

## A Predictive Estimate of the 2010 Census Count for California

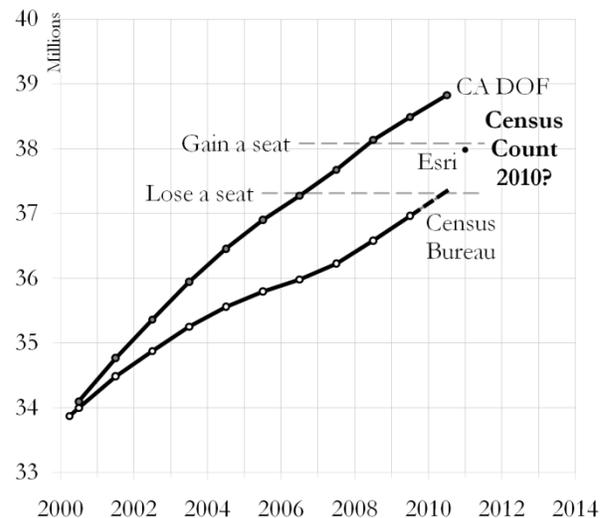
by *John Pitkin & Dowell Myers*

Great uncertainty surrounds the size of the California population, as illustrated in Figure 1. There is a 1.5 million difference between the estimate prepared for the state by the Census Bureau (36,961,664 in July 2009) and the estimate prepared for the same date by the Demographic Research Unit of the California Department of Finance (38,487,889 in July 2009). Note that the DOF has newly issued a July 2010 estimate of 38,826,898 but there is no comparable figure for comparison prepared by the Census Bureau. Meanwhile the Esri mapping and data company has prepared a April 2010 estimate of 37,983,948 for California's population. An analysis by Election Data Services based on this figure showed that California would retain its 53 seats in the House of Representatives, but if 99,396 additional residents were counted in the state, California would gain an additional seat. Conversely, if 668,713 fewer residents than expected were counted, the state would lose a seat. A smaller population would also equate to millions fewer dollars in revenue sharing and other economic impacts. Thus the actual size of the California population has great political and economic significance.

On December 21, the Census Bureau will report the 2010 census count for all states, resolving the political uncertainty about reapportionment. One might assume that the count of residents will fall somewhere in the middle of the various estimates, but there is concern that net coverage of the population might be lower than in the 2000 census, thus yielding a lower population total than suggested by any of these estimates. If this undercoverage turns out to be equal in all states, the relative distributions of political representation and revenue sharing would be unaffected.

Even if this happens, however, an understanding of recent population changes in California and their implications for future growth requires consistent estimates of the population over time. Ideally these estimates would be true counts, but in practice we rely on estimates that have a fixed, constant level of coverage of the population. Both the DOF and Census Bureau current population estimates work forward from the same Census 2000 base population, which serves as a key benchmark for coverage of the population. Yet the latest estimates in these two series differ so widely that at least one of them, and possibly both, must now deviate substantially from the Census 2000 standard of population coverage and, more importantly, must have substantially under or over-estimated the amount and pace of population change since 2000.

**Figure 1.** Uncertainty of the California 2010 Census Count & the Divergent Annual Estimates Post-2000



Source: U.S. Census Bureau, California Department of Finance, & Election Data Services

The 2010 Census will provide another key benchmark for gauging changes, but will not provide all the answers about population change because of its limited content, due to a much-restricted questionnaire, and because it will embed a new coverage ratio that may be less complete than in 2000. Without the long-form questions that were dropped, the 2010 Census will shed little light on the numbers of immigrants and migrants from other states to California, which are major components of change since the 2000 census.

An alternative means of learning the size of the population at the national level has been newly applied by the Census Bureau. Independently of the census count and its own current estimates, the Bureau has developed a formal demographic analysis (DA) that produced alternative estimates of the U.S. population on April 1, 2010 (released December 6, 2010). This demographic analysis was prepared, as for previous censuses going back to 1940, by tracking the number of recorded births in previous years, subtracting the cumulative number of recorded deaths, and accounting for international in and out-migration. The demographic analysis amounts to an historical accounting of the population and offers an independent criterion for gauging the accuracy and coverage of the census count, both in total and for specific populations defined by sex, race, age, and, for the first time in 2010, Hispanic origin. More importantly for our purposes, DA also provides new insight into the sources of population change since the last census in 2000.

