

ARIZONA DEPARTMENT OF ADMINISTRATION

Office of Employment and Population Statistics

Arizona State and County Population Projections, 2012-2050: Methodology Report

December 7, 2012

TABLE OF CONTENTS

1. BACKGROUND	1
2. METHODOLOGY.....	1
2.1 The Cohort-Component Model.....	2
2.2 Component Modules of the Projections Model	3
2.2.1 Mortality Module.....	4
2.2.2 Net Migration Module	5
2.2.3 Fertility Module.....	7
2.2.4 Projected Population Module	10
2.2.5 Aggregations	11
3. DATA INPUTS.....	12
3.1 Population	13
3.2 Survival Rates.....	13
3.3 Fertility Rates	13
3.4 Migration Rates.....	13
3.4 Controls.....	13
4. THE MODELING PROCESS.....	14
4.1 Fertility Rates	14
4.1.1 Projection of Fertility Rates	14
4.2 Mortality Rates.....	15
4.2.1 Projection of Life Expectancies	16
4.3 Special Populations	18
4.4 Migration.....	18
4.4.1 Projected Migration Controls.....	19
4.4.2 Subdivision of County Migration Controls	21
4.4.3 Net Foreign Migration Distribution Rates.....	21
4.4.4 Net Domestic Migration Rates.....	23
4.4.4.1 Implied Migration Method.....	23
4.4.4.2 Census Migrants 1995-2000 Method	24
4.4.4.3 Adjustments to Net Domestic Migration Rates	24
5. ACKNOWLEDGMENTS.....	25

1. BACKGROUND

Arizona State and County Population Projections (2012 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.

2. METHODOLOGY

The Arizona Population Projections Model is a Cohort-Component model. This model is used to project population for 10 race/ethnic groups in 16 geographical areas (Arizona state and its 15 counties) over a projection period of 40 years. The 10 race/ethnic groups are the combinations of five race groups by Hispanic origin. The five race groups are: White, Black, Native American, Asian (including Native Hawaiian and Pacific Islander), and Other (including two or more races). Each of these races is delineated into Hispanic origin and non-Hispanic origin to make a total of 10 race/ethnic groups.

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.

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 10 ethnic groups categorized by race and Hispanic origin. Movement of population from one time period to the next is accomplished by adding births and net immigration and subtracting deaths and net out-migration to each cohort (see Figure 1). The basic projection equations are:

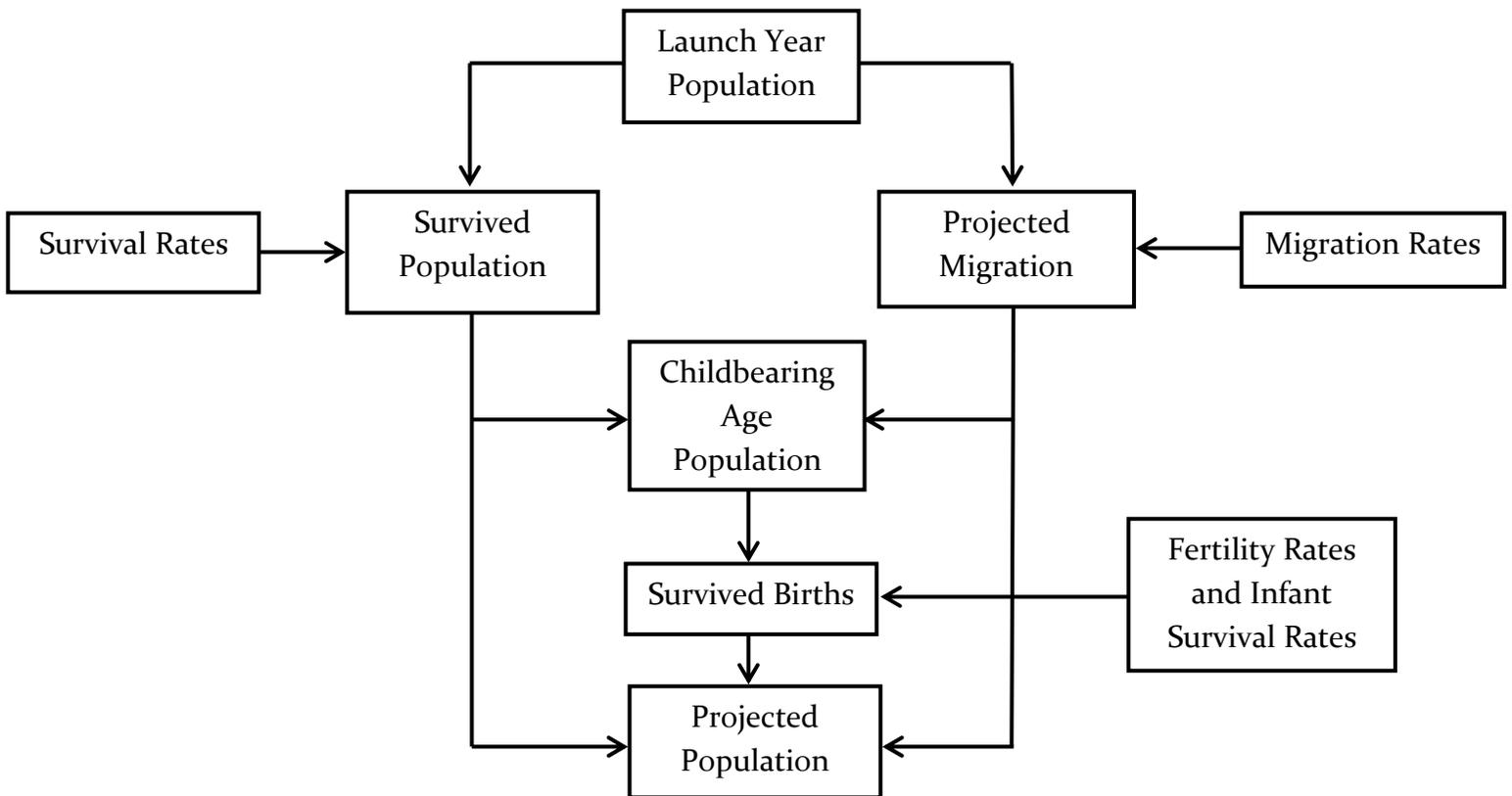
$$P_{0,S,E,t+1} = B_{S,E} - ID_{S,E} \quad \text{Age 0}$$

