

JANUARY 2016 | NUMBER 42

Using Linked Survey and Administrative Data to Better Measure Income

Implications for Poverty, Program Effectiveness, and Holes in the Safety Net

BY BRUCE D. MEYER, UNIVERSITY OF CHICAGO AND NATIONAL BUREAU OF ECONOMIC RESEARCH;
AND NIKOLAS MITTAG, CERGE-EI/CHARLES UNIVERSITY

Survey data are used for many purposes and have become one of the most important sources of information for policymakers and researchers. A large share of the empirical research in economics and other social sciences relies on survey data as underlined by the hundreds of thousands of citations to surveys such as the Current Population Survey (CPS). Additionally, many of the statistics that are frequently used to design and evaluate policies at both the national and state level, such as the rates of unemployment and health insurance coverage, rely on household survey data. The CPS is the source of these statistics, as well as official income distribution and poverty statistics. The survey is also extensively used to determine the effects of transfer programs on the income distribution, program participation rates, and the extent to which individuals are missed by specific programs or missed by the safety net entirely. However, the usefulness of the information in the CPS and other household surveys depends on its accuracy, which unfortunately has been declining. In our research, we link the CPS data to admin-

istrative records from New York State to examine the implications of survey errors, particularly inaccurate welfare program reporting and imputation, for key income-based statistics. We find that inaccurate survey data badly distort our understanding of the income distribution, poverty, and the effects of government programs.

We focus on the implications of misreporting and imputation errors for three types of questions that the government and researchers try to answer using survey data such as the Current Population Survey. First, a large literature examines measures of hardship and the distribution of household income among those with few resources. These statistics supply us with vital information on the prevalence and extent of material deprivation among the worst-off in the population. Most well-known is the annual official income and poverty report; the official poverty rate is also one of the most-cited government statistics in the popular press. Many other scholars have used these data to calculate poverty or income distribution measures at the bottom.

A second prototypical question asks how the addition of the income from specific programs alters the poverty rate or other measures of material deprivation. These calculations provide estimates of the poverty-reducing effects of policies and identify which types of individuals benefit. Such analyses for more than a dozen government programs can be seen in the annual Supplemental Poverty Measure report, the 2015 version being the U.S. Census.

A third important question is who is missed by transfer programs. This information may point to failings of the safety net to reach many of those it is intended to help. Earlier work on this issue focusses on the share of single mothers who neither work nor receive income from government transfer programs. These papers often conclude that a large share of single mothers is missed by our safety net. In each of these prototypical cases, we find that survey errors, mainly the misreporting of government transfer receipt and amounts, but also nonresponse and inaccurate imputation, lead to a greatly distorted view of the situation of those with the fewest resources and the effects of transfer programs.

While the problems of measurement error and nonresponse are not new, they are two of several characteristics of household surveys that have been getting worse over time. Fewer households respond to interviewers (unit nonresponse), and fewer who respond agree to answer income questions (item nonresponse). Item nonresponse rates have been rising over the past 25 years and are on the order of 20 to 30 percent, or higher, for both earnings and government transfers. For some transfer programs, imputed dollars account for 24 to 36 percent of total dollars received in the CPS in 2012, and this nonresponse particularly affects measures of poverty and the tails of the income distribution, which is our population of interest. And even when these households do respond, they are less likely to give accurate answers.

In light of these problems, some researchers have questioned the accuracy of income data for the poor and suggested that consumption data would provide a better benchmark, suggesting that a major source of the recent large discrepancy between income and consumption measures of poverty and low incomes is likely to be the underreporting of transfer income. While an issue for many other variables, the measurement-error problem is particularly severe for transfer programs, with receipt missed for over one-third of housing assistance recipients, 40 percent of Supplemental Nutrition Assistance program (SNAP) recipients, and 60 percent of Temporary

Assistance for Needy Families (TANF) and General Assistance recipients. Even among those who correctly report receipt, average amounts received in the CPS fall short of the true amounts by 6 percent for SNAP, 40 percent for TANF and General Assistance, and 74 percent for housing assistance. A few studies attempt to correct for program underreporting, but most do not. While past evidence on this issue has been based on a mix of aggregate and linked microdata, here we are able to calculate directly the effects of misreporting.

A major difficulty in evaluating the extent of survey errors and their consequences is that one needs an external measure of truth to compare to survey responses. Some previous studies have used reinterviews, information from other surveys, or administrative records to validate survey responses. In our work, we replace survey responses on the receipt and amount of government transfers with administrative records for four income transfer programs linked to the survey. The administrative records are extremely accurate (they contain actual payments made; they are validated by the agency; and definitions are comparable to survey definitions). The administrative data are linked to survey data at the individual level with a high match rate because validated social security numbers are required to receive three of the programs.

Overall, correcting for misreporting sharply changes key results from survey data. Using the administrative variables, poverty and inequality are lower than officially reported, program effects are larger, and fewer individuals have fallen through the safety net. Incomes below the poverty line, particularly below half the poverty line, are substantially understated in the CPS. While the underreporting as a share of income becomes smaller as income rises, substantial dollars are missed even toward the middle of the income distribution. Throughout the distribution, correcting for underreporting makes a larger difference to household resources than including reported noncash benefits. Underreporting of transfer receipt also makes government anti-poverty policies appear much less effective: in the corrected data the poverty-reducing effect of all programs combined is nearly doubled, while the effect of housing assistance is tripled. Both the understatement of household income and the poverty-reducing effect in the survey are even more pronounced for some subpopulations that are at particular risk of deprivation.

The understatement is particularly large for single mothers: correcting for survey errors increases their

overall poverty reduction due to the four programs by 11 percentage points, amplifying the poverty-reducing effect of public assistance more than sixfold and that of housing assistance more than tenfold. In addition, we find that the fraction of nonworking single mothers missed by government transfers is much lower than previously reported. This underlines that the coverage of the safety net is better than the survey suggests. While our specific results pertain to one survey (the CPS) in one state (New York), it is very likely that our findings are more general. New York is a large state and demographically similar to the United States, and while the safety net in New York is more extensive than in other states, previous studies have found misreporting rates to be lower in New York than

in other states. The extent of misreporting of transfer programs is higher in the CPS than in other surveys, but none of the major surveys is without substantial measurement error in program receipt. Thus, our results suggest substantial biases to similar analyses in alternative datasets and other states.

NOTE

This research brief is based on Bruce Meyer and Nikolas Mittag, “Using Linked Survey and Administrative Data to Better Measure Income: Implications for Poverty, Program Effectiveness and Holes in the Safety Net,” National Bureau of Economic Research Working Paper no. 21676, October 2015.
