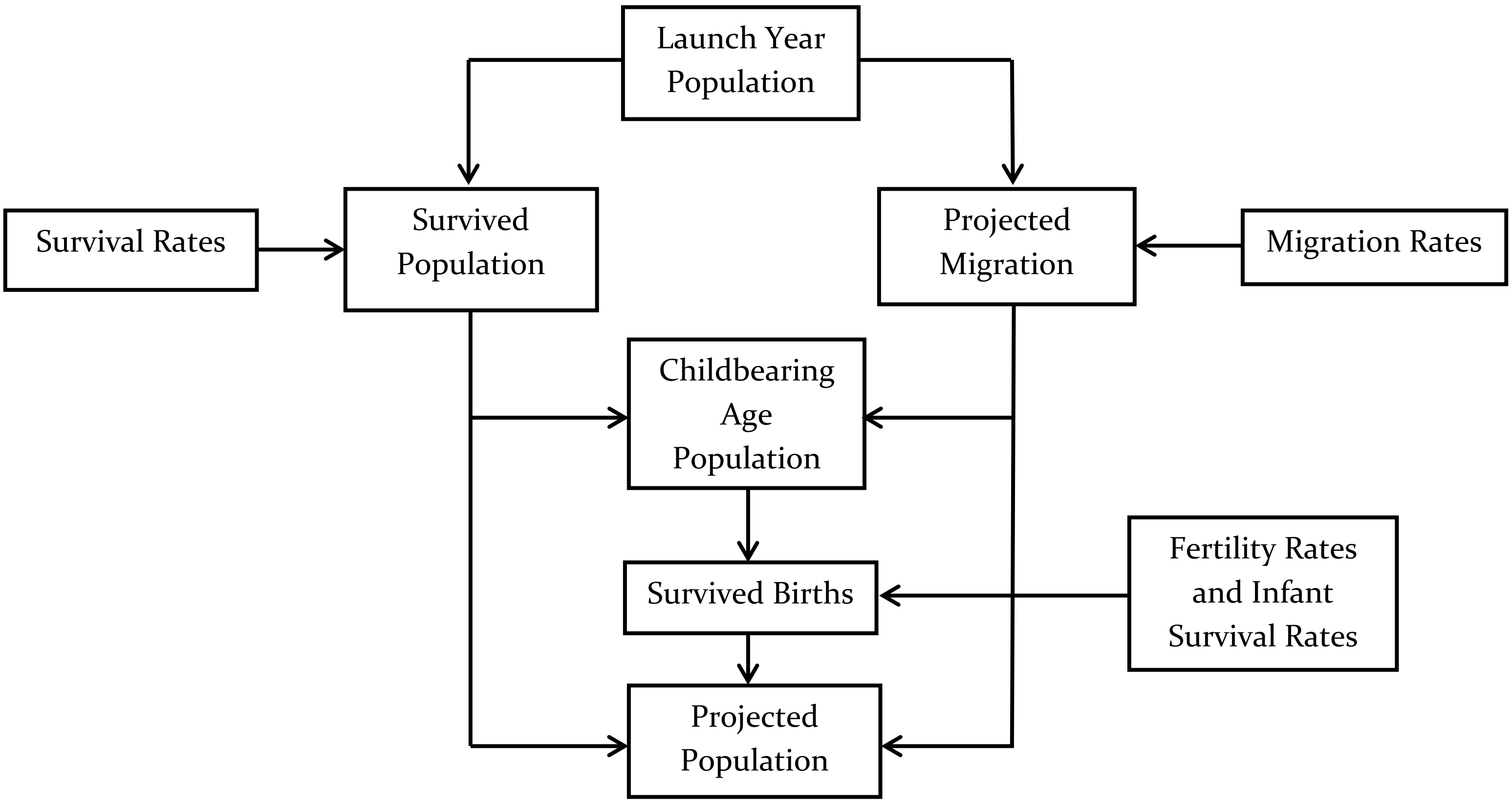
$$P_{0,S,E,t+1} = B_{S,E} - ID_{S,E} \qquad Age\ 0$$

$$P_{x+1,S,E,t+1} = P_{x,S,E,t} - D_{x,S,E} + DNM_{x,S,E} + INM_{x,S,E} \qquad Age\ 1\ to\ 84$$

$$P_{85+,S,E,t+1} = P_{84,S,E,t} + P_{85+,S,E,t} - D_{84,S,E} - D_{85+,S,E} + DNM_{84,S,E} + INM_{84,S,E} + DNM_{85+,S,E} + INM_{85+,S,E} \qquad Age\ 85+$$

x is age in the launch year; $x+1$ is age in the target year; t is launch year; $t+1$ is target year; S is the sex; E is race/ethnic group; P is total population; B is births between t and $t+1$; ID is infant deaths between t and $t+1$; D is deaths over age 1 between t and $t+1$; DNM is domestic net migration over age 1 between t and $t+1$; and INM is international net migration over age 1 between t and $t+1$.

Figure 1: Overview of the Cohort-Component Method



The basic projection cycle is the single year from year t to year $t + 1$ (projection interval). Each sex and race/ethnic group is represented by 85 single-year age groups, which range from age 0 (under 1 year of age) through age 84, plus a terminal group of ages 85+. During a projection cycle, each cohort in the launch year population is moved ahead both one year in time and one year in age. The result is the cohort aged $x + 1$ at time $t + 1$ in the projected population. Thus, the fundamental program operations consist of advancing each cohort a single year by subtracting the projected deaths and adding the projected net migration. The terminal age group (85+) projection takes into account the populations 84 and 85+ in the launch year. For the first year of life, the program develops a new cohort, the age group 0 in the projected population. The number of infant deaths reduces the births that occur during the projection interval.

The female projection is done first in order to project births, infant deaths, and population age 0 for each sex, followed by the male projection. This sequence is repeated for each race/ethnic group. Totals for both sexes are derived by adding males and females. Projections for all race/ethnic groups are determined using a bottom-up approach by summing individual race/ethnic groups. The State projection is also determined in a bottom-up manner by summing individual counties. For analytical purposes, we also create an independent State projection based on state-wide assumptions.

An actual projection involves moving the population ahead over a number of years (the projection period). This may mean deriving a current (postcensal) estimate by updating a census benchmark, or it may involve an actual projection representing a future year. In either case, the program operations are the same, and the terms launch year and target year define the populations at the beginning and at the end of the projection cycle. A projection extended over a number of years consists of a sequence of repetitive cycles (projection series), with the launch year population of one cycle representing the target year population of a previous cycle. Over each cycle, the program operations are the same. To avoid needless repetition, this document will describe the operation of a single cycle.

2.2 Component Modules of the Projections Model

The model contains four main modules: 1) mortality, 2) net migration, 3) fertility, and 4) projected population. Modules 1-3 create projections of each component of population change, while the last module uses the projected components of change to derive the target year population from the launch year population. The model retains the computations for both uncontrolled projections and projections that utilize the sex and race/ethnic-specific control totals.