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

$$P_{85+,S,E,t+1} = P_{84,S,E,t} + P_{85+,S,E,t} - D_{x,S,E} - D_{84,S,E} - D_{85+,S,E} \pm DNM_{84,S,E} \pm INM_{84,S,E} \pm DNM_{85+,S,E} \pm INM_{85+,S,E} \quad \text{Age 85 +}$$

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 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 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 or subtracting 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 ethnic group. Totals for both sexes are derived by adding males and females. Projections for all ethnic groups are determined using a bottom-up approach by summing individual 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 ethnic-specific control totals.

Prior to entering Module 1, special populations are removed from the launch year population:

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

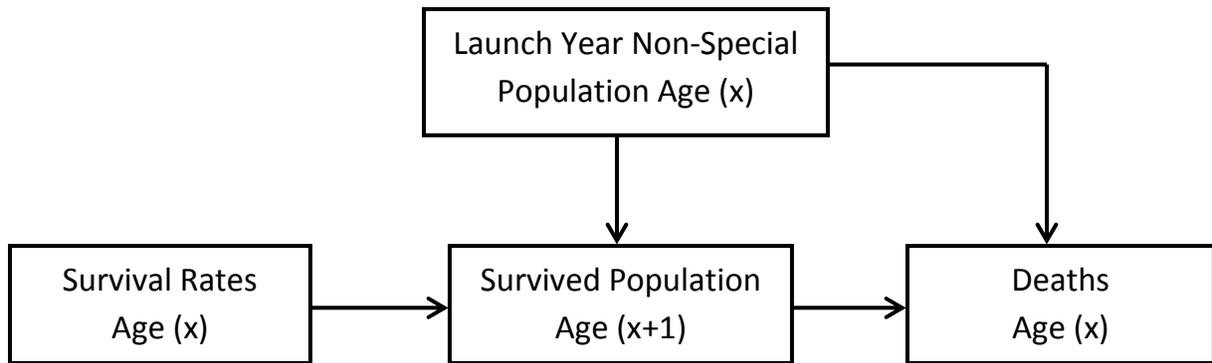
where NSP is non-special population; P is total population; and SP is special population.

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 ($SP_{x,S,E,t+1}$).

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:

$$SURVNSP_{x+1,S,E,t+1} = NSP_{x,S,E,t} * S_{x,S,E,t} \quad (x = 0 \text{ to } 83)$$

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

where $SURVNSP$ is survived non-special population; NSP is non-special population; and S 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.

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 group³:

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

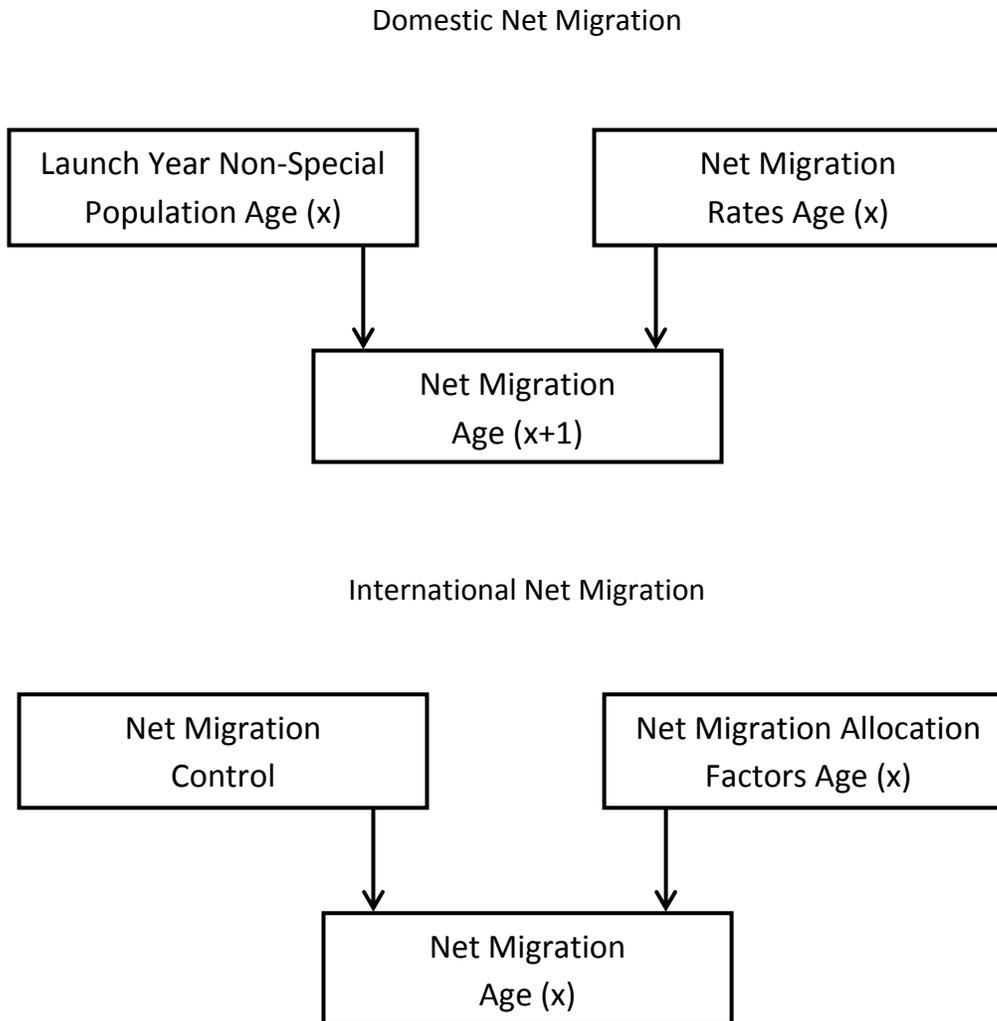
where *DNM* is projected domestic net migration of the non-special population and *DNMRATE* the launch year domestic net migration rate. For example, $DNM_{15,S,E}$ represents the population change due to domestic migration of the non-special population aged 15 in the launch year who are 16 in the target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

