California fertility: measurement and trends

Ethan Sharygin, California Department of Finance

June 1, 2015: 26th USC-SCAG Demographic Workshop

This brief describes the fertility data used by the Demographic Research Unit, empirical trends in fertility, and methods of fertility estimation for small areas.

California is a geographically and demographically diverse state– in it can be found a microcosm of many of the issues faced by applied demographers around the country. In many counties, the core demographic processes– births, deaths, migration– occur so infrequently that the underlying risks faced by the population cannot be reliably estimated from the raw data.[1]

These data are necessary for evaluating program needs and coverage, for historical analysis, and for generating accurate counts of the current and future population of the state. We address these problems by tackling quality issues in the raw data, and using new model-based estimates for small counts.

[1] Counties are the local unit of analysis we use most– but we have previously been unable to estimate age-specific fertility rates with such granularity.

Fertility data: challenges and innovations

Birth certificates contain an assortment of variables not generally present in aggregate birth statistics.[2] For example, individuals are allowed multiple race responses for themselves and the child's other parent; they may additionally report their education level, birthplace, smoking history, industry and occupation, height and weight, and their participation in pregnancy-related health or nutrition programs. Some percent of these variables are missing each year, but they are each present often enough that multiple imputation of missing responses is feasible (Fig 1).[3]

[2] Data to estimate fertility also exist in the June supplement to the CPS since 1971, the annual ACS since 2006, and decennial Censuses since 1940– see www.census.gov/hhes/fertility/about/

[3] Raghunathan et al. "A multivariate technique for multiply imputing missing values using a sequence of regression models", *Survey Methodology* 27: 2001.

With complete multiple race detail and Hispanic ethnicity imputed, we move on to calculating rates from birth counts. There are at least two sources of complete mid-year population counts for the California population: those from the Demographic Research Unit (DRU) and another set provided by the Census Bureau for use by the National Center for Health Statistics (NCHS).

One issue in the study of fertility is that population counts are not currently generated for all possible multiple race combinations– DRU collects multiple-race individuals in a single catch-all category, and NCHS bridges multiple-race individuals into a single race that they are most likely to choose if forced to self-identify by only one race (using a regression model fit from the National Health Interview Survey where several variations of the race question are asked).[4] Since it

[4] Ingram et al. "United States Census 2000 population with bridged race categories", *National Vital Statistics Reports*, National Center for Health Statistics: 2003.

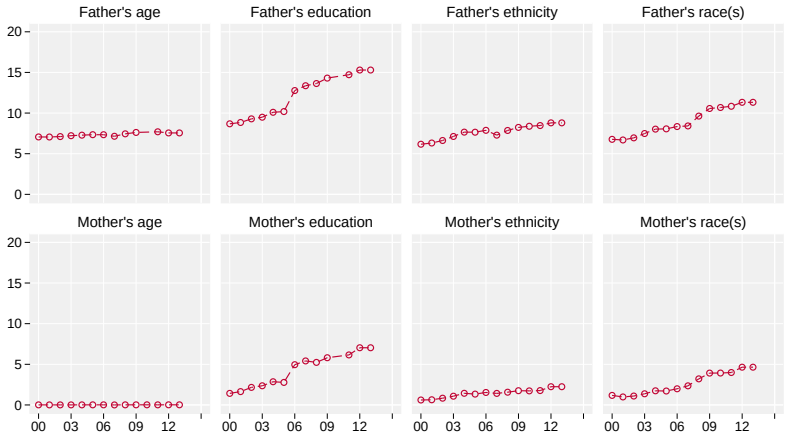


Figure 1: Percent of missing values for key variables in California birth certificates. *Missing values are imputed in the data used for studying fertility.*

is impracticable to track fertility rates for 30 or more combinations of multiple races, the latter approach is preferred here. The bridged race approach is based on the theory that demographic rates are more likely to be correlated with an individual's bridged race; in other words, that multiracial people have more in common with those of a particular other single race than to all other multiracial people.

Fertility in California has been shifting from younger to older ages: not only have teen birthrates been falling, but the birth rate to women ages 20–30 have also been rapidly declining (Fig 2). This is partly made up for by increased fertility at older ages, especially 30–39.[5] For example, fertility of women age 35 used to be half that of those age 20– these rates have nearly converged. Overall fertility has declined since the start of the series, with a brief rebound that coincided with the housing boom and that was experienced mostly by young women.

