Errors in Reporting and Imputation of Government Benefits and Their Implications
Pablo Celhay, Pontificia Universidad Católica de Chile
Bruce D. Meyer, University of Chicago and NBER
Nikolas Mittag, CERGE-EI

October 31, 2017

Abstract
We document the extent, nature, and consequences of survey errors for receipt of cash welfare
and food stamps in three major U.S. household surveys linked to administrative program records. Our
results confirm earlier evidence of frequent misreporting of program receipt, particularly failure to report
receipt. The surveys poorly capture patterns of participation in multiple programs, even though there is
not much evidence of program confusion. Error rates are larger among imputed observations, which also
account for a large share of false positive errors. Many household characteristics have significant effects
on errors in reporting receipt, both false positives and false negatives. Among others, we find large
differences in survey error by race, ethnicity, and income. We provide evidence on the consequences of
these errors for models of program receipt. Estimated effects of, among others, income and race are
biased downward. On the positive side, our results confirm a tendency to attenuation in models with
misclassified dependent variables that preserves many qualitative results in the survey estimates. We
then examine error due to item non-response and imputation, as well as whether imputation improves
estimates. Item non-respondents have higher receipt rates than the population, even conditional on many
covariates. The joint distribution of program receipt and these covariates differs between item non-
respondents and the entire population, but also between accurate and imputed receipt among item non-
respondents. Thus, neither excluding item non-respondents nor using their imputed values yields
consistent estimates. For a specific model, binary choice models of program receipt, estimates from the
linked data favor excluding item non-respondents over using their imputed values. The biases in each case
are well predicted by the error patterns we document, so such analyses can help researchers make more
informed decisions on the use of imputed values.

Key words: Linked Survey Data, Measurement Error, item non-response, imputation, program receipt.

JEL Codes: C81, D31, I32, I38

Any opinions and conclusions expressed here are those of the authors and do not necessarily represent the views
of the New York Office of Temporary and Disability Assistance (OTDA) or the U.S. Census Bureau. We thank
Michael Davern and the audience at the Joint Statistical meetings for their comments. Mittag is thankful for
financial support from the Czech Academy of Sciences (through institutional support RVO 67985998) and the Czech
Science Foundation (grant no. 16-07603Y). The data analysis was conducted at the Chicago Census Research Data
Center by researchers with Special Sworn Status and the results were reviewed to prevent the disclosure of
confidential information. Celhay: School of Government, Pontificia Universidad Católica de Chile, Avda. Vicuña
Mackenna 4860, Macul, Santiago, Chile, pacelhay@uc.cl. Meyer: Harris School of Public Policy Studies, University
of Chicago, 1155 E. 60th Street, Chicago, IL 60637, bdmeyer@uchicago.edu Mittag: CERGE-EI, joint workplace of
Charles University and the Economics Institute of the Czech Academy of Sciences, Politických vězňů 7, Prague 1,
110 00, Czech Republic. nikolasmittag@posteo.de