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 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.

Figure 3: Net Migration Module



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 for domestic net migration is:

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

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

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

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

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

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

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

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:

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

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

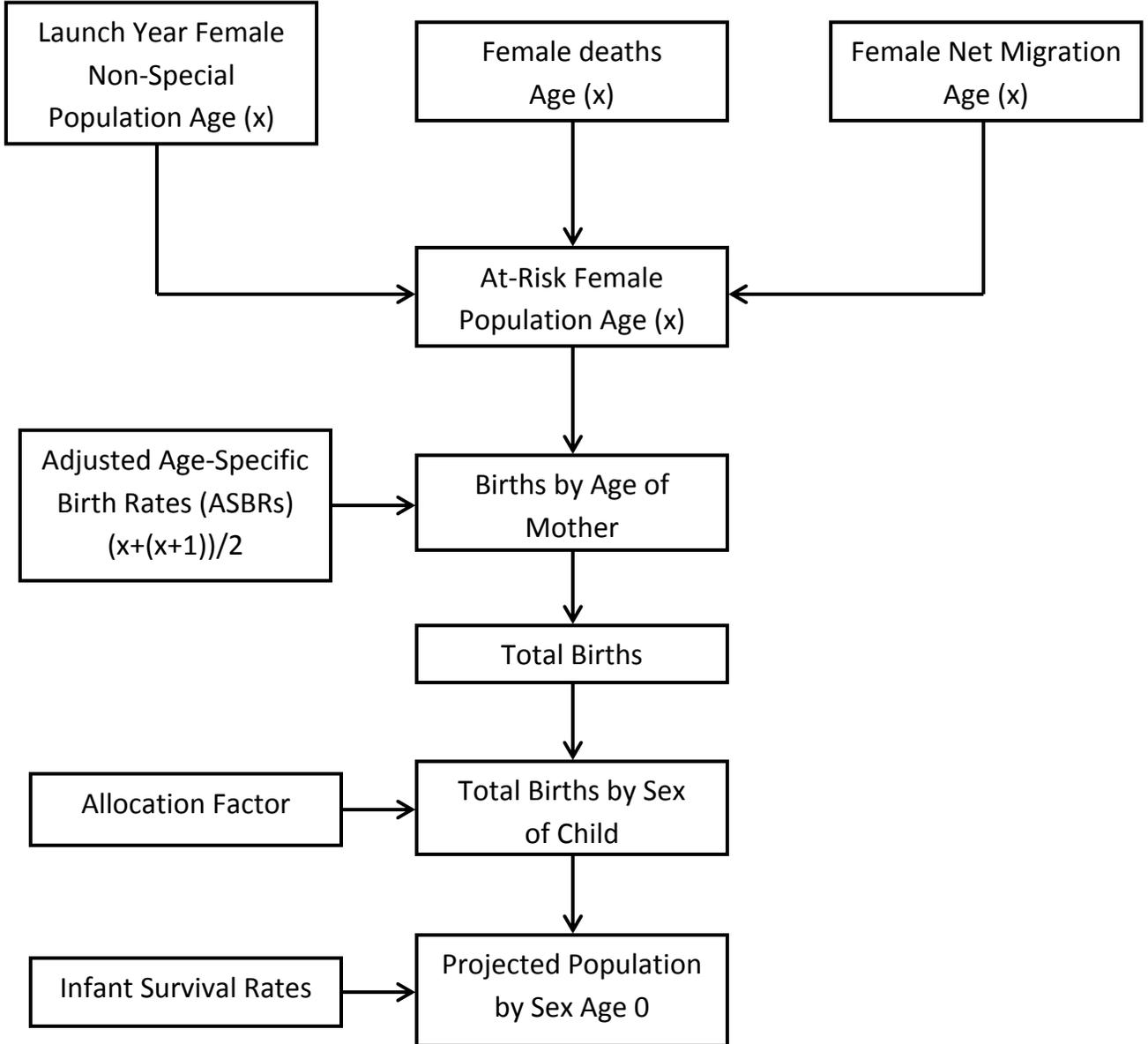
$$INETMIG_{S,E} \approx \sum INM_{x,S,E}$$

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:

$$ADJASBR_{x,E,t} = \frac{(ASBR_{x,E,t} + ASBR_{x+1,E,t})}{2} \quad (x = 14 \text{ to } 44)$$

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

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

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

$$MB_E = B_E * PBM_E$$

$$FB_E = B_E * PBF_E$$

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

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

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

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

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, since it is assumed woman live throughout half of the projection interval.⁴ The at-risk population is further adjusted by adding or subtracting the 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.