Prior to entering Module 1, special populations (described in Section 4.4) are removed from the launch year population:

$$NSP_{x,S,E,t} = P_{x,S,E,t} - SP_{x,S,E,t}$$

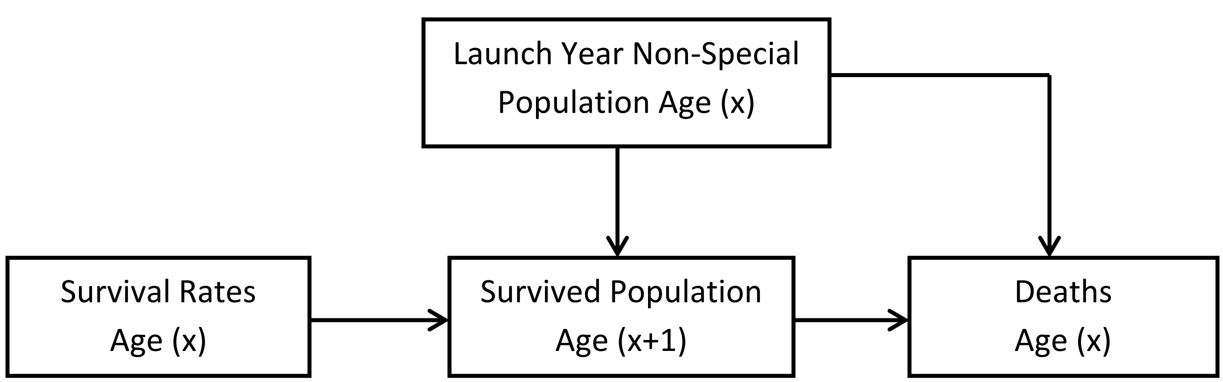
where NSP is non-special population; P is total population; and SP is special population.

SP is added back in Module 4 to complete the population projection. SP can be held constant, or the user can input an externally-derived projection $\left(SP_{x,S,E,t+1}\right)$.

2.2.1 Mortality Module

This module is devoted to computing the survived population in the target year and deaths during the projection interval as shown in Figure 2. Deaths are projected for ages 0 and older in the launch year. Infant deaths are determined in the fertility module.

Figure 2: Mortality Module



The survived population is determined by multiplying the launch year non-special population by the projected survival rate for each age group:

$$SURVNSP_{x+1,S,E,t+1} = NSP_{x,S,E,t} * SR_{x,S,E,t} \qquad (x = 0 \; to \; 83)$$

$$SURVNSP_{85+,S,E,t+1} = NSP_{84,S,E,t} * SR_{84,S,E,t} + NSP_{85+,S,E,t} * SR_{85+,S,E,t}$$

where $SURVNSP$ is survived non-special population; NSP is non-special population; and SR is launch year life table survival rate.

Deaths from t to $t+1$ are computed by subtracting the survived population age $x+1$ from the launch year population age x:

$$D_{x,S,E} = NSP_{x,S,E,t} - SURVNSP_{x+1,S,E,t+1}$$

where D is deaths to the non-special population.

For example, $D_{0,S,E}$ represents the population age 0 who did not reach age 1 in the target year and $D_{85+,S,E}$ represents the population age 85 and older who did not reach age 86 and older in the target year.

If a control for deaths is implemented, a ratio is used to adjust the age-specific deaths computed above, along with the infant deaths. The adjusted deaths will sum to the control within rounding error. The adjustment routine for deaths is:[1]

$$DFAC_{S,E} = \frac{DEATHS_{S,E}}{\sum D_{x,S,E} + ID_{S,E}}$$

$$CNTD_{x,S,E} = D_{x,S,E} * DFAC_{S,E}$$

$$CNTID_{S,E} = ID_{S,E} * DFAC_{S,E}$$

$$DEATHS_{S,E} \approx \sum CNTD_{x,S,E} + CNTID_{S,E}$$

$$(Unless\ indicated, \sum.\ is\ the\ sum\ over\ all\ Age\ Groups\ in\ all\ equations).$$

where $DFAC$ is the adjustment factor; $DEATHS$ is the control, ID is infant deaths; $CNTD$ is controlled deaths by age; and $CNTID$ is controlled infant deaths.

---

[1] Infant deaths determined from births based on ASBRs depend on female deaths used to compute the female population at-risk of having a child. Therefore, any adjustment to deaths would result in a new projection of infant deaths. In other words, adjustment of deaths requires infant deaths, and infant deaths depend on the adjustment of deaths. This issue is covered in the fertility module section.

2.2.2 Net Migration Module

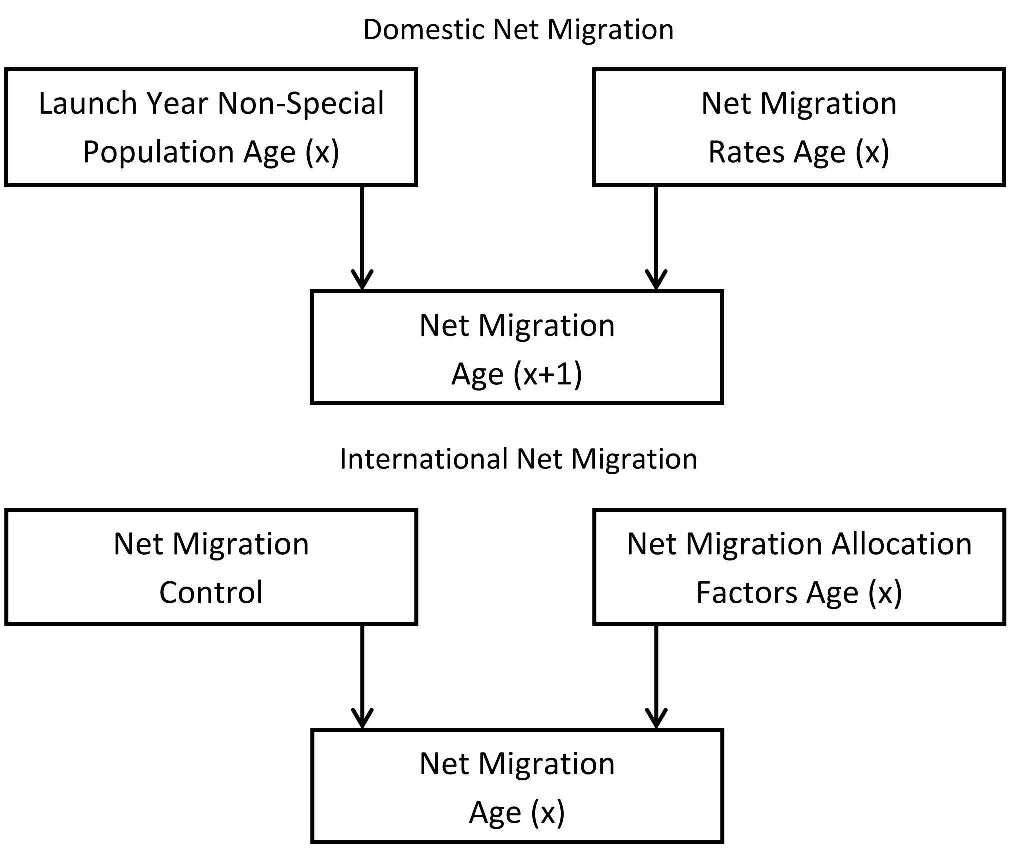
This module is devoted to computing the impact of domestic and international migration on population change as shown in Figure 3. Net migration is projected for ages 0 and older in the launch year.[2]

Population change due to domestic migration from t to $t+1$ is determined by multiplying the launch year non-special population by the projected net migration rate for each age group[3]:

$$DNM_{x,S,E} = DNMRATE_{x,S,E} * NSP_{x,S,E,t}$$

where DNM is projected domestic net migration of the non-special population and $DNMRATE$ the launch year domestic net migration rate. For example, $DNM_{15,S,E}$ represents the population change due to domestic migration of the non-special population aged 15 in the launch year who are 16 in the target year.

Figure 3: Net Migration Module



[2] The assumption is that infants born during the projection interval move with the parent(s) and are not projected separately.

[3] An alternative approach is to apply the net migration rate to the survived launch year population. As long as the migration rates are computed using the population at the beginning of the migration interval, applying them to the launch year population is acceptable and is easier to implement.

