Under-reporting of means-tested program receipt in the American Community Survey: The case of California

Sarah Bohn, PPIC
Caroline Danielson, PPIC
Matt Levin, PPIC
Shannon McConville, PPIC

Latest draft: September 27, 2013

Survey under-reporting of means-tested programs is well-documented (Wheaton, 2007; Meyer, Mok and Sullivan 2009). This paper investigates the degree of under-reporting in the Supplemental Nutrition Assistance Program (SNAP) and Temporary Aid to Needy Families (TANF) programs for the California sample of the American Community Survey (ACS), survey years 2009 through 2011. To assess under-reporting, we compare self-reports aggregated up to cells defined by geographic location and by demographic characteristics to detailed California administrative data on participation. We focus on the ACS since it is a relatively new survey. As a result, the extent of under-reporting is not as well validated as other household surveys such as the Current Population Survey (CPS) and the Survey on Income and Program Participation (SIPP). The ACS, however, is becoming more widely used due to its substantially larger sample size and the ability to produce estimates at smaller geographic levels, such as states and counties. For this reason, the ACS is often used in state-based supplemental poverty measure (SPM) research, which requires, among other things, high-quality estimates of means-tested program participation (Isaacs et al., 2012; Levitan et al., 2012; Bohn, Danielson, Levin, Mattingly, and Wimer, 2013).

Several household surveys collect information on participation and benefit amounts for means-tested programs, including TANF and SNAP. However, comparing weighted survey totals to available aggregate administrative totals, it is clear that many people do not accurately report participation and benefits received from these public programs. And this problem of under-reporting has persisted, and for some programs increased, over time and across surveys (Wheaton 2007; Meyer, Mok, and Sullivan 2009). Some programs, particularly SNAP and TANF, have particularly high false negative rates. In our recent work, preliminary estimates yield ACS under-reporting rates of 51 percent for TANF and 34 percent for SNAP in California in 2011 (Bohn et al, 2013).

The ACS differs from other household surveys in several ways that could result in different levels of misreporting of program participation, including the methods of data collection (mainly mail-in) and less detailed questions on public programs. For example, the ACS contains only one question on SNAP participation. In particular, the survey respondent is asked only whether anyone in the household had received Food Stamps or a Food Stamps benefit card in the past 12 months. The survey does not include any information on number of people in the household receiving benefits, length of time in program, nor benefit amounts (as in the CPS and SIPP). Lastly, the ACS household concept and respondent may generate differential under-reporting. For example, for TANF, only persons age 15 or older are asked about TANF and General Assistance cash aid, and the response is likely to include any aid going to children.

While particularly problematic for the SPM research (see Bohn et al., 2013), under-reporting of SNAP and TANF presents hurdles for researchers in a variety of settings. We have obtained custom tabulations of administrative data that permit a robust evaluation of under-reporting in the California sample of the ACS. Given its size and diversity, results for California may be more generalizable than similar evaluations in other states.

Methodological Approach

Two general methods have been used to assess the degree of misreporting of program participation and benefit receipt. One relies on comparisons between survey reported information and aggregate administrative totals, while the other uses one-to-one matches between survey and administrative data at the individual or household level. Meyer and Goerge (2011) examine program participation in the SNAP/FSP for two states (Illinois and Maryland) using one-to-one matches between state administrative records and two household surveys (ACS and CPS) at the individual/household level. They find large levels of under-reporting in both surveys (35% in ACS and 50% in CPS) and also identify household characteristics correlated with under-reporting including age and race/ethnicity.