⁴ 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.

If 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:

$$\begin{aligned}
 NSP_{0,S,E,t+1} &= B_{S,E} - ID_{S,E} && \text{Age 0} \\
 NSP_{x,S,E,t+1} &= NSP_{x,S,E,t} \pm DNM_{x,S,E} \pm INM_{x,S,E} && (x = 1 \text{ to } 85+) \\
 SP_{x,S,E,t+1} &= SP_{x,S,E} && \text{Default } (x = 0 \text{ to } 85+) \\
 P_{0,S,E,t+1} &= NSP_{0,S,E,t+1} + SP_{x,S,E,t+1} && \text{Age 0} \\
 P_{x,S,E,t+1} &= NSP_{x,S,E,t+1} + SSP_{x,S,E,t+1} && (x = 1 \text{ to } 85+)
 \end{aligned}$$

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:

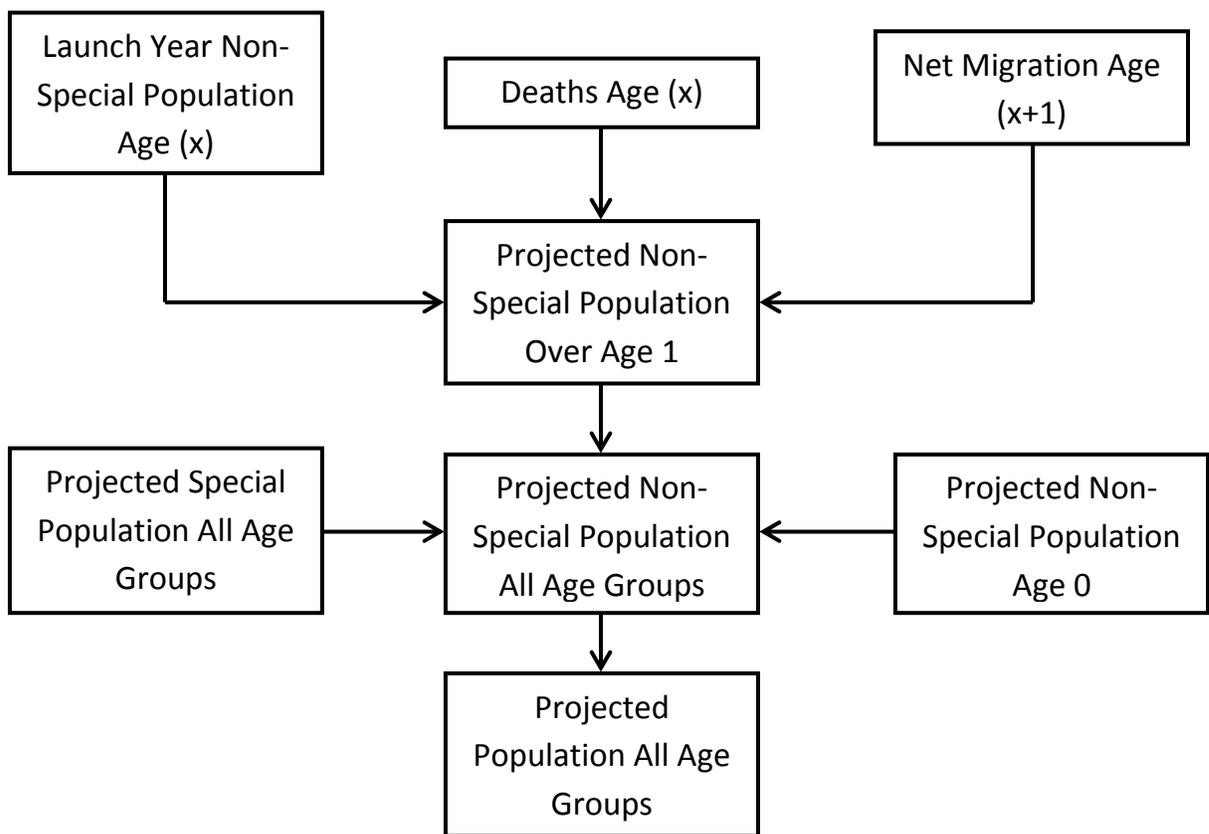
$$\begin{aligned}
 D_{S,E} &= \sum D_{x,S,E} + ID_{S,E} && (x = 1 \text{ to } 85+) \\
 DNM_{S,E} &= \sum DNM_{x,S,E} \\
 INM_{S,E} &= \sum INM_{x,S,E} \\
 NSPCH_{S,E} &= \sum NSP_{x,S,E,t+1} - \sum NSP_{x,S,E,t} \\
 NSPCH_{S,E} &= B_{S,E} - D_{S,E} \pm DNM_{S,E} \pm INM_{S,E} \\
 NSPCH_E &= \sum_S B_{S,E} - D_{S,E} \pm \sum_S DNM_{x,S,E} \pm \sum_S INM_{x,S,E}
 \end{aligned}$$

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 ethnic groups are the sum of the individual 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:

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

$$P_{x,E,t+1}^S = \sum_S P_{x,S,E,t+1}^S$$

$$P_{x,S,t+1}^S = \sum_E P_{x,S,E,t+1}^S$$

$$P_{x,t+1}^S = \sum_S P_{x,S,t+1}^S$$

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 10 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 ethnic group. An additional survival rate is required to compute infant deaths. This makes 87 survival rates for each sex and 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 ethnic group-specific controls to project international migration.