All groups have seen a decline in fertility during the past decades, with a minor exception being the period of fertility rebound among Hispanic women during 2004-2006. California fertility is below replacement level, which means that the state's population growth comes from net in-migration and from aspects of the age structure that are conducive to population growth (in particular, a youthful population that is subject to high rates of fertility).

Education is another important predictor of fertility that is less often examined. Data on exposure– approximated by mid-year population by education level– has been limited below the state level until recent years. With the ready availability of population data with geographic and education detail from the ACS,[6] More years spent in education and delayed career start may explain a large portion of the

[5] Births to women 40+ have increased rapidly, but remain a very small part of the fertility picture.

[6] ACS is preferred to CPS, despite the latter's longer time series, because of the availability of data below the state level.

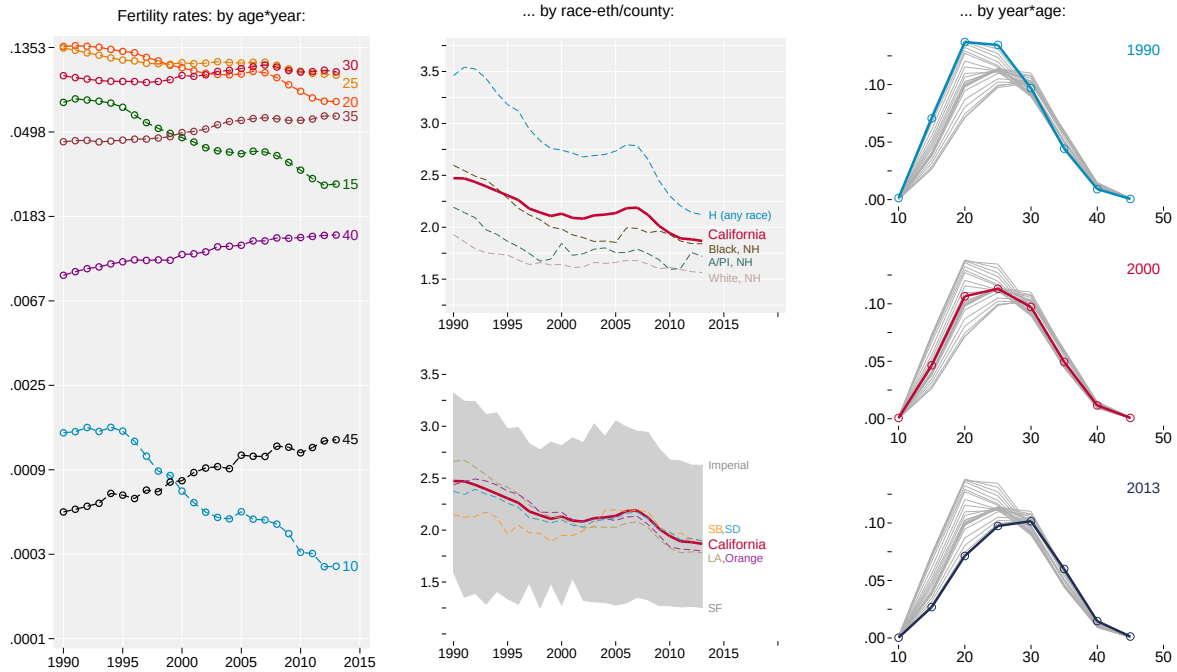


Figure 2: Trends in age-specific fertility rate and total fertility rate: California, 1990–2013. *Race-ethnicity and county figures are total fertility rate (TFR).*

shifting age pattern of fertility to a later modal age of childbearing.

Fertility trends by education since 2007 are presented below (Fig 3). Fertility by education has generally been more stable than by race/ethnicity. Education, then, may be a highly useful lens through which to predict future fertility trends: the story of changing fertility is as much a story of movement between relatively stable age-specific fertility rates by education levels as it is of secular fertility decline. Fertility generally declines with the level of women's education, with the exception that women with an incomplete college education have the lowest fertility of any education group.