Meyer, Mok & Sullivan (2009) provide a comprehensive examination of reporting rates across ten programs and five surveys using aggregate administrative totals as the comparison. Their results show consistent and increasing under-reporting for several programs, including TANF and SNAP, across several surveys. They suggest that refinements of administrative aggregates based on demographic characteristics and regions could form the basis for adjustments to correct for under-reporting. Some studies have indeed found evidence that household type and race/ethnicity of householder are correlated with misreporting. (Meyer and Goerge, 2011; Kirlin et al, 2013)

Our approach is a hybrid of those described above. While we do not have individually matched data, we have detailed administrative data for various subgroups of California SNAP and TANF participants. In particular, we can match participants based on county, racial/ethnic group, age, household composition, and joint program participation. This detail is an improvement over aggregate totals provided in most administrative reporting. Furthermore, our custom administrative data has been constructed to match as closely as possible to the SNAP and TANF questions in the ACS. In particular, we count actual program participants based on any participation during the year and length of time on aid. We also adjust administrative counts to account for participants moving across the state within the study period and/or changing household composition.

This paper will examine the rate of misreporting in the ACS within detailed demographic subgroups, as determined by the administrative data. We will model the relationship between misreporting of program participation and of benefit amounts as a function of respondent and reported household characteristics, including race, county of residence, family composition, and employment. The parameter estimates will yield some evidence of differential propensity to underreport. Finally, we will take the evidence from both of these exercises and assess the extent to which we can apply an under-reporting correction made using mean counts of monthly participants in SNAP and in TANF—entirely publicly available data—and obtain similar results.

Data and Preliminary Results

Our key data source consists of custom tabulations from the longitudinal statewide administrative database for California that records monthly receipt of SNAP and TANF for individuals. In other words, these are not publicly available data, but neither do we have access to administrative records that are individually matched to survey records. This database, known as the Medi-Cal Eligibility Determination System (MEDS), does not contain the dollar amounts received, only whether an individual participated in the program. We aggregate these counts to cells defined by characteristics of the unit (number of adults, number of children, county, race, etc.). We use these data to create a distribution of months on aid over each year, as well an unduplicated count of persons and units ever on the program. These tabulations were created in collaboration with the California Department of Social Services.

In the corresponding ACS data, we take pains to split up Census households into units that adhere as closely as possible to the definitions used for program eligibility determination. This involves a number of judgment calls with regard to relationships between individuals in the ACS. For example, an adult sibling who lives with a mother-child dyad would not be required to apply for SNAP together with the others if the three do not generally prepare and eat meals together. Employing the convention used by other state SPM researchers, we split households into the maximum number of units possible according to program rules (Isaacs et al., 2011a; NYC Center for Economic Activity, 2012). Essentially, we keep nuclear families intact, but move related and unrelated adults into their own units (along with any of their children). In addition, we move foster children into single person units and assign them SNAP receipt. TANF units are assumed to be the same as SNAP units in preliminary work, a decision we plan to test further. We also plan to further test the convention of maximum unit creation. We then define SNAP and TANF receipt according to self-reported information within these redefined units.

Known limitations include the fact that with these data we only detect net under-reporting, and cannot disentangle under-reporting separately from over-reporting. However, previous research using one-to-one matches indicates that false negatives are a much bigger issue than false positives (Meyer and Goerge, 2011). Second, we cannot directly investigate the sources of under-reporting listed above by, for example, comparing response rates for subgroups defined by duration of program receipt. However, we may be able to address this particular source of under-reporting by exploiting our administrative counts of time on aid.

The left-hand panel of Table 1 presents preliminary tabulations of ACS data and administrative data across race/ethnicity, one demographic group of interest. We find overall statewide under-reporting of SNAP participation of 34 percent, but differential rates across racial/ethnic group, ranging from 25 percent to 52 percent. Similarly, the right-hand panel of Table 2 presents the same tabulations for ten large counties in California, further highlighting the variation in under-reporting and the improvements in adjustment possible with more refined aggregations of administrative data.

The right-hand panels of Tables 1 and 2 show the adjustment we make in the ACS to correct for underreporting, which reduces the under- (or over-) count in the aggregate to below 5 percent in almost all cases. We will explore other demographic dimensions similarly. We will then examine all dimensions simultaneously in a multivariate, regression framework. With both raw under-reporting counts and correlational estimates in hand, we will then assess the quality of adjustments of under-reporting using pure, publicly available data in California. These data consist of mean caseload counts for each county in the state. Finally, we will conduct the identical set of analyses for the TANF program. Taken as a whole, our final results will be informative for the general effort to correct for survey under-reporting and specifically for the creation of state-level SPMs.