If a control for domestic net migration is implemented, the Plus-Minus method is used to adjust the age-specific domestic net migration computed above. Unlike deaths, domestic net migration can have positive and negative values across age groups. In this case, two adjustment factors should be used to account for the positive and negative values separately. The Plus-Minus method will also work when the domestic net migration has the same sign for every age group. The adjusted domestic net migration will sum to the control within rounding error. The adjustment routine[4] for domestic net migration is:

$$ABSUM_{S,E} = \sum \left| DNM_{x,S,E} \right|$$

$$SUM_{S,E} = \sum DNM_{x,S,E}$$

$$PFAC_{S,E} = \frac{ABSUM_{S,E} + (DNETMIG_{S,E} - SUM_{S,E})}{ABSUM_{S,E}}$$

$$NFAC_{S,E} = \frac{ABSUM_{S,E} - (DNETMIG_{S,E} - SUM_{S,E})}{ABSUM_{S,E}}$$

$$IF\ DNM_{x,S,E} \geq 0 \quad then\ CNTDNM_{x,S,E} = DNM_{x,S,E} * PFAC_{S,E}$$

$$IF\ DNM_{x,S,E} < 0 \quad then\ CNTDNM_{x,S,E} = DNM_{x,S,E} * NFAC_{S,E}$$

$$DNETMIG_{S,E} \approx \sum CNTDNM_{x,S,E}$$

where $ABSUM$ is sum of the absolute value of the uncontrolled domestic net migration estimates; SUM is sum of the uncontrolled domestic net migration estimates; $DNETMIG$ is the control; $PFAC$ is the adjustment factor for positive values; and $NFAC$ is the adjustment factor for negative values.

International net migration is projected differently than domestic net migration. Instead of international net migration rates, allocation factors are used to distribute the international net migration control to age groups. The factors represent the proportion of the projected international migration in each age group. If no control is supplied, the international migration projection will be zero. The projection of international net migration is:

$$If\ INETMIG_{S,E} = 0\ then\ INM_{x,S,E} = 0$$

$$If\ INETMIG_{S,E} \neq 0\ then\ INM_{x,S,E} = INETMIG_{S,E} * AFAC_{x,S,E}$$

$$INETMIG_{S,E} \approx \sum INM_{x,S,E}$$

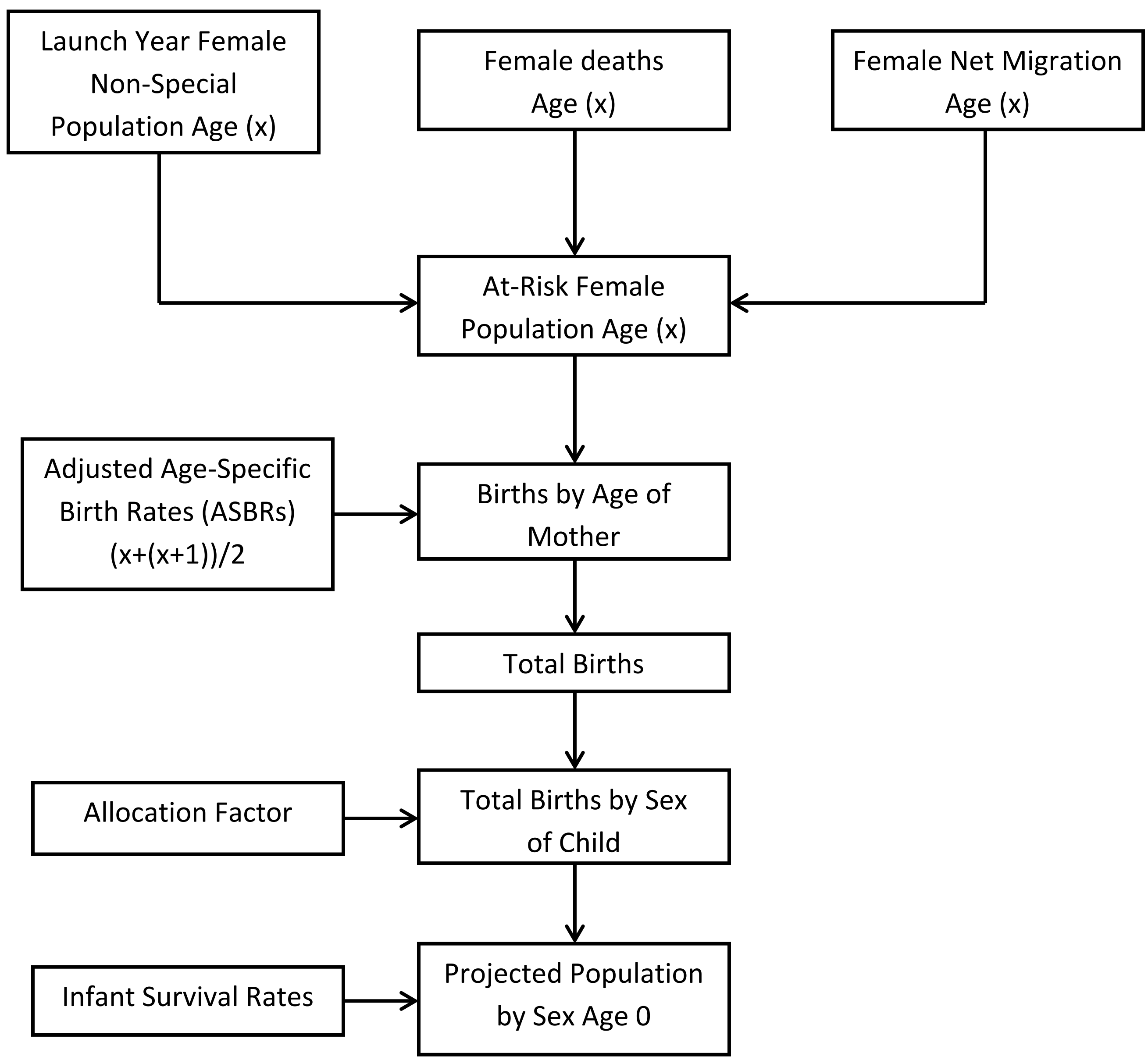
[4] This routine has been improved to remedy negative factors that can occasionally occur when the control value and estimated value are very different in magnitude.

where $INETMIG$ is the control; INM is projected international migration; and $AFAC$ is international migration age allocation factor.

2.2.3 Fertility Module

The fertility module is devoted to projecting births and infant deaths during the projection interval and the population age 0 in the target year for males and female as shown in Figure 4.

Figure 4: Fertility Module



This is accomplished in three steps. First, the at-risk female population by age is multiplied by the launch year ASBRs; these results are summed to obtain total births (*at-risk* means females of childbearing age). Second, total births are allocated between males and females using proportions. Finally, using infant survival rates, births are survived to obtain the projection for age 0 and infant deaths. These computations are:

$$ADJASBR_{x,E,t} = \frac{\left(ASBR_{x,E,t} + ASBR_{x+1,E,t}\right)}{2} \qquad (x = 14 \; to \; 44)$$

$$ATRISKFNSP_{x,E,t} = FNSP_{x,E,t} - (0.5 * FD_{x,E}) \pm FCNTDNM_{x,E} \pm FINM_{x,E}$$

$$B_{x,E} = ADJASBR_{x,E,t} * ATRISKFNSP_{x,E,t}$$

$$B_E = \sum B_{x,E}$$

$$MB_E = B_E * PBM_E$$

$$FB_E = B_E * PBF_E$$

$$MNSP_{0,E,t+1} = MB_E * MINFS_{x,E}$$

$$FNSP_{0,E,t+1} = FB_E * FINFS_{x,E}$$

$$MID_E = MB_E - MNSP_{0,E,t+1}$$

$$FID_E = FB_E - FNSP_{0,E,t+1}$$

where $ADJASBR$ is adjusted birth rate; $ATRISKFNSP$ is at-risk female population; $FNSP$ is female non-special population; FD is uncontrolled female deaths; $FCNTDNM$ is controlled female domestic net migration; $FINM$ is female international net migration; B is births; MB is male births; FB is female births; PBM is proportion of births that are male; PBF is proportion of births that are female; $MNSP$ is male non-special population age 0 in the target year; $FNSP$ is female non-special population age 0 in the target year; $MINFS$ is male infant survival rate; $FINFS$ is female infant survival rate; MID is male infant deaths; and FID is female infant deaths.