The new Census Bureau demographic analysis relies on estimates of immigration from the American Community Survey (ACS), a source of information not available in earlier decades. In the demographic analysis for the 2000 Census the international migration estimates for 1990-2000 were especially problematic due to the large number of undocumented immigrants that were counted. Since the ACS covers all immigrants, regardless of whether they have legal documents from the Immigration Service

(and are therefore counted in the Office of Immigration Statistics data), the ACS estimates of immigration are expected to prove more accurate and complete than those in earlier decades.

It should be understood that neither the ACS nor its predecessor survey, the Census 2000 Supplemental Survey, can be used as an independent check on the Census Bureau's current population estimates, since both are controlled to those estimates. They do however provide a new and reliable basis for estimating both immigration and domestic migration at the state level for the 2000-2009 period, and these are key building blocks for population estimates derived by our predictive method based on demographic analysis.

### **Constructing a Components and ACS-Based Estimate for California**

This report introduces two parallel estimates of 2000-2010 demographic changes and the resultant April 1, 2010 Census day population for the state of California. As with the earlier California estimates cited here, this estimate is benchmarked to the 2000 census, but unlike earlier estimates, it works forward from the Census 2000 population incorporating nine and a quarter years of migration estimates from the ACS (and its predecessor, the Census 2000 Supplementary Surveys). To this is added in births and subtracted out deaths, both of which are reported in vital records data. The remaining components are filled in with modeled rates derived from analysis of survey data and other sources. The components for periods not covered by these data (Jan.-Mar. 2010 for births and deaths and mid 2009-2010 for migration) are estimated. Our estimates of international migration adopt two of the ACS-based series used by the Census Bureau in its DA estimates of the national resident population on Census Day.

The ACS is the new *de facto* standard for measuring migration at the state as well as national level, domestic as well as foreign. Migration, prior residence and place of birth are no longer asked in the Census itself. Unlike the Census, the ACS provides annual data on mobility—not once a decade—and offers more consistent, continuous measures of migration flows than we have had before, e.g., not just income tax filers, and includes vital demographic details. It is likely to become recognized as the most authoritative source for measuring migration components. There are however questions about the reliability and interpretation of these data, and, for immigration, which item to use. The present study's integration of ACS-based migration data into the population via demographic accounting provides a basis for cross-checking all the components with one another and with the new census.

The analysis presented here amounts to a prediction of what should be the count in California, given the observed or estimated demographic components, and given the assumption that the coverage ratio in 2010 is the same as in the 2000 census. This predictive analysis constructs an estimate by building the population from its components, one year at a time, working forward from the Census 2000 base period. The variance between our estimate and the count reported by the Census Bureau is subject to postanalysis with regard to each of its underlying components. With release of fuller census details in spring and summer 2011 it will be possible to investigate which age groups, genders, or race/ethnic groups account for the variance between the predicted and actual counts.

The present analysis is limited to the total population, Hispanic or Latino population, and non-Hispanic population, all separately calculated for males and females. We have chosen to present the estimates in terms of Hispanic or Latino identity because Latinos account for a large share of the population growth in California, including roughly half of all births and half of all

immigration. Estimates for the other major racial groups will be reported in coming months.

Two alternative estimates are provided, varying only in the components for net domestic migration, foreign-born immigration, and foreign-born emigration. Also captured is the very small number of U.S. citizens at birth who move into or out of the United States. Even with the ACS there is substantial uncertainty about these various immigration and migration components, due not only to sampling variability, but also to respondent error or artifacts of questionnaire design. The ACS includes two different questions that yield information about movement into and out of California. One question asks place of residence one year ago (ROYA), which is useful for measuring internal migration of residents between states as well as from abroad. The other question asks year of entry for migrants from abroad (YOE). The latter question yields an estimate of immigration in the last year that is somewhat higher than yielded by the question on ROYA. The Census Bureau's new demographic analysis estimates also uses these two questions in alternative series and, following the Bureau, we also make use of the sampling variability in the resulting data to help define a higher estimate series. Whereas the Bureau produces five alternative series, two of which are fairly extreme and one of which merely occupies the middle, we have elected to offer only two series, one credible on the high side and the other on the low side. Our upper estimate for California uses the upper bound of a 90% confidence interval for the ROYA question on domestic migration and the upper bound on the YOE question for immigration. Our lower estimate uses the reported results of the ROYA question for immigration (which is lower than reported by YOE) and uses the reported results for ROYA for domestic migration. In constructing our lower population estimate we do not make use of a lower bound of the confidence interval for this question because there is substantial credible information of higher domestic migration from other sources,