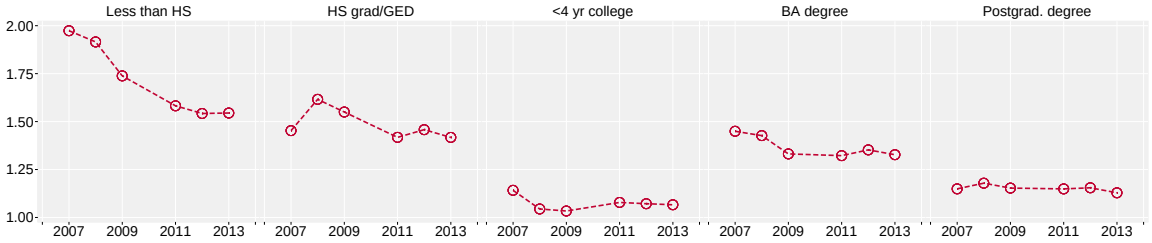


Figure 3: Total fertility rate by education: California, 2007–2013. *Total fertility rate by education is calculated for the population age 25+, thus representing fewer additional lifetime births than TFR for the entire female population age 15-49.*

Model-based fertility

WITH THE COMMON CURRENCY of bridged race for numerators and denominators, we fit a model to the observed rates to enable study of small areas. We use a regression model for count and exposure data with random slopes and intercepts in order to model fertility in small areas where few births occured.[7] The model at calculates a single fertility rate for each race, a overall fertility level for each county, and a coefficient for the effect of county-level covariates such as the poverty rate and percent of women with college education. These effects are estimated using data from the 5 years preceding each individual year, and combined into a count of predicted births for each bridged race and county population. As an initial assessment of model fit, a table of the model-generated versus actual births is included here (Table 1).

[7] a similar model is described and documented in Kulkarni et al., "Falling behind: life expectancy in US counties from 2000 to 2007 in an international context", *Popul Health Metr*, 9:1, 2011.

Year	Actual Births	Model-Based	Difference (%)
2001	525,000	522,362	0.50
2002	526,917	526,379	0.10
2003	538,585	533,708	0.91
2004	542,576	539,732	0.52
2005	546,613	543,761	0.52
2006	560,126	556,105	0.72
2007	562,874	564,034	0.21
2008	548,296	560,147	2.16
2009	524,967	543,815	3.59
2010	509,979	522,638	2.48
2011	500,337	502,125	0.36
2012	502,105	493,452	1.72
2013	502,105	493,865	1.64

Table 1: Actual and model-based fertility, California 1995-2013. *Note the evidence of delayed fertility 2008–2013. The model over-predicted births during the recession and under-predicted during the recovery. The current model does not include business cycle information. Future iterations will include processes that can capture economic interactions with fertility.*

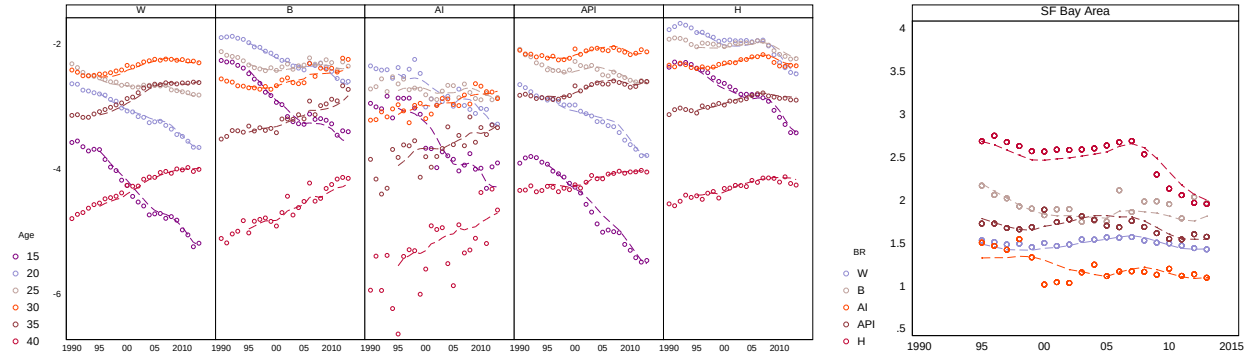
For counties with large populations, the model lends heavy weight to the observed trends (Fig 4). In California's sparsely populated rural counties, the model borrows strength from neighboring counties, and from data on individuals of the same age and race in other regions. For the purposes of presentation, the following figures collapse the county estimation results into 8 ad-hoc regions: SF Bay Area, Greater Sacramento, Far Northern, Central Valley, Sierra Nevada, Central & South Coasts, and Inland Empire (Fig 4).