Several of these equations require further elaboration. Females pass from one age group into another during the projection interval. Because they spend half the projection interval in one group and half in the next higher age group (on average), the proper ASBR is the average rate of these two groups. In addition, some of the original members of the cohort die and others move in or move out. The launch year female non-special population in each age group is reduced by one-half of the uncontrolled deaths, because it is assumed these women live throughout half of the projection interval.[5] The at-risk population is further adjusted by adding or subtracting the

[5] Using the uncontrolled deaths eliminates the circularity involved in estimating infant deaths when the control for deaths is instituted, as described in Footnote 1. A more precise approach would use an iterative algorithm in this situation, but there is no substantive difference in the projection using uncontrolled deaths and it keeps the programming much simpler. In practice, controls for deaths are generally instituted during the postcensal period prior to the launch year and not over the projection horizon.

female domestic and international net migration. The domestic and international net migration is not reduced by one-half, like the deaths, but they are at risk of moving for one-half of the projection interval. The key issue is not where the baby is born, but where it is in the target year. A reasonable assumption is the baby will move with the parent(s) regardless of where the baby was born. By linking the baby and parents in this manner, the birth and migration of infants is treated in one-step.

When a control for births is implemented, it replaces the total births determined using ASBRs and the at-risk female population. The controlled births are split into males and females, survived to age zero in the target year, and used to compute infant deaths based on the equations shown above.

2.2.4 Projected Population Module

The projected population module combines the launch year non-special population with the results from the fertility, mortality, and net migration modules to develop the target year non-special projection by age. The special population projection is then added to complete the projection process as shown in Figure 5. The default for special population is to hold it constant at launch year values, but an independent projection can be defined for any year. The equations for this module are:

$$NSP_{0,S,E,t+1} = B_{S,E} - ID_{S,E} \qquad Age\ 0$$

$$NSP_{x,S,E,t+1} = NSP_{x,S,E,t} \pm DNM_{x,S,E} \pm INM_{x,S,E} \qquad (x = 1\ to\ 85+)$$

$$SP_{x,S,E,t+1} = SP_{x,S,E} \qquad Default\ (x = 0\ to\ 85+)$$

$$P_{0,S,E,t+1} = NSP_{0,S,E,t+1} + SP_{x,S,E,t+1} \qquad Age\ 0$$

$$P_{x,S,E,t+1} = NSP_{x,S,E,t+1} + SSP_{x,S,E,t+1} \qquad (x = 1\ to\ 85+)$$

NSP is non-special population; B is births, ID is infant deaths; DNM is domestic net migration; INM is international net migration; SP is special population; P is total population.

The components of population change are computed by aggregating the non-special population, deaths, and domestic and international net migration over age groups along with the projection of births. Components of change are computed for males, females, and both sexes as follows:

$$D_{S,E} = \sum D_{x,S,E} + ID_{S,E} \qquad (x = 1\ to\ 85+)$$

$$DNM_{S,E} = \sum DNM_{x,S,E}$$

$$INM_{S,E} = \sum INM_{x,S,E}$$

$$NSPCH_{S,E} = \sum NSP_{x,S,E,t+1} - \sum NSP_{x,S,E,t}$$

$$NSPCH_{S,E} = B_{S,E} - D_{S,E} \pm DNM_{S,E} \pm DNM_{S,E}$$

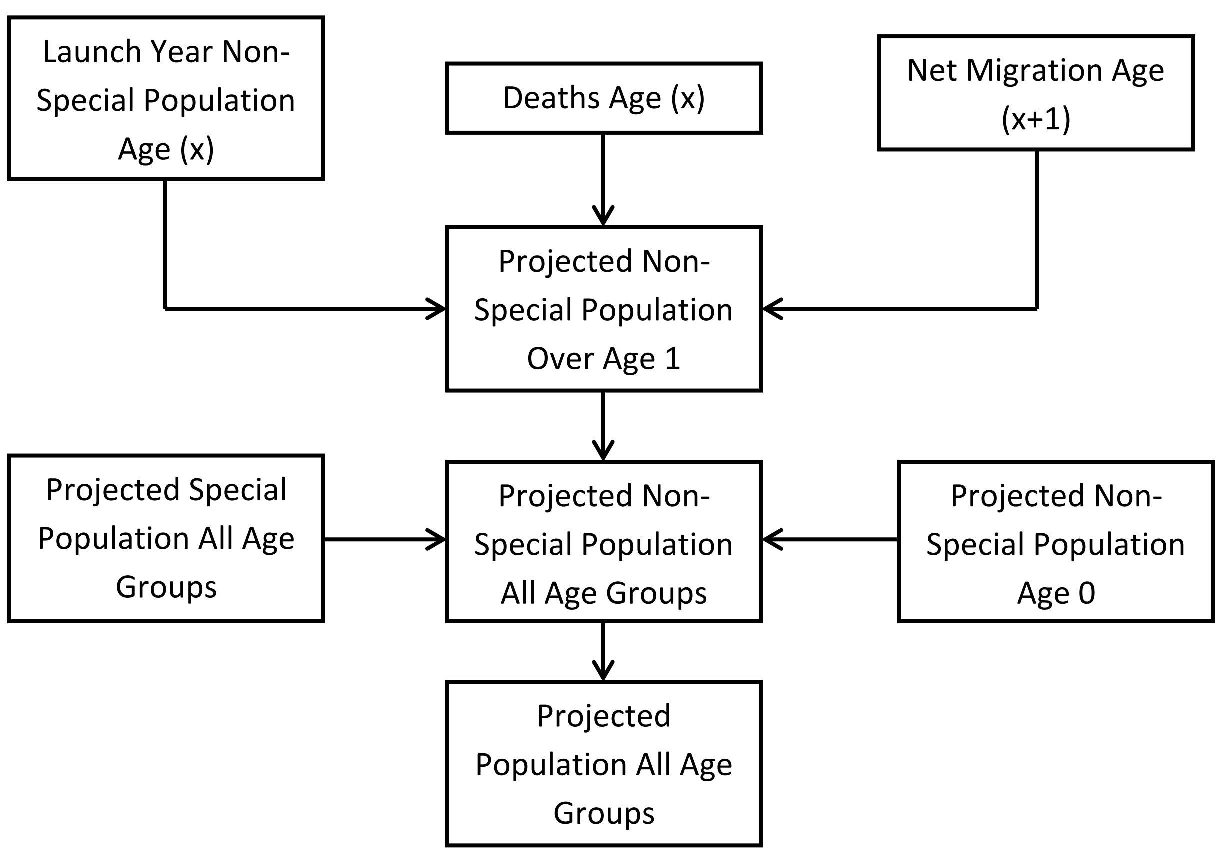
$$NSPCH_{E} = \sum_{S} B_{S,E} - D_{S,E} \pm \sum_{S} DNM_{x,S,E} \pm \sum_{S} DNM_{x,S,E}$$

$$where \sum_{S}. \; is\ the\ sum\ of\ Males\ and\ Females,$$

$NSPCH$ is change in the non-special population over the projection interval.