such as emphasized in the DOF estimates of the California population.

## Comparing Estimates of the 2010 Census Count for California

Table 1 presents our two estimates of the population of California, showing the components of change between Census Days in 2000 and 2010. The high estimate is 38,090,000 and the low estimate is 37,017,000. The spread of 1,073,000 is about one-third less than the 1.5 million difference between the Census Bureau and DOF estimates for California's population. When compared to the divergent estimates shown in Figure 1, our low estimate is very close to the Census Bureau estimate. Compared to the Bureau's latest estimate, when projected from July 2009 to April 2010, our low estimate is lower by 231,000, less than one percent ( $-0.62\%$ ) below the Bureau estimate. Conversely, our high estimate lies closer to the DOF estimate, falling about 652,000 below that ( $-1.68\%$ ). Thus, the one million spread between our high and low estimates corresponds to the lower half of the spread between DOF and Census Bureau estimates and hangs 231 thousand below it.

The major factors producing the divergence between our upper and lower estimates are related to domestic migration and immigration. Under the higher estimate, we estimate 414,000 less domestic outmigration from California over the last decade than expected under the lower population estimate. This is about one-quarter less domestic outmigration. Under the higher scenario we also estimate 699,000 more immigration over the last decade, about one-third more than estimated under the lower population estimate. The explanations for these differences were explained above as the choice of alternative measurements in the ACS.

The above differences between high and low estimates are computed to be roughly similar for Hispanic and non-Hispanic populations. Under either estimate series, the Hispanic popu-

lation accounts for a larger share of the components of growth than its share of the total population. Whereas, Latinos were 32.4% of the state's population in the 2000 census, they have accounted for 48.0% of the births and only 14.1% of deaths. Being younger, they are also much more geographically mobile and account for about 42% of the net outmigration and 47% of the immigration. The net result over the decade is that we estimate that the Hispanic population increased its share of the total in the state from 32.4% in 2000 to 36.2% or 36.5% in 2010, under our low and high estimate series, respectively. This estimate of the Hispanic share is slightly lower than what might have been expected for 2010 based on the most recent estimates prepared by either the Census Bureau (37.0%, 2009) or the Department of Finance (37.2%, 2008). The difference might be partially explained by the sharp decline in both births and immigration in the last two years of the decade, trends not yet incorporated in other estimates.

It remains to be seen where the reported census count for California will fall, whether it will lie within the spread between our two estimates or possibly fall below both the Census Bureau estimate and our own low estimate. Comparison to the Esri population estimates relied upon by Election Data Services for estimating apportionment impacts suggests potentially significant impacts. If the California census count falls to our lower estimate, and if other states' populations hold as expected, a census count as low as our lower estimate would cause California to lose one seat in the House of Representatives. Conversely a count that reached our upper estimate would produce a gain of one seat. However, we expect that if either extreme of our two estimates were achieved, that would likely indicate a similar directional shift in the count of other states, mitigating to some extent the relative shift in California.