3.4 Controls

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

4. THE MODELING PROCESS

4.1 Fertility Rates

Fertility rates by mother's age group, race/ethnicity and county of residence were computed for the year centered on the 2010 Census Date (October 1, 2009 - September 30, 2010). 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 were also computed. Comparisons were made between counties and state, race/ethnic groups, and the total population.

Out of 150 combinations (i.e. 15 counties x 10 race/ethnic groups), nearly three-quarters (110) had less than 150 births. Birth rates calculated for these combinations were not reliable. Due to a mismatch of race/ethnic group between Census and birth data, the state-level rates were not applicable, either. Therefore, a substitution scheme was developed and applied accordingly.

Age-specific birth rates 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. Both the White Hispanic group and Other Hispanic group used the rates computed for Hispanics of any race. For the remaining 6 race/ethnic groups, the ratios of Non-Hispanic ASBR to Total ASBR, or the ratios of Hispanic ASBR to Total ASBR (depending on the Hispanic origin) were applied to the ASBRs of each race to obtain the new race/ethnic rates. The H and NH ratios were based on the total population (all races) ASBR.

The rates obtained from the substitution for all race/ethnic groups were applied to the female non-special population as of July 1, 2010 to estimate births in Fiscal Year 2011. Using the actual births and projected births, an adjustment factor was computed and applied to the rates from the substitution to obtain the new ASBRs and TFRs. 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 Arizona⁵.

4.1.1 Projection of Fertility Rates

To project the ASBRs over the projection horizon, historical US fertility trends were studied. US ASBRs and TFRs between 1990 and 2010 were collected for five race/ethnic groups (White NH, Black NH, American Indian or Alaska Native, Asian or Pacific Islander, and Hispanic (of any

⁵ One set of state rates were computed. These state rates were used for all counties but were calibrated to actual county births for FY2011.

race)). Studying the trends, it was found that White NH TFRs had been stable with a slight upward trend over the 21-year period. However, TFRs dropped substantially since 2007.

The US Census Bureau's 2008 projection made an assumption that by 2100, fertility levels of all race groups will converge to that of the White NH group. We made the same assumption. Based on historical births in Arizona and consensus among technical advisors, the TFR of all race/ethnic groups was set to converge to 1.9 in the year 2100, and corresponding 1-year ASBRs were created for input into the model.

Proportions were used to divide the 2100 White NH TFR into 5-year age group ASBRs. For each historical year, the proportion for each White NH age group was computed by dividing the White NH ASBR by the White NH TFR. After careful trend analysis of White NH ASBR proportions, the ASBRs proportion trend was extended to 2030 as some of the results obtained were not reasonable beyond 2030. By applying these proportions to the TFR of 1.9 in 2100, the ASBRs for 2100 were obtained at the 5-year age group level. The 5-year ASBRs were transformed into 1-year ASBRs using a cubic spline interpolation procedure.

Since an unusually deep decline in birth rates have been observed since 2007, it is reasoned that there might be a rebound in the next few years. Preliminary birth data for 2012 indicated that birth rates have stabilized, which further supports the conjecture of a rebound. Therefore, it is assumed that in the next five years, birth rates will rebound 50 percent of the way back to 2007 levels and will stay at those levels for the following five years until 2022. Using the 2022 ASBRs for each race/ethnic group and the 2100 White NH ASBRs, the ASBRs over the entire projection horizon were calculated by linear interpolation (stopping at 2050 since 2050 is the end of the projection period). These ASBRs were used in our Medium Series Projections.

For the High Series, we assumed that fertility rates will rebound 100 percent of the way to pre-recession high levels in the next five years and stay at these levels for five more years. In 2100, the TFR of all race groups will converge to 2.4.

For the Low Series, fertility rates will not rebound from today's levels. In 2100, the TFR of all race groups will converge to 1.4.

4.2 Mortality Rates

Life tables were constructed by race, ethnicity, and sex using Census 2010 population and deaths between 10/1/2009 and 9/30/2010 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.

Out of 374 possible life tables ((10 race/ethnicity groups + 1 total population) x 2 sexes x 17 places), 193 were computed at the 5-year level, 23 were computed at the 10-year level (because there was not enough data at the 5-year level), and the remaining 158 could not be computed at either level. Given the scarcity of data, other methods explored to account for the missing cases proved futile. 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 Census and Arizona vital statistics data. The substitutions are as follows:

- US 2007 Black for Non-Hispanic & Hispanic Black
- US 2007 Native American for Non-Hispanic & Hispanic Native American
- US 2007 Asian for Non-Hispanic & Hispanic Asian
- AZ Hispanic for White Hispanic & Other Hispanic
- AZ Non-Hispanic for Other Non-Hispanic

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.

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, 2010 non-special population to estimate the total number of deaths by sex and race/ethnic group for FY2011. Using actual deaths from 2003-2010, the number of deaths in FY2011 was predicted by a linear regression model. 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 Arizona⁶.

4.2.1 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.