These equations reflect the uncontrolled projection information. If controls are used, the controlled values for births, deaths, infant deaths, domestic net migration, and foreign net migration would replace their counterparts in the above equations and are stored separately in the database.

Figure 5: Projected Projection Module



2.2.5 Aggregations

Higher-level totals are built using a bottom-up approach. Totals for both sexes are derived by adding males and females. Projections for all race/ethnic groups are the sum of the individual race/ethnic groups. The State projection is the sum of the counties. These aggregations are illustrated using total population by age (x) in the target year (P_{t+1}), but they work the same for any variable. Totals created by summing over age groups are not shown.

The computations for individual areas (*A*, for State or counties) are:

$$P_{x,E,A,t+1} = \sum_{S} P_{x,S,E,A,t+1}$$

$$P_{x,S,A,t+1} = \sum_{E} P_{x,S,E,A,t+1}$$

$$P_{x,A,t+1} = \sum_{S} P_{x,S,A,t+1}$$

These computations for the bottom-up State projections (P^S) are:

$$P^S_{x,S,E,t+1} = \sum_{A} P_{x,S,E,A,t+1}$$

$$P^S_{x,E,t+1} = \sum_{S} P^S_{x,S,E,t+1}$$

$$P^S_{x,S,t+1} = \sum_{E} P^S_{x,S,E,t+1}$$

$$P^S_{x,t+1} = \sum_{S} P^S_{x,S,t+1}$$

3. DATA INPUTS

Launch Year populations, both Total and Special (i.e. military personnel and dependents, prisoners, college students in dorms), form the starting point of a projection series, and various rates and proportions are used to compute the components of change. With the exception of the Launch Year total population, all of the data elements described below can be modified to reflect changing conditions during the projection series.

3.1 Population

The Launch Year total population is stratified by age, sex, and six race/ethnic groups. Each sex and race/ethnic group is arrayed into 85 single-year ages, from age 0 to age 84, and a final group ages 85 and over. The Special Population is stratified the same way as the Total Population. Special Populations complicate the projection process because their change is not determined by the same factors that affect fertility, mortality, and migration; consequently, they often follow trends that differ from the rest of the population and often have different demographic characteristics as well. These demographic differences can have a substantial impact on the projection of the components of change. Another characteristic of special populations is they often do not age in place as other population groups; therefore, their age structure may remain relatively stable over time.

3.2 Survival Rates

Survival rates are used to compute deaths and are derived from a life table. In a single-year model, the survival rate represents probability of a cohort surviving from one year to the next. Its complement is the probability of dying. One-year survival rates are needed for each age, sex, and race/ethnic group. An additional survival rate is required to compute infant deaths. This makes 87 survival rates for each sex and race/ethnic group in the model.

3.3 Fertility Rates

Age-Specific Birth Rates (ASBR) for individual ages from 15 to 44 are used to project births. A schedule of ASBRs is needed for each race/ethnic group. When there is no control, the birth rates are applied to the Launch Year female population adjusted for deaths and migration during the projection horizon. The proportions of births that are male (PBM) and female (PBF) are used to project male and female births respectively.

3.4 Migration Rates

Net migration rates for ages 1+, sex, and race/ethnic group are used to project domestic migration. Net migration rates use the local area (State or County) population as the population at risk in the denominator. These rates are based on the population at the beginning of the migration interval. The domestic migration projections are derived by applying these rates to the Launch Year population.

The projection of international net migration uses allocation factors by single year of age, which represents the share of the total net international migration; therefore, these factors sum to 1.0

over all ages. Separate factors for each sex and race/ethnic group are used in conjunction with sex and race/ethnic group-specific controls to project international migration.

3.5 Controls

Controls were employed for total net migration. Because age-specific net migration can be negative, zero, or positive, a two-factor controlling routine was used. For adjusting births or deaths, a single-factor routine was sufficient.

4. THE MODELING PROCESS

4.1 Launch Year Population

The launch year population is the point where the projections series begins. It typically is constructed from the most recent decennial census data (base population). Census 2020 was an atypical census, occurring in the midst of a worldwide pandemic which caused substantial operational changes and delays. In addition to challenges with the enumeration, the Census Bureau established a major change in its disclosure avoidance system to protect data confidentiality. This system, Differential Privacy, has contributed to delays in the release of Census 2020 data.

Census 2020 population at the detailed levels needed for the projections model will not be available until August 2023. Without this data source, OEO needed an alternative base population. Through a special data request, OEO obtained the Census Bureau's Vintage 2021 "Blended Base" population estimates for April 1, 2020. An infographic describing this dataset can be found at https://www.census.gov/library/visualizations/2022/comm/creating-the-vintage-2021-blended-base.html .

In simple terms, the blended base population uses Census 2020 populations at aggregate levels and distributes those using data that was produced before Census 2020 occurred. This was a necessary step to estimate post-Census populations because detailed Census 2020 data was not available even to internal Census Bureau divisions.

The blended base population for Arizona and its counties was obtained at the sex and single year of age level in both 11 race groups and 31 race groups. These data were aggregated to match the format of the projections model and adjusted with a simple factor to match OEO's official postcensal population estimates for July 1, 2021. Several drafts were produced using this dataset as the base population where anomalies emerged in certain single ages, particularly in Maricopa County.

To remedy this, OEO researched building a different base population using the projected 2020 population from OEO's 2018 projections series in conjunction with the Census Bureau's blended base. The projected 2020 population had the benefit of smoother transitions between single ages and allowed certain non-special populations to correctly age in place. A new base population for the model was constructed by controlling the projected 2020 population to blended base totals by county, race, sex, and voting age, thereby taking advantage of a robust age distribution and Census 2020 enumerated data[6]. A simple factor adjustment was applied to bring the population in line with OEO's official postcensal population estimates for July 1, 2021.

4.2 Fertility Rates

Fertility rates by mother's age group, race/ethnicity and county of residence were computed for the year centered on the 2020 Census Date (October 1, 2019 - September 30, 2020). Rates for the Balance of State (i.e. all counties excluding Maricopa, Pinal and Pima), were also computed.

The rates excluded the <15 and 45+ age groups. However, the births by females aged <15 were included in the 15-19 age group while the births by females aged 45+ were included in the 40-44 age group, with no adjustments made to the base populations of the rates.

Total Fertility Rates (TFRs) by county and race/ethnicity also were computed. Comparisons were made between counties and state, race/ethnic groups, and the total population.

Birth rates often are unreliable because of small cell sizes and also because of a mismatch of race/ethnic grouping between population and birth data; population data did not have an equivalent "Other" race category to what was collected on birth records. Therefore, a substitution scheme was developed and applied accordingly.

Age-specific birth rates (ASBRs) were computed for White Non-Hispanics (White NH) and Other Non-Hispanics (Other NH) combined. These rates were substituted for the White NH group and the Other NH group. No substitutions were made for the remaining race groups.

The ASBRs were applied to the female non-special population as of July 1, 2021, to estimate births in Fiscal Year 2022. Using the actual births and estimated births, an adjustment factor was computed and applied to the rates to obtain the new ASBRs. This adjustment is a calibration process that ensures the fertility rates closely produce the number of births that occur in reality. Calibration was performed for each county and Arizona[7].

4.2.1 Projection of Fertility Rates

U.S. historical fertility trends were stable with a slight upward trajectory for approximately 20 years starting in 1990. However, TFRs dropped substantially in 2007 and have not recovered. This was first caused by a decrease in fertility of Hispanic women. Low TFRs have persisted partly as a consequence of another trend where ASBRs for younger women, especially teens and those in their early 20s, have fallen while those of women older than 30 have increased.