**Table 1.** Estimates of the Population of California on April 1, 2010, and Components of Change for 4/1/2000 to 4/1/2010, by Hispanic Origin and Sex (in thousands)

	Total population			Hispanic		
	Both sexes	Male	Female	Both sexes	Male	Female
<b>Low Series</b>						
<b>Total Population on 4/1/2000 (Census)</b>	<b>33,872</b>	<b>16,875</b>	<b>16,997</b>	<b>10,967</b>	<b>5,615</b>	<b>5,352</b>
Births	5,449	2,779	2,670	2,618	1,342	1,276
Deaths	2,350	1,162	1,188	331	186	144
Net domestic migration	(1,512)	(759)	(753)	(643)	(354)	(289)
Foreign-born immigration	2,139	1,080	1,059	1,010	550	460
Foreign-born emigration	563	311	252	206	126	80
Net native migration*	(17)	(26)	10	(21)	(16)	(5)
<b>Total Population on 4/1/2010 (CDP low estimate)</b>	<b>37,017</b>	<b>18,475</b>	<b>18,542</b>	<b>13,393</b>	<b>6,825</b>	<b>6,569</b>
<b>High Series</b>						
<b>Total Population on 4/1/2000 (Census)</b>	<b>33,872</b>	<b>16,875</b>	<b>16,997</b>	<b>10,967</b>	<b>5,615</b>	<b>5,352</b>
Births	5,449	2,779	2,670	2,618	1,342	1,276
Deaths	2,350	1,162	1,188	331	186	144
Net domestic migration	(1,099)	(553)	(545)	(463)	(256)	(208)
Foreign-born immigration	2,838	1,433	1,405	1,341	730	611
Foreign-born emigration	603	333	270	221	135	86
Net native migration*	(17)	(26)	10	(21)	(16)	(5)
<b>Total Population on 4/1/2010 (CDP high estimate)</b>	<b>38,090</b>	<b>19,011</b>	<b>19,079</b>	<b>13,889</b>	<b>7,094</b>	<b>6,796</b>

Notes: Net native migration includes born abroad of American parents, Armed Forces overseas, and migration to and from U.S. possessions.

Source: USC Population Dynamics Research Group, 2010 Census Year California Profile Project, John Pitkin & Dowell Myers.

## References

California Department of Finance, "California County Race / Ethnic Population Estimates and Components of Change by Year, July 1, 2000–2008." Series E-3. Sacramento, California, June 2010.

California Department of Finance, "California County Population Estimates and Components of Change by Year, July 1, 2000-2010." Series E-2. Sacramento, California, December 2010.

Election Data Services. 2010. "New Population Estimates Show Slight Changes For 2010 Congressional Apportionment, With A Number of States Sitting Close to the Edge," October 15, 2010, accessed December 10, 2010 from [www.electiondataservices.com/images/stories/pdf/NR\\_Appor2010ESRI\\_finalwTableMap.pdf](http://www.electiondataservices.com/images/stories/pdf/NR_Appor2010ESRI_finalwTableMap.pdf)

Robinson, J. Gregory. 2010. Coverage of Population in Census 2000 Based on Demographic Analysis: The History Behind the Numbers. Paper presented at the U.S. Census Bureau: 2010 Demographic Analysis Technical Review, Suitland, MD, Jan. 8, 2010.

U.S. Bureau of the Census. 2009. "Methodology for the State and County Total Resident Population Estimates (Vintage 2009): April 1, 2000 to July 1, 2009." Washington, DC. Accessed at <http://www.census.gov/popest/topics/methodology/2009-st-co-meth.pdf>

U.S. Bureau of the Census. 2010. *ACS Design and Methodology Report*. Revised. Washington, DC. Accessed at [http://www.census.gov/acs/www/methodology/methodology\\_main/](http://www.census.gov/acs/www/methodology/methodology_main/)

U.S. Bureau of the Census. 2010. "The Development and Sensitivity Analysis of the 2010 Demographic Analysis Estimates." Released Dec. 6, 2010.

## John Pitkin

John Pitkin is senior research associate in the USC Population Dynamics Research Group and the principal architect of the predictive estimates and the California Demographic Futures projections. He also is president of Analysis and Forecasting, Inc., Cambridge, Massachusetts

## Dowell Myers

Dowell Myers is professor in the School of Policy, Planning, and Development and director of the USC Population Dynamics Research Group. He is the principal investigator directing several projects on housing and demographic change, immigrant assimilation, and demographic projections.

## Acknowledgments

Mi Young Kim provided outstanding research assistance with the American Community Survey. The authors also thank Sarah Mawhorter for her assistance with report design and for her great patience in working under tight deadline during a holiday period. The Haynes Foundation is gratefully acknowledged for supporting primary research on the consistency of data informing the 2010 census year.