A method was explored where SSA's projected rate of change in survival rates between 2010 and 2050 were applied to Arizona 2010 data. Since applying the rate of change to survival rates could result in values greater than one, data transformations were performed so that calculations could be carried out on death rates. The SSA's projected rate of change in death rates was applied to Arizona's 2010 death rates by age, sex, and race/ethnic group to obtain death rates for 2011-2050. These death rates were converted back to survival rates. Unfortunately, this process did not provide good control over the target year life expectancies.

⁶ 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 obtained by the regression model.

Therefore, the method used to project life expectancies relied on the numerical difference in SSA's projections instead of the rate of change. In principle, the difference between SSA's life expectancy in 2010 and 2050 was added to Arizona's 2010 life expectancy to obtain the Arizona life expectancy in 2050. A problem arises given that SSA publishes life tables for the total population while Arizona's projections are based on 10 race/ethnic groups, i.e. not every group will experience the same improvement in life expectancy as reported by SSA.

According to the Census Bureau's 2008 population projection, different race/ethnicity and sex groups are projected to have the following number of years in life expectancy improvement between 2010 and 2050:

- Hispanic (of any race): Male 3.5, Female 2.6
- Non-Hispanic Black: Male 8.9, Female 7.2
- Non-Hispanic (all other races): Male 4.6, Female 4.2

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 2010 life expectancy were lower than the SSA's 2010 total population life expectancy, an upward adjustment was made; if the race/ethnicity/sex group's 2010 life expectancy were higher than the SSA's 2010 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 2010 total population life expectancy and the Arizona race/ethnicity/sex group's 2010 life expectancy. To obtain the 2050 life expectancy for a particular group, the formula below was used.

$$2050 \text{ Life Expectancy (LE)} = \text{AZ 2010 LE with Substitution \& Calibration} + \\ \text{SSA Change} + \text{Group Adjustment}$$

The previously calibrated survival rates were adjusted to match the 2050 life expectancies and were used as the final survival rates for 2050. Linear interpolation was used to produce the survival rates for 2011-2049. These survival rates were used for the Medium Series projections.

To obtain survival rates and life expectancies for the Low Series projections, the 2030 SSA life table was substituted for the 2050 life table. For the High Series projection, the 2070 SSA life table was substituted for the 2050 life table. The logic is that in the next 40 years, instead of making improvements in survival rates and life expectancies that SSA projected for 2050, improvements for the Low Series would only reach the level that SSA projected for 2030; improvements for the High Series would reach the level that SSA projected for 2070. The group adjustment for the Low Series was 1/5 and 1/3 for the High Series.

4.3 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 2008-2010 3-year data. The proportion of total population that is active military and their dependents was calculated. These proportions were applied to the launch year population to estimate the number of military and their dependents⁷. The resulting number was added to the Census 2010 Special Population and held constant for the projection horizon.

Some counties have a large number of college students who do not live in dorms and thus, were not accounted for in the Census 2010 special population. Also, the census special population was 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 17-30 for Coconino, Pima, Maricopa, and Arizona.

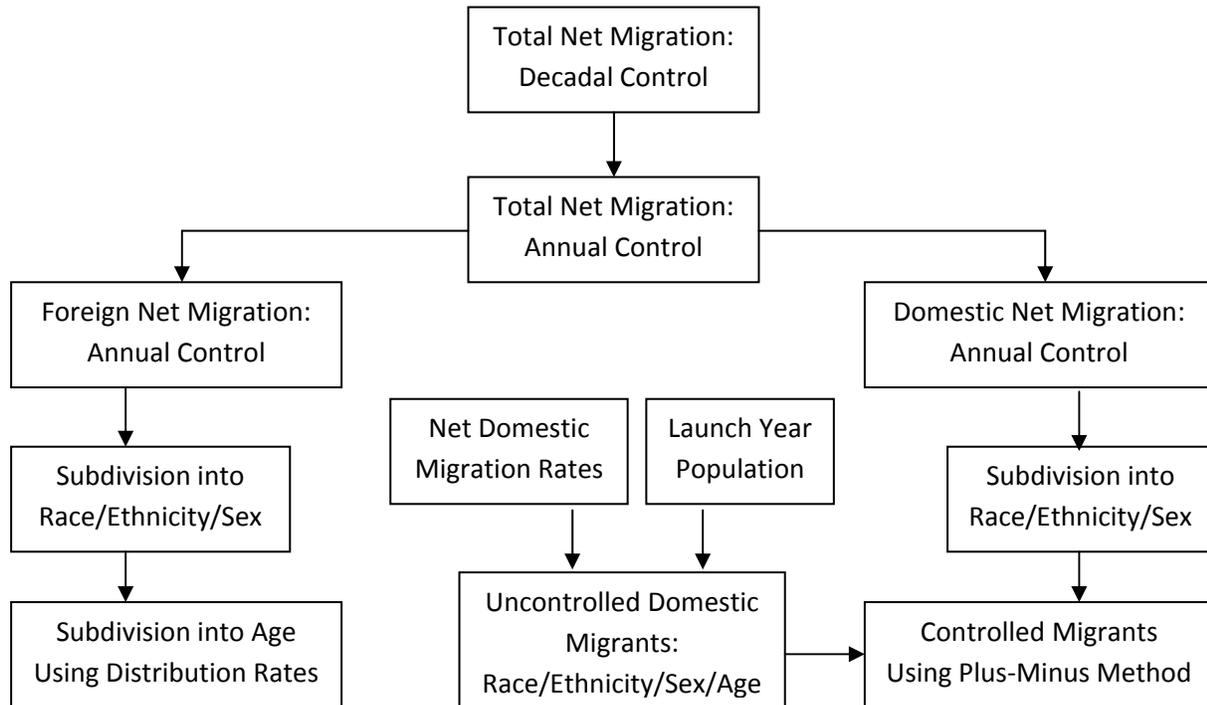
For Coconino and Pima, a method was devised to compare the census total population to the natural cohort of ages 17-30. The difference between the two is assumed to be special population. Specifically, a line was drawn from the census population of age 16 to age 31. This represented the natural cohort of residents in the county. For each year of age, the natural cohort was subtracted from the census population to obtain the estimated special population. All negative results were replaced with zero. The distribution of special population in Coconino and Pima combined was then used to redistribute the special population totals for ages 15-24 in Maricopa and Arizona.