---

[6] This hybrid construction of base population was not used in Coconino and Pima counties because the non-special populations did not correctly age in place. The blended base version was used.

[7] One set of state rates were computed. These state rates were used for all counties but were calibrated to actual county births for FY2021 for a set of base rates. The base rates were then calibrated to FY2022.

These trends are also observed in Arizona. In 2020, the TFR for all women was 1.66, with the value for White NH and Other NH women being even lower at 1.53. Hispanic women, who tend to have greater fertility than non-Hispanic women, had a TFR of only 1.85. Although these values are below the replacement rate, it is not expected that they will rebound. The projections model holds the TFR (and ASBRs) for all race groups constant for all years in the horizon.

Not all counties exhibit the same patterns in fertility as the state. Some have always, and will continue to, experience either higher or lower fertility than the state as a whole. To preserve these relationships, the model holds each county's specific rates and TFRs constant for all years.

4.3 Mortality Rates

Life tables were constructed by race, ethnicity, and sex using vintage 2021 blended base population estimates and deaths between October 1, 2019, and September 30, 2020 for Arizona, each county, and smaller counties combined (other than Maricopa, Pima, and Pinal). Life expectancies at birth from these life tables were also computed. Deaths that were caused by COVID-19 were excluded from these calculations because an increased number of deaths due to a temporary pandemic is not expected to continue into the future.

Because of the scarcity of data in some cells, not all life tables could be computed, and some were unreliable. Therefore, we made substitutions for race/ethnic groups where the number of known deaths was inadequate and also for those where there were mismatches in race/ethnic grouping between population data and Arizona vital statistics data. The substitutions are as follows:

- US 2019 Black for Non-Hispanic Black
- US 2019 Native American for Non-Hispanic Native American
- US 2019 Asian for Non-Hispanic Asian
- AZ Non-Hispanic for White Non-Hispanic & Other Non-Hispanic

For Arizona data, the 5-year death rates were transformed into 1-year death rates using cubic spline interpolation. A linear adjustment was made where the splines produced negative rates and where the increase/decrease between two 5-year age groups was not monotonic. One-year survival rates were computed from the adjusted 1-year death rates using basic life table functions. Survival rates from the US 2019 data were created from life tables published by the National Center for Health Statistics (NCHS). Although data were available for 2020, they were highly skewed by the increased number of deaths caused by COVID-19. Using the most recent death data unaffected by the pandemic (2019) provided a better gauge for future patterns in mortality.

A calibration process similar to that used for fertility rates was then performed. The survival rates were subtracted from one to create a new "death rate". The single-year "death rates" were applied to the July 1, 2020 non-special population to estimate the total number of deaths by sex and race/ethnic group for FY2021. An adjustment factor equal to total model deaths/total estimated deaths was calculated, applied to the "death rates," and subtracted from one to yield new survival rates. The calibration was performed for each county and Arizona[8]. Deaths attributed to COVID-19 were removed from the calibration process to keep from overestimating mortality in future years. When the calibration for the launch year was performed, this same logic was used, and only non-COVID deaths in FY2022 were included.

4.3.1 Projection of Life Expectancies

The Social Security Administration's (SSA) Office of the Chief Actuary published projected life tables for the total population for every 10 years up to 2100. The difference between SSA's life expectancy in 2019 and 2060 was estimated to be four years for males and 3.4 years for females.

A problem arises given that SSA publishes life tables for the total population while Arizona's projections are based on six race/ethnic groups. Not every group will experience the same improvement in life expectancy as reported by SSA.

According to the Census Bureau's 2017 population projection, different race/ethnicity and sex groups are projected to have the following number of years in life expectancy improvement between 2017[9] and 2060:

- Hispanic (of any race): Male 6.4, Female 4.8
- Non-Hispanic Black & American Indian: Male 11.0, Female 8.8
- Non-Hispanic White & Asian: Male 6.8, Female 5.5

These race/ethnicity groupings also do not correspond directly to those used in Arizona's model, and it is difficult to reconcile these changes with those produced by the SSA. However, a useful pattern was observed; improvement in projected life expectancy is inversely related to the current life expectancy. If a group currently has a relatively low life expectancy, its projected improvement is relatively high. To respect this pattern, the projected improvement in life expectancy was adjusted (either upward or downward) based on current life expectancies of each race/ethnicity/sex group. That is, if a race/ethnicity/sex group's 2020 life expectancy were lower than the SSA's 2019 total population life expectancy, an upward adjustment was made; if the race/ethnicity/sex group's 2020 life expectancy were higher than the SSA's 2019 total population life expectancy, a downward adjustment was made. After some experimentation, the group adjustment was set to equal 1/4 of the difference between SSA's 2019 total population life expectancy and the Arizona race/ethnicity/sex group's 2020 life expectancy. To obtain the 2060 life expectancy for a particular group, the formula below was used.

[8] The substitutions were used to create a set of state rates. These state rates were used for all counties but were calibrated to specific county deaths in FY2021.

[9] Data for 2019 was unavailable.

$$2060\ Life\ Expectancy\ (LE) = AZ\ 2020\ LE\ with\ Substitution\ \&\ Calibration\ + SSA\ Change\ +\ Group\ Adjustment$$

The calibrated base survival rates were adjusted to match the 2060 life expectancies and were used as the final survival rates for 2060 at the state level. At the county level, the 2060 life expectancies were obtained by adjusting the state value according to the calibration factor of the base rates in 2020. After calibrated rates for FY2022 were produced, linear interpolation was used to create the survival rates for 2023-2059.

The base survival rates used to project the long-term survival rates were different than those described earlier in section 4.3. Even after excluding deaths from COVID-19, the number of deaths in FY2020-FY2022 were much higher than what would be expected based on historical trends. Because of this, deaths in these three years were removed as outliers, and a linear model was fit to data from 1990-2019 for all geographies. New total deaths were estimated using the trendlines and served as the calibration controls for the base rates in FY2021 and launch year rates in FY2022.

Current data suggest that deaths are likely to be above normal in the short term as we transition away from the pandemic. Because of this, the original calibrated rates for FY2022 are used for an additional year and then linearly move towards the long-term trend in 2024.

4.4 Special Populations

For the projections model, military persons in Group Quarters (GQ), military persons in households and their dependents, people in adult correctional facilities, and college students in dormitories were treated as Special Population.

The number of military persons and their dependents not in GQ were estimated using the ACS 2015-2019 5-year data. The total population in households where at least one person was active in the military was calculated. Because of the variability of survey results in small areas, the state distribution of military and dependents by race, sex, and age was used in all counties and controlled to each PUMA's total military and dependents population[10]. The resulting number was added to the base 2020 special population and held constant for the projection horizon.

Active military and their dependents are the only part of special population that is from households. The three other groups in special population reside in group quarters and their base population values are controlled to match the Census 2020 enumerated counts. These groups are persons in college dormitories, adult correctional facilities, and military barracks.

Some counties have a large number of college students who do not live in dorms and thus, were not accounted for in the Census 2020 special population. Also, the census special population was

[10] ACS data is categorized by Public Use Microdata Areas (PUMA) which sometimes consist of more than one county. Military and their dependents in the Gila+Pinal PUMA were assigned to Pinal. Military and their dependents from the Cochise, Graham, Greenlee, and Santa Cruz PUMA were assigned to Cochise.

only available in 5-year age groups, causing more people to be distributed to ages 15-17 than there should be given that most college students are 18 and older. To more accurately reflect the student population, adjustments were made to the special population ages 18-30 for Coconino and Pima.

