

Who's right? A novel use of survey and administrative data to assess
state population estimates

Hans Johnson, johnson@ppic.org
Joseph Hayes, hayes@ppic.org
Public Policy Institute of California

The Census Bureau and many states report annual estimates of state populations. Some states rely on the Census Bureau's estimates, while others produce their own estimates. The estimates are used for many purposes, including revenue disbursement and program planning. In some states, the estimates are used to develop expenditure caps. In this study, we exploit both the large sample size of the American Community Survey and a survey item that is new to the survey (as compared to the decennial census) to evaluate the accuracy of state population estimates for highly populated states. Specifically, we compare state vital statistics data on the number of women giving birth to the American Community Survey findings regarding the number of women giving birth in the prior year. Births and the number of women giving birth are thought to be universally recorded, or nearly so, in state administrative data. If a state's vital statistics totals are higher than those derived from the American Community Survey, then we have some evidence that the Census Bureau population estimate for the state, at least among this segment of the population, is too low. Alternatively, if the vital statistics totals from state administrative data are lower than those of the Bureau, we have evidence that the Bureau's estimates are too high. By reweighting the ACS sample using state population estimates, we can similarly evaluate estimates produced by the states.

This approach is straightforward and simple in concept, although complicated in practice. The American Community Survey asks respondents whether any females in the

household ages 15 to 50 have given birth in the past year. Census Bureau population estimates are used to weight the responses, and thus we can compare the weighted totals of the number of females giving birth based on the American Community Survey with administrative data to assess state population estimates. However, survey responses to the question on whether or not a woman gave birth in the past year might suffer from a number of biases, including telescoping and perhaps even recall bias. In this paper, we consider whether and for which groups such biases exist, identifying both the extent and direction of the biases that would be necessary to alter our findings.

For the nation's most populated state, California, where the discrepancy between the Bureau's estimate and the state's own estimate is particularly large, we use our method not only to determine which estimate appears to be more accurate, but also to evaluate potential biases in the survey responses. California's vital records on births have the advantage of containing a rich set of demographic variables, including educational attainment, parity, nativity, and racial and ethnic classifications that are aligned to census categories. The state's large population results in a large sample of women giving births; the 2005 ACS includes over 5,000 women who reported giving birth in the prior year.

Estimates produced by the Census Bureau and the state agency responsible for developing estimates, the Demographic Research Unit of the Department of Finance, differ widely. The Census Bureau estimates the state's population at 36,458,000 as of July 2006, almost one million lower than the California Department of Finance estimate of 37,444,000 for the same date. Population growth implied by these projections differ even more, of course, with the state's own estimated growth since the 2000 census almost 40 percent higher (3.6 million versus 2.6 million).

The first section of the paper will describe methods used by the Census Bureau and state agencies to develop population projections. The second section will identify which states develop their own estimates and will compare Census Bureau estimates with states' own estimates. The third section describes our data and methods in detail, while the fourth applies those methods to estimates for the largest states. The fifth section of the paper considers the case of California in detail.

This paper will highlight how the American Community Survey can be used by demographers to evaluate state population estimates. The results of this research could have important implications for improving estimates of state populations, and will shed new light on the accuracy of those estimates.