4.4 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 the required race/ethnicity/sex/age groups. This process is depicted in Figure 6 and described in the following sections.

⁷ 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.

Figure 6: Development of Migration Controls



4.4.1 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. The contributions from Tom Rex of Arizona State University (ASU) and staff at Maricopa Association of Governments (MAG) deserve special credit. Based on his analysis of historical migration (including domestic migration, legal and undocumented immigration) and trend in retirement migration, Tom Rex proposed decadal migration controls for the next four decades and annual migration controls for the current decade for each county. He subsequently revised those controls based on his analysis of the current labor force imbalance. These controls were used as starting points for discussion among members of the Council for Technical Solutions (CTS) and other experts. MAG staff proposed alternative ways of analyzing the data.

Staff at the Office of Employment and Population Statistics (EPS) analyzed the proposed controls and feedback from various sources. For the current decade, it was decided that the average of Tom Rex’s two versions of controls would be used. For the 2020s and 2030s, a three-decade trend prediction, adjusted by the factor between actual and predicted 2000s’ values (rounded to the nearest 50,000), was used for the Phoenix Metro area. Other counties with net positive migration were adjusted proportionately from Tom Rex’s controls. Counties with negative net migration remained unchanged from his adjusted numbers. For the 2040s, net migration was held constant from the 2030s.

The two preceding paragraphs do not imply endorsement of the assumptions or projection outcome by either Tom Rex or MAG. EPS staff bear sole responsibility for the final decision on the assumptions and results.

In order to distribute the decadal controls into annual migration controls, the following assumptions were made for the current decade:

- Arizona is gaining population through net migration as of 2012.
- Net migration will increase at a moderate pace in the next two to three years.
- Net migration will be higher in the mid to late part of this decade than presently observed.
- There will be a couple of years where net migration is higher than the long-term trend.

For later decades, annual controls were produced from decadal controls using Karup-King coefficients for subdivision into tenths. Within the Phoenix Metro area, the relative share of migration between Maricopa County and Pinal County were adjusted. Pinal's share was assumed to change from the current 15 percent to about 33 percent in 2050⁸.

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 preliminary projection of net foreign migration was used along with AZ data on legal permanent residents to project foreign migration controls. Two methods were used by the Census Bureau to project US foreign migration in 2060. Linear interpolation between 2010 and 2060 was used to estimate annual foreign migration for both methods, and an average was taken of the interpolated results. Arizona's share of US legal permanent residents was then extrapolated from data compiled by the Department of Homeland Security and applied to the averaged Census projections to obtain annual foreign migration controls⁹. The extrapolation of Arizona's share ran to 2030 and then remained constant until 2050 in the Medium series.

For the Low series, the current decade's total net migration controls were set to 70 percent of the Medium series (for positive controls) or to 130 percent (for negative controls). Subsequent decades' controls were assumed to equal the 1980s level (rounded to the nearest 50,000). Arizona's share of US foreign migration was assumed to stay constant at the 2011 level for the entire projection horizon.

⁸ Historically, Pinal's share was tiny. It shot up during the 2000s, and reached a high of 80% in 2010 based on EPS estimates, or more than 300% based on IRS data. Many arguments suggested that those high shares are aberrations and cannot be sustained, and EPS agrees. There is evidence pointing to readjustment of the shares. While residential completions have increased in FY2012 in Maricopa County, they continue to decline in Pinal. Pinal's share of residential completions in the metro area in FY2012 is 13%, down from 28% in FY2006. In the long run, however, most people agree that Pinal still has potential. Thus, its share is allowed to go up over time.

⁹ Annual foreign migration controls were created at the state level. County shares, calculated from intercensal estimates by EPS and the Census Bureau, were used to subdivide the state's annual foreign migration controls.

The High series adjusts the Medium series controls by 125 percent (for positive controls) or 75 percent (for negative controls). Subsequent decades used a five-decade trend to predict the Phoenix Metro area (rounded to the nearest 50,000). Other counties with positive controls were adjusted proportionately. Negative controls remained unchanged for the 2020s and were set at 80 percent of the previous decade for the 2030s and 2040s. The trend of Arizona's share of US foreign migration was allowed to run to 2050.

4.4.2 Subdivision of County Migration Controls

After the annual migration controls were calculated, a procedure was established to subdivide those controls into the required 20 race/ethnicity/sex groups. The procedure followed for domestic migration and foreign migration were the same; the county total was proportionately split into 20 groups based on an assumed distribution. However, the distribution used for each type of migration differed.

ACS 5-year PUMS data for 2005-2010 were used to model the distribution of foreign migrants across the 20 groups. Originally, only data on ACS immigrants were used. However, these data produced distributions that did not coincide with other demographic information. Analysis was then expanded to include all foreign born persons who entered Arizona in 2005 or later. These distributions made much more sense and were used to split the annual foreign migration control into smaller levels.

Domestic migration from several ACS 3- and 5-year PUMS were analyzed but also produced unusable distributions. The Census 2010 total population data provided a much more believable scenario and was used to distribute the annual domestic migration control for each county into smaller groups.

Adjustments to the domestic distribution were needed to more accurately capture the movement of Native Americans within the state. For Coconino and Navajo counties, the projected net migration was positive. However, it is expected that a net loss of Native Americans should occur based on past trends. To ensure that a net loss was realized for these two counties, a constant value for net outmigration was used. Because all other counties, except Maricopa, had very little movement of Native Americans, the proportion of migrants assigned to the Native American race groups was set to zero. The proportions of the remaining 18 race/ethnicity/sex groups were then adjusted to sum to 100 percent.

4.4.3 Net Foreign Migration Distribution Rates

Using 1995-2000 Census and 2007-2009 ACS migration data, proportionate distributions by single year age groups for the 20 ethnic group/sex combinations were created. These distribution rates were used to partition the net foreign migration controls into age groups.

Greater reliance was placed on the 1995-2000 Census data because of the greater variability and unreasonable trends observed in the ACS data, especially for Blacks, Native Americans, and Asians, when stratified by age, sex, and Hispanic origin.

The 1995-2000 data did not contain the 1-4 age group. So the ACS 2007-2009 data were used to estimate the 1-4 age group for 1995-2000 using the ratio of the shares of 1-4 to 5-9 from the total net foreign migration. Where possible, the race-specific ratio of the shares was used to estimate the proportion in each race group. The resulting proportionate distributions were adjusted to sum to one.

State-level factors were used to create the distributions for Hispanics and Non-Hispanics, males and females within these ethnic categories, and to allocate 5-year age groups into single year of age. These state factors were used for Maricopa, Pima, Pinal, and the balance of the State because of data limitations and they provided more consistency in the distributions across geographic areas. The steps below detail how the foreign distribution rates were computed.

1. The proportionate distributions by 5-year age groups were averaged for each race from the Census and ACS data.
2. The age distribution for each race was split into initial Hispanic and non-Hispanic origin distributions using the age-specific ratios of the Hispanic to total share and non-Hispanic to total share. The same ratios were used for each race group given data limitations and also because they adequately captured the tendency for Hispanic foreign migrants to be younger than non-Hispanic foreign migrants. They were adjusted to sum to one.
3. The age distribution for Hispanics was split into initial male and female distributions using the age-specific ratios of the male Hispanic to Total Hispanic share and female Hispanic to Total Hispanic share. The same ratios are used for Hispanics for each race group given data limitations and because they adequately captured the variation by sex in the Hispanic population. They were adjusted to sum to one.
4. The age distribution for non-Hispanics was split into initial male and female distributions using the age-specific ratios of the Male non-Hispanic to Total non-Hispanic share and female non-Hispanic to total non-Hispanic share. The same ratios are used for non-Hispanics for each race group given data limitations, and they adequately captured the variation by sex in the non-Hispanic population. They were adjusted to sum to one.
5. The 5-year shares were then allocated into single years using the proportion of each 5-year age group contained in its single year of age. The factors were based on the net foreign migrants from 1995-2000.

Net foreign migration distribution rates were used twice in the development of the total net migration data inputs. First, the rates were used in the process of developing the net domestic migration rates. They were applied to an annualized foreign migration estimate and subtracted

from the annualized total implied net migration to yield annualized net domestic migration. Second, the rates were directly applied to the projected foreign net migration controls up to and including 2050.

4.4.4 Net Domestic Migration Rates

Calculation of the net domestic migration 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

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

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.

4.4.4.1 Implied Migration Method

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

$$\text{Implied Migration} = \Delta\text{Population} - \Delta\text{Births} + \Delta\text{Deaths}$$

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

The steps below delineate how implied migration was calculated:

1. For each 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.

4.4.4.2 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 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. Initial estimate of the net migration by age by each sex and 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 ethnic group.

4.4.4.3 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 counties¹⁰ 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 seventeen¹¹. 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 county.” The adjustment was also based on survival analysis.

5. ACKNOWLEDGMENTS

The EPS 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.

Members of the Council for Technical Solutions:

Anubhav Bagley, MAG
Dr. Robert Carreira, Cochise College
Dr. Alberta Charney, U of A
Tracy Clark, ADOT
Chris Fetzer, NACOG
Mark Griffin, CAG
Randy Heiss, SEAGO
Khaleel Hussaini, ADHS
Caroline Leung of PAG
Dr. Kooros Mahmoudi, NAU
Sharon Mitchell, WACOG
Tom Rex, ASU

¹⁰ Apache, Gila, La Paz, Navajo, and Santa Cruz.

¹¹ The method assumes that all changes in the population are due to migration. Deaths are not considered as they are a small part of the population change for 7-17 year-olds.

Other Contributors:

Bill Schooling, former Arizona State Demographer (January 2009 – January 2012)

Jeff Tayman, University of California, San Diego, Consultant

Scott Bridwell (MAG)

Jesse Ayers (MAG)

Keith Killough (ADOT)

Travis Ashbaugh (CAG)

Jack Tomasik

Many other individuals also reviewed draft projections or attended pre-release presentations. Their comments, criticisms, and encouragements are greatly appreciated.