In these two counties, a method was devised to compare the census total population of ages 18-30 to the natural cohort of those aged 17. The difference between the population aged 17 and the population of each single year of age is assumed to be special population. This provided a better estimate of the size of special population in these age groups, ensuring that an appropriate proportion of persons aged in place.

Special population in dorms accounted for growth in the college population, as evidenced by enrollment figures from five major universities in Arizona[11]. The annualized growth rate from 2010-2021[12] for each university was applied to the Census dormitory population of the counties where each institution was located[13]. Arizona experienced its peak number of births in 2007, and these children would reach college age (18 years) in 2025. To account for this, the dormitory population was allowed to grow until 2025. It is then held constant from 2025-2060.

Using data reported from the regional Councils of Government (COGs), the average population of correctional facilities and military barracks between 2010 and 2019 was calculated. This value was assigned to the GQ population in adult correctional facilities and military barracks in 2025, with the level of both groups held constant from 2025 onward. Values for 2021-2024 were interpolated.

The growth in special population is in addition to net migration controls discussed in the following sections; it is exogenous to the model and not part of the migration calculated from year to year.

4.5 Migration

Development of net migration data inputs involved several steps. Controls for total net migration were developed using historical data, which were then subdivided into net foreign migration, net domestic migration, and race/ethnicity/sex/age groups. This process is depicted in Figure 6 and described in the following sections.

4.5.1 Projected Migration Controls

The trend in total net migration for each decade from 1950 onward was analyzed in conjunction with economic history to project decadal migration controls for the 2010s, 2020s, 2030s, and 2040s. Controls were produced using two methods. In the first method, decadal migration

---

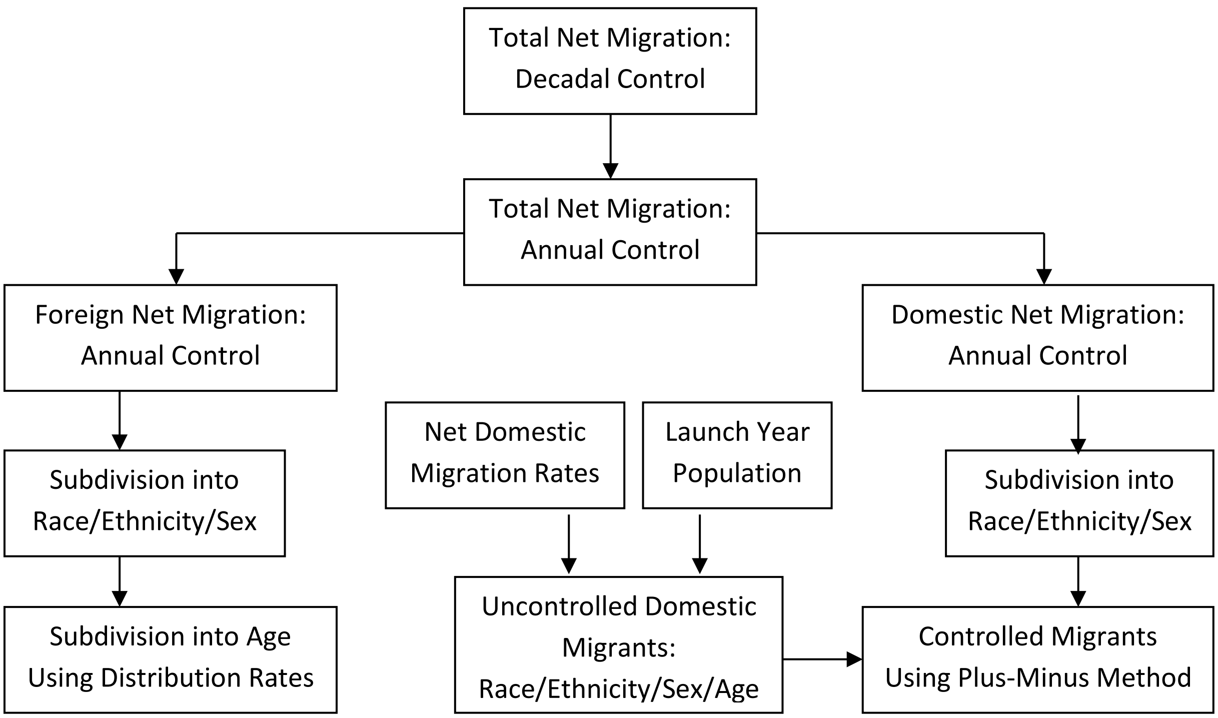
11 University of Arizona, Arizona State University, Northern Arizona University, Embry-Riddle Aeronautical University, and Grand Canyon University

12 Data from the Arizona Board of Regents Twenty-First Day Headcount

13 Pima, Maricopa, Coconino, and Yavapai

controls for the next four decades were set to the average decadal net migration implied between 1990 and 2020 in each county. The decadal controls were subdivided into annual net migration using osculatory interpolation. The second method applies the average county share over 20 years (2001-2020) of state net migration to the state total from the first method.

Figure 6: Development of Migration Controls



The average of both methods was taken and used as the starting point for discussion among members of the Council for Technical Solutions (CTS) and other stakeholders. Feedback was received on a case-by-case basis and adjustments were made in accordance with local knowledge of economic development and city planning. The long-term trend in Metro Phoenix assumed that Pinal County will approach 33 percent of the share of net migration in 2060[14]. Net migration in some smaller counties was interpolated through the current decade to meet their long-term trends. Apache, Greenlee, La Paz, Navajo, and Santa Cruz counties will reach zero net domestic migration in 2040, which will stay constant to 2060.

[14] In FY2022, Pinal County's share of net migration in Metro Phoenix was 17.2%. Migration controls for Maricopa and Pinal counties were adopted from analyses done by Maricopa Association of Governments. In this series, Pinal's share of net migration in Metro Phoenix reached 32.7% in 2060.

Annual controls for each county and the state were separated into annual domestic controls and annual foreign controls. Foreign controls were calculated first and then subtracted from the total net migration controls to produce the domestic controls.

The Census Bureau's vintage 2017 projection of net foreign migration was used along with Arizona and U.S. data on legal permanent residents to project foreign migration controls. The results from two methods were taken and used as the state foreign migration controls. In the first method, Arizona's share of U.S. legal permanent residents from 1995-2020, compiled by the Department of Homeland Security, was fit to a logarithmic model. The share projected by this model was applied to the applied to the Census projections to obtain annual foreign migration controls[15]. In the second method, Arizona's average share of U.S. migration[16] over 10 years (2011-2021) was applied to the Census projections.

The state control is distributed to each county based on the 10-year annual average proportion of state foreign migration with a few adjustments. Mohave's and Yavapai's proportions were set to zero due to their small negative values. The proportions for all other counties were adjusted so that the total distribution summed to 100%.

4.5.2 Subdivision of County Migration Controls

County migration controls were split into 12 race/ethnicity/sex groups. The procedure followed for domestic migration and foreign migration were the same; the county total was proportionately split into 12 groups based on an assumed distribution. However, the distribution used for each type of migration differed, as described in Sections 4.5.3 and 4.5.4.

Survey estimates from the 2015-2019 ACS 5-year PUMS were the only source for finding the proportion of domestic migrants by race and sex. The estimates were unreliable and problematic when used to distribute the controls. This was also the case in the 2018 projections series. To resolve this issue, OEO adopted the same logic as in 2018; the Census total population data provided a much more feasible scenario and was used to distribute the annual domestic migration control for each county into smaller groups. Census 2020 data was used, with some adjustments to more accurately capture the movement of Native Americans within the state.