Figure 4: Empirical and model-based fertility: regions of California, 1990–2013. *Model-based fertility estimated at county level and aggregated by region; see following pages for figures.*

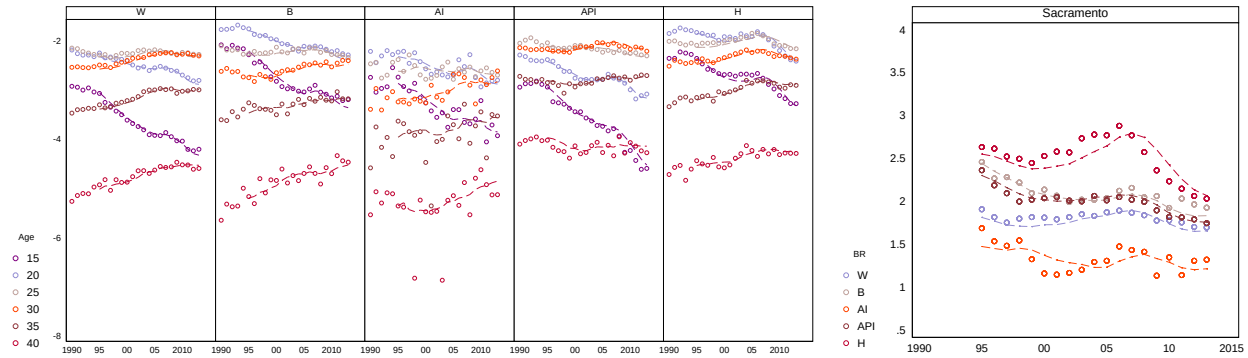
A FINAL USE of these data is to anticipate future trends in fertility. At present, there is no consensus theory in the social sciences that could tell us whether fertility should recover in the future or continue to fall. Given results of ongoing research into the causes of fertility decline, there may be slightly more evidence favoring reasons for continued decline.[8] Projections to date have assumed stable or recovering fertility. These assumptions will be revisited in future revisions.

[8] W Lutz et al., "The Low Fertility Trap Hypothesis: Forces that may lead to further postponement and fewer births in Europe." *European Demographic Research Papers*, 4:5, 2006.

