Correcting for Misreporting of Government Benefits

Author(s): Nikolas Mittag

Source: *American Economic Journal: Economic Policy*, May 2019, Vol. 11, No. 2 (May 2019), pp. 142-164

Published by: American Economic Association

Stable URL: https://www.jstor.org/stable/10.2307/26641368

American Economic Journal: Economic Policy 2019, 11(2): 142–164
https://doi.org/10.1257/pol.20160618

Correcting for Misreporting