These adjustments were retained from the 2018 model. Because many counties had very little movement of Native Americans, the proportion of domestic migrants assigned to the Native American race groups was set to zero[17]. The proportions of the remaining 10 race/ethnicity/sex groups were then adjusted to sum to 100 percent.

---

15 The Census Bureau's Vintage 2017 projections were first adjusted using the difference between the projected value in 2017 and the official estimated value in 2017 published by the Census Bureau. This brought the level of foreign migration down for the entire series to reflect lower than expected foreign migration.

16 Data were calculated from the Census Bureau's population estimates from vintages 2019 and 2021.

17 Counties with nonzero domestic migration of Native Americans were Apache, Coconino, Maricopa, and Navajo.

ACS 5-year PUMS data for 2015-2019 were used to model the distribution of foreign migrants across the 12 groups. To better reflect recent migration patterns, data analysis included only foreign-born persons who entered Arizona in 2015 or later.

4.5.3 Net Foreign Migration Distribution Factors

The same ACS 2015-2019 PUMS data were also used to model the distribution of foreign migrants across all ages. Sparse data on migrants by race and sex produced unreliable distributions in several PUMAs. In these cases, the state distribution was adopted and controlled to the estimated total for the PUMA. If the size of a race group was 500 or greater, the PUMA specific distribution factors were used.

4.5.4 Net Domestic Migration Rates

Domestic migration from the ACS 2015-2019 PUMS were analyzed but produced unusable distributions. Expressed as rates, too many anomalies by race, sex, and age were present to be used in the model. With no other viable data sources available (because of the delay in Census 2020 data releases), OEO decided to use the domestic migration rates from the 2018 projections series. These rates were much more robust because they were developed using data from Census 2000, Census 2010, and vital statistics from ADHS.

Creating these rates was an incredibly detailed process. Two distinct sets of rates were produced using two distinct methods. The average rates of both sets were used in the model. The first method required the estimation of annualized implied total net migration, annualized foreign net migration, and annualized net domestic migration. Simply described, the net domestic migration rate is produced using the formula

$$Domestic\ Migration\ Rate = \frac{(Annualized\ Total\ Net\ Migrants - Annualized\ Foreign\ Migrants)}{Population\ Base}$$

The second method made use of the migration information from Census 2000 for the period 1995-2000. Details of each method are in the following sections.

<u>4.5.4.1 Implied Migration Method</u>

Due to the lack of accurate direct migration data, implied migration for the decade was computed using both 2010 and 2000 Census Populations, Births, Deaths, and the formula

$$Implied\ Migration = \Delta Population\ -\ Births\ +\ Deaths$$

where Δ is the change between 2000 Census and 2010 Census dates.

The steps below delineate how implied migration was calculated:

1. For each race/ethnic/sex group, implied net migration was calculated for the state, Maricopa, Pima, Pinal, and Balance of State.
2. The ethnicity of intercensal births was adjusted to compensate for suspected differential classification of Hispanic births. The adjustment was guided by historical data from single year ACS PUMS from 2000-2010.
3. A ratio adjustment was performed on the Census 2010 population under 10 years of age to reflect undercounting of young children. The ratio compares the population of children under 10 from the Census Bureau's Demographic Analysis to the Census 2010 population.
4. The implied net migration over 10 years was then annualized.
5. The estimated net foreign migrants are subtracted from the annualized net migration to obtain the annualized domestic net migration.
6. A population denominator is calculated and used to produce a net domestic migration rate.

4.5.4.2 Census Migrants 1995-2000 Method

Using the tabulation of migrants from Census 2000, state-level factors were used to create 5-year migration rates for Hispanics and Non-Hispanics and males and females within these race/ethnic categories. These state factors were used for Maricopa, Pima, Pinal, and the Balance of the State because of data limitations and to provide more consistency in the age-specific rates across geographic areas. The detailed steps are as follows:

1. Estimated age-specific rates for Hispanics and non-Hispanics using the age-specific ratios of the Hispanic to total rate and non-Hispanic to total rate. The same ratios were used for each race group given data limitations, and they adequately captured the tendency for Hispanic domestic migrants to be younger do not have increased rates in the retirement ages compared to non-Hispanic domestic migrants.
2. The age-specific rate for Hispanics was split into male and female rates using age-specific ratios of the Male Hispanic to Total Hispanic rate and female Hispanic to Total Hispanic rate. The same ratios were used for Hispanics for each race group given data limitations, and they adequately captured the variation by sex in the Hispanic population.
3. The age-specific rates for non-Hispanics were split into initial male and female rates using age-specific ratios of the Male non-Hispanic to Total non-Hispanic rate and female non-Hispanic to Total non-Hispanic rate. The same ratios were used for non-Hispanics for each race group given data limitations, and they adequately captured the variation by sex in the non-Hispanic population.
4. Using cubic spline interpolation, the 5-year age group rates were interpolated into single year ages.

5. An initial estimate of the net migration by age by each sex and race/ethnic group was generated using the Census 2010 data, and the Plus-Minus method was used to adjust the rates, so they match the calibration control. The same adjustment factor(s) was applied to each sex and race/ethnic group.

4.5.4.3 Adjustments to Net Domestic Migration Rates

The assumption was that averaging the two sets of rates above would produce more reasonable migration rates than using any one method alone. However, even after averaging, tests revealed that further adjustments were needed.

Three types of adjustments were made. The first, and the simplest, was applied to ages 70 and older in all counties and the state. All positive migration rates for ages 70 and older were replaced with zero. This adjustment was needed because implied migration showed net outmigration occurring in older ages, which was not reflected by the averaged rates. Net outmigration among the very old is widely recognized among the senior living community. The hypothesis is that as out-of-state retirees become older and frailer, some of them move out of state to be closer to their children. This is solidly supported by migration figures implied by decennial census data and death statistics.

The second adjustment was applied to five counties[18] for the college-age population and was needed because the "balance of state" migration rates did not adequately capture the out-migration in this age band. The adjustment was based on survival analysis. The population for 2009 was estimated using the 10-year migration rate between age seven and seventeen[19]. An annual migration rate was then calculated using the Census 2010 population and estimated 2009 population and replaced the averaged rates.

The last adjustment was made to ages 37+ in Gila and 45+ in La Paz. Rates were recalculated by annualizing the implied net migration between Census 2000 and Census 2010 and dividing by a population base. This was needed because the age patterns of migration in these counties were demonstrably different than those of the "balance of sate." The adjustment was also based on survival analysis.

[18] Apache, Gila, La Paz, Navajo, and Santa Cruz

[19] The method assumes that all changes in the population are due to migration. Deaths are not considered as they are a very small part of the population change for 7-17 year-olds.

5. ACKNOWLEDGMENTS

The OEO staff would like to thank the following people for their contributions to the projections process. They devoted many hours of discussion and data analysis and provided guidance for model design that resulted in high quality data inputs and results. The affiliation listed for each person is as of the time of their contribution and may or may not be current.

Current and former members of the Council for Technical Solutions:

Travis Ashbaugh, Central Arizona Association of Governments
Jesse Ayers, Maricopa Association of Governments
Scott Bridwell, Maricopa Association of Governments
Tracy Clark, Arizona Department of Transportation
Chris Fetzer, Northern Arizona Council of Governments
George Hammond, University of Arizona
Michael Huff, Northern Arizona Council of Governments
Roland Hulse, Western Arizona Council of Governments
John Merideth, Southeastern Arizona Governments Organization
Caroline Leung, Pima Association of Governments
Hyunsoo Noh, Pima Association of Governments
Matt Roach, Arizona Department of Health Services

Many other individuals also reviewed draft projections or attended many of our presentations on the progress of the projections. Their comments, criticisms, and encouragements are greatly appreciated.