SF Bay Area: ASFR and TFR by Mother's Bridged Race, 1990-2013

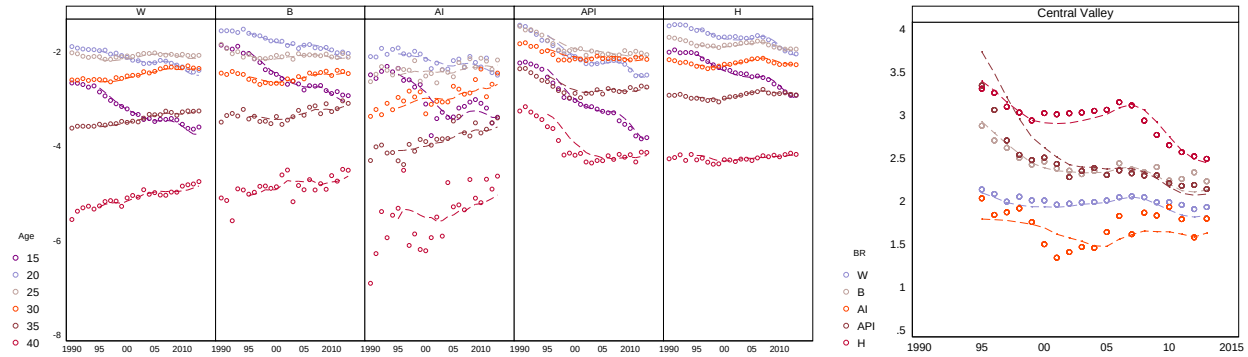


Sacramento: ASFR and TFR by Mother's Bridged Race, 1990-2013

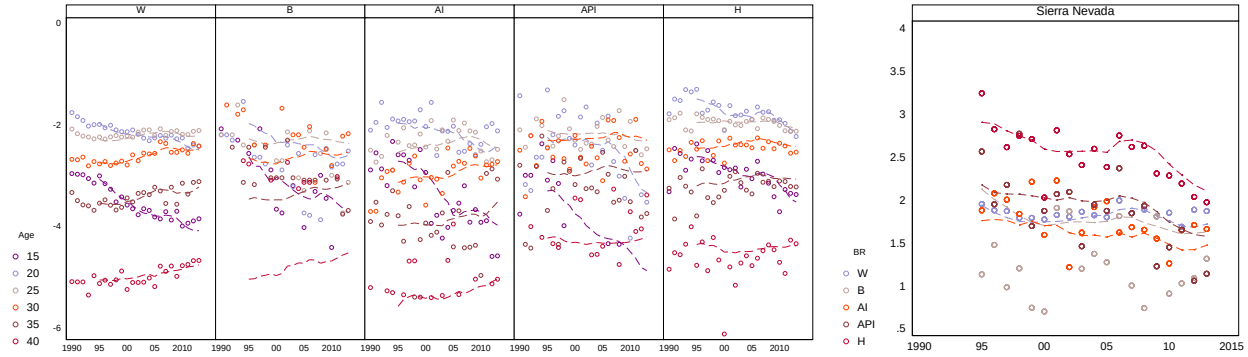


Far North: ASFR and TFR by Mother's Bridged Race, 1990-2013

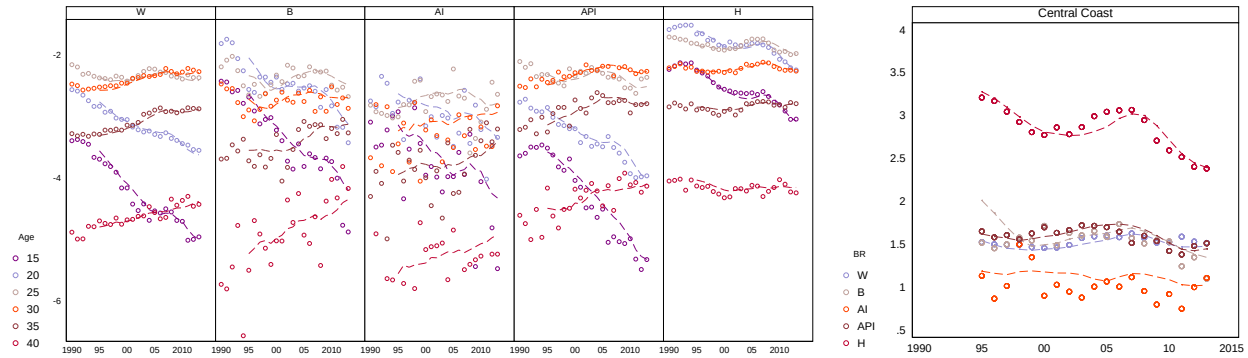
Central Valley: ASFR and TFR by Mother's Bridged Race, 1990-2013



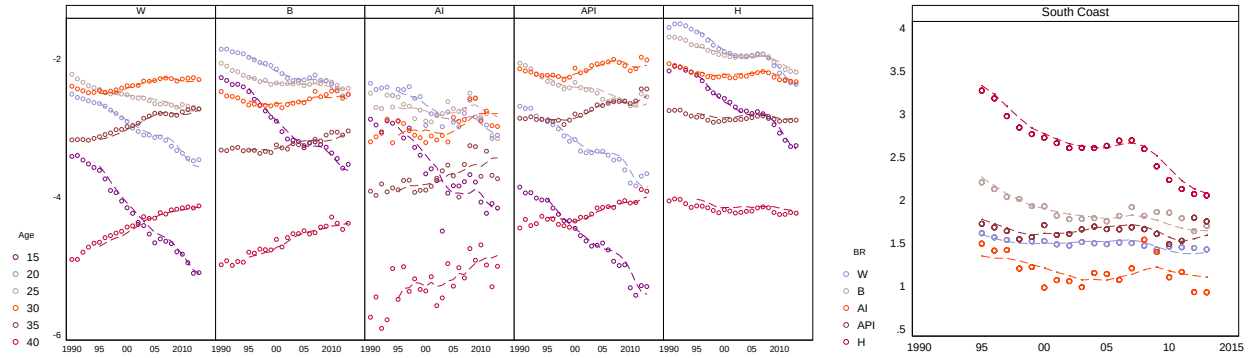
Sierra Nevada: ASFR and TFR by Mother's Bridged Race, 1990-2013



Central Coast: ASFR and TFR by Mother's Bridged Race, 1990-2013



South Coast: ASFR and TFR by Mother's Bridged Race, 1990-2013



Inland Empire: ASFR and TFR by Mother's Bridged Race, 1990-2013