Accuracy of Medicaid reporting in the ACS: Preliminary results from linked data

Michel Boudreaux[†], Kathleen Thiede Call[†], Joanna Turner[†], Brett Fried[†], Brett O'Hara[‡]

† State Health Access Data Assistance Center
2221 University Avenue SE
Minneapolis, MN 55414

‡ U.S. Census Bureau
4600 Silver Hill Rd
Washington, MD 20233

Proceedings of the 2013 Federal Committee on Statistical Methodology (FCSM) Research Conference

Acknowledgements

This research is funded by a contract with the U.S. Census Bureau. We appreciate the contributions and guidance of staff at the U.S. Census Bureau: Jennifer Cheeseman Day (Social, Economic, and Housing Statistics Division) and Amy O'Hara, Leticia Fernandez, James Noon, and Sonya Rastogi (Center for Administrative Records Research and Applications).

Disclaimer: This paper is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

Introduction

The "Medicaid undercount" refers to the discrepancy between administrative counts of Medicaid enrollment and estimates from survey data. Nearly all state and federal surveys estimate fewer Medicaid enrollees than is described by enrollment records. In the 2005 Current Population Survey (CPS) 40.8% of people known to have Medicaid from administrative records were not reported as having Medicaid in the CPS (SNACC Phase V, 2010). In the 2002 National Health Interview Survey (NHIS) the undercount was 33.5% (SNACC Phase IV, 2009) and in the 2003 Medical Expenditure Panel Survey Household Component (MEPS/HC) it was 17.5% (SNACC Phase VI, 2010).[1] State sponsored surveys often fare better, but still tend to undercount Medicaid by 12-26% (Call, Davern, Klerman & Lynch, 2013). This mismatch suggests that survey data, relative to what is known from administrative data, give biased estimates of key policy measures such as the share of the population covered by Medicaid or those lacking health insurance. Although survey data on Medicaid is likely biased, unlike administrative data, surveys provide a wide array of policy relevant covariates such as access to health services, health status, and race and ethnicity, and surveys are the only source of information about the uninsured and the eligible but not enrolled.

This working paper presents preliminary results from a collaboration between SHADAC and the U.S. Census Bureau that extends prior research on the CPS, NHIS and MEPS/HC (Call et al., 2013; Davern, Klerman, Baugh, Call, & Greenberg, 2009a; SNACC Phases II-VI, 2008-2010)[1] to the American Community Survey (ACS). The ACS began collecting information on health insurance coverage in 2008. Since that time the ACS has become an important source of information for monitoring health insurance coverage and evaluating health policy.[2] The ACS is a unique data asset because its large sample size makes it possible to produce statistically reliable single year health insurance estimates at the national, state and sub-state levels (Davern, Quinn, Kenny & Blewett, 2009b). It can also

[1] There are six SNACC reports (Phase I – Phase VI) conducted as part of a multi-agency collaboration linking surveys to administrative data. The agencies include SHADAC, the Centers for Medicare and Medicaid Services, The Department of Health and Human Services Assistant Secretary for Planning and Evaluation, the National Center for Health Statistics, and the U.S. Census Bureau. The estimates for the NHIS, CPS and MEPS/HC come from Table 3 of the Phase IV, V and VI reports that include only explicit responses (imputed and edited values are excluded).
[2] See descriptions of the use of the ACS for monitoring health insurance at the state level from SHADAC (http://www.shadac.org/shap/technical-assistance/data/chartbooks) and Kaiser (http://kff.org/uninsured/state-indicator/total-population-2/) and for use creating inputs into models of how the Affordable Care Act (ACA) will impact states see Sonier 2012 and Buettgens et al. 2013.

be used to monitor important, but relatively small sub-groups like minority children in poverty. Other federal surveys lack the necessary sample size to adequately monitor these geographic and demographic groups on an annual basis.

Compared to more complicated questionnaires like the NHIS or MEPS/HC, the ACS has a simpler health insurance question which could potentially contribute to misclassifying Medicaid enrollees – it lacks state-specific program names, lacks a verification question[3], and has a "laundry list" response option. However, preliminary results in this working paper provide evidence that the ACS "undercount" is in line with other surveys that measure health insurance coverage. Yet caution should be used when comparing ACS results with other surveys as the ACS question captures Medicaid and other means-tested coverage, without the ability to separate out Medicaid coverage like the other surveys.

The only other evidence on the extent to which the ACS misclassifies coverage is limited to the 2006 ACS content test (O'Hara 2009).[4] This working paper is the first research on the Medicaid undercount using the full production ACS. In this paper we describe the coding of Medicaid and other means-tested insurance in the ACS and people known to be enrolled in Medicaid or expansion Children's Health Insurance Program (CHIP) coverage on the day of the survey according to administrative enrollment data. Results are presented by broad demographic characteristics (i.e., age, poverty, and state of residence). We also report the upper bound of bias to estimates of uninsurance attributable to misclassification of Medicaid.

Methods

To gauge the extent of the Medicaid undercount in the ACS, we link the 2008 ACS to the Medicaid Statistical Information System (MSIS) using a procedure developed by the Census Bureau that anonymizes the data and preserves confidentiality so it can be used for statistical purposes and research. Our approach to the analysis, outlined below, is consistent with prior linkage projects exploring the undercount in federal surveys (Davern et al., 2009a; Call et al., 2013).

MSIS: The MSIS data contain days enrolled per month for people in Medicaid and CHIP programs that operate as part of a state's Medicaid program (so called expansion CHIP or M-CHIP). The MSIS tracks enrollment in all benefit categories including full benefits and partial benefits (i.e., only specific services such as emergency care are provided). States provide these data to the Centers for Medicare and Medicaid Services (CMS) which in turn cleans the data and notes unresolvable anomalies (Czajka, 2012). Prior to linking the MSIS to the ACS, we de-duplicate the MSIS so that each record corresponds to an individual person and we remove deceased individuals.

ACS: The ACS is a general population survey conducted by the Census Bureau on an annual basis. The survey is primarily conducted by mail, with telephone and in-person non-response follow up. The Census Bureau's internal version of the 2008 ACS contains over 4 million person records. To ensure the confidentially of respondents we use the internal version of the ACS, accessed through the Census Bureau's Minnesota Research Data Center.[5]

The health insurance question shown in Figure 1 collects information about health insurance status on the day of interview. The question classifies people into one or more insurance types and those who report not having any type listed (except for Indian Health Service which is not considered to be comprehensive coverage) are considered to be uninsured. Those that are coded "Yes" to item (d.) are considered to have Medicaid or another means-tested public coverage type. In this working paper our focus is on survey response. Therefore, we include only explicit reports of Medicaid or another means-tested public coverage (imputed and edited responses are omitted).[6] Our estimates reflect the civilian noninstitutionalized population (U.S. Census Bureau, 2008).

[3] Many surveys ask respondents that do not report any health insurance type to verify that they lack any coverage. The ACS does not.
[4] In January through March of 2006, the ACS conducted a test of new and modified survey content prior to implementation in 2008. Two versions of a health insurance coverage question were evaluated to determine which worked best in an ACS questionnaire environment (Nelson and Ericson, 2007).
[5] While our analysis is based on the internal, restricted use ACS data, the Census Bureau provides a public version of the ACS which has been augmented to protect confidentiality. The so called "Public Use Microdata Sample (PUMS)" is accessible from American Fact Finder or the Integrated Public Use Microdata Series (IPUMS).
[6] We include the same tables of estimates in Appendix B for all values including imputations and edits.