Improving Health Insurance Estimates: Using the American Community Survey to Estimate Health Insurance Coverage for Counties in the United States

by Bethany DeSalvo, Mark Bauder, and Sam Szelepka

This paper investigates the health insurance estimates released by the U.S. Census Bureau's Small Area Health Insurance Estimates (SAHIE) program. These model-based estimates provide data on the number of people with and without health insurance coverage for every county in the United States by age, sex, and income group. The program recently changed from using the Current Population Survey measures of health insurance coverage to using measures from the American Community Survey (ACS). This change substantially lowered the margins of error for the estimates. This paper examines these policy-relevant estimates using demographic and spatial methods, including the new income-to-poverty ratio categories pertinent to the Patient Protection and Affordable Care Act.

Model-based estimates provide more precision than survey-only estimates. The modeling utilized by the SAHIE program incorporates covariates from census and administrative data that improve the variance of the estimates. The SAHIE model expands on the Fay-Herriot model typically used for small area estimation (Fay & Herriot, 1979). The model is formulated in a hierarchical Bayesian framework. The posterior means serve as the point estimates and both the posterior means and variances integrate for a normal approximation to calculate symmetric 90-percent confidence intervals. The half-widths serve as the margins of error. The datasets utilized in this model include the ACS, demographic population estimates, aggregated federal tax returns, participation records for the Supplemental Nutrition Assistance Program (SNAP), formerly known as the Food Stamp program, County Business Patterns, Medicaid and Children's Health Insurance Program (CHIP) participation records; and Census 2000. The models are area-level because they invoke aggregated survey estimates and administrative data rather than individual survey and administrative records.

The model produces estimates of health insurance coverage for demographic and income groups within counties and states. Estimates by sex (female, male, both), race/ethnicity (all races, non-Hispanic White, non-Hispanic Black, Hispanic), age (0-18, under 65, 18-64, 40-64, 50-64), and income groups are provided. Also, county estimates by the same age, sex and income groups, but

not by race are estimated. Income groups are defined by the income-to-poverty ratio (IPR) – the ratio of family income to the appropriate federal poverty level. Estimates are produced for the following IPR groups: all income, 0-138%, 0-200%, 0-250%, 0-400%. States classify a threshold of either 200 or 250 percent of the federal poverty level as being eligible for state provided healthcare services. For the past few years, SAHIE has provided states with either the 200 or the 250 IPR, based on the requirements of each state. However, for 2008 and 2009, the incorporation of the ACS into our estimates allowed us to provide precise estimates of 200 percent and 250 percent, plus two additional IPR groups of 138 and 400, which are relevant to the Patient Protection and Affordable Care Act for each state. This Act helps families gain access to health care by allowing Medicaid to cover families with incomes up to 138 percent of the poverty line. Also, families with incomes above the level needed to qualify for Medicaid but below 400 percent of the poverty line can receive tax credits that will help them pay for health coverage in the health insurance exchanges.

When the SAHIE program started in 2005, the Annual Social and Economic Supplement of the Current Population Survey (CPS ASEC) were the only health insurance survey estimates available from which to build model-based estimates. SAHIE incorporated three years of CPS ASEC averaged estimates. Unfortunately, the CPS ASEC data could not produce reliable estimates for all counties. Not all counties are in the survey, and most counties in the survey have small sample sizes. In 2008, the ACS included a health insurance question which provide insurance estimates for all geographic areas with a population of 65,000 or more (U.S. Census Bureau, 2009). The ACS contains approximately 30 times the number of addresses as those from the single-year CPS ASEC. Direct survey estimates based on a 5-year accumulation from the ACS for geographic areas with population of 20,000 or less will not be available until 2013.Ohara & Bauder (2010) compared the reliability of estimates from a CPS-based and an ACS-based model for state and observed the ACS provides more useful or refined estimates of uninsured populations for policy makers and other stakeholders. We have found similar results for the county-level data.

Improving these estimates is important for several reasons, including the impact that understanding the coverage of health insurance with a precise measure for every county in the

United States (except Kalawao, HI) is important for public policy. Production of these estimates is partially funded by the Breast and Cervical Cancer Early Detection Program (NBCCEDP) of the Centers for Disease Control (CDC). The CDC requires an estimate of the percent eligible and participation rates for screening crossed by various demographic groups.

Fay, R., and Herriot, R. (1979). Estimates of income for small places: An application of James-Stein procedures to census data. *Journal of the American Statistical Association, 74,* 269-277.

O'hara, B., and Bauder, M. (2010). Databases for Estimating Health Insurance Coverage for Children: A Workshop Summary. Committee on National Statistics: CNSTAT. http://www.nap.edu/openbook.php?record_id=13024&page=165.

U.S. Census Bureau. (2009). *Design and Methodology American Community Survey.* Washington, DC: U.S. Government Printing Office.