Table 1. SNAP under-reporting in ACS before and after adjustment by race/ethnicity, California (2011)

		No Adjustment			Under-reporting adjustment			
	Admin totals (MEDS)	Self-reported SNAP receipt, unweighted	Self-reported SNAP receipt, weighted	Percent under-reporting, no adjustment	Sample assigned SNAP receipt, unweighted	Adjusted SNAP receipt sample, unweighted	Adjusted SNAP receipt total,	Percent under/over reporting after adjustment
White	657,362	3,965	431,291	-34.4%	1,940	5,905	649,127	-1.3%
Hispanic	1,117,547	7,225	834,672	-25.3%	2,576	9,801	1,136,141	1.7%
Black	400,214	1,578	191,515	-52.1%	1,442	3,020	374,891	-6.3%
Other	294,681	1,701	166,793	-43.4%	1,265	2,966	295,380	0.2%
California	2,469,804	14,469	1,624,271	-34.2%	7,223	21,692	2,455,539	-0.6%

Table 2. SNAP under-reporting in ACS before and after adjustment by county, California (2011)

	Admin totals (MEDS)	No Adjustment			Preliminary Adjustment			
		Self-reported SNAP receipt, unweighted	Self-reported SNAP receipt, weighted	Percent under-reporting, no adjustment	Sample assigned SNAP receipt, unweighted	Adjusted SNAP receipt sample, unweighted	Adjusted SNAP receipt total, weighted	Percent under/over reporting after adjustment
Los Angeles	706,972	4,075	415,110	-41.3%	2,547	6,622	685,687	-3.0%
San Diego	152,950	897	94,717	-38.1%	493	1,390	150,464	-1.6%
Riverside	142,604	934	108,465	-23.9%	317	1,251	143,079	0.3%
San Bernardino	192,150	831	120,535	-37.3%	407	1,238	186,046	-3.2%
Orange	122,463	842	89,089	-27.3%	314	1,156	122,759	0.2%
Sacramento	125,141	722	85,492	-31.7%	269	991	119,577	-4.4%
Fresno	116,051	722	79,359	-31.6%	239	961	110,395	-4.9%
Alameda	85,042	458	53,545	-37.0%	258	716	81,674	-4.0%
Kern	77,033	426	51,689	-32.9%	180	606	74,287	-3.6%
Santa Clara	66,631	429	46,538	-30.2%	174	603	69,237	3.9%

NOTE: The table displays results for ten counties with the largest SNAP caseload totals. All counties and county-groups identifiable in the ACS (41 total) are calculated in the identical manner.

References

Bohn, Sarah, Caroline Danielson, Matthew Levin, Marybeth Mattingly, and Christopher Wimer. 2013. "The California Poverty Measure: A New Look at the Safety Net", Public Policy Institute of California.

Isaacs, Julia, Joanna Young Marks, Katherine Thornton, and Timothy Smeeding. 2011. "The New Demography of Poverty: The Wisconsin Poverty Measure and Effects of Federal and State Policies in Wisconsin." Working paper prepared for Population Association of America annual meetings.

Kirlin, Mirel, Schroeder, and Wiseman. 2013. "Misreporting of Food Stamp Participation in a Household Survey: Results from a single state pilot study of the National Health and Nutrition Examination Survey."

Meyer and Goerge. 2011. "Errors in Survey Reporting and Imputation and their Effects on Estimates of Food Stamp Participation." CES Working Paper.

Meyer, Mok, and Sullivan. 2009. "The Underreporting of Transfers in Household Surveys: Its Nature and Consequences." NBER Working Paper.

NYC Center for Economic Opportunity. 2012. "The CEO Poverty Measure, 2005-2010." Working paper.

Wheaton, Laura. 2007. "Underreporting of Means-Tested Transfer Programs in the CPS and SIPP." 2007 Proceedings of the American Statistical Association, Social Statistics Section [CD-ROM], Alexandria, VA: American Statistical Association: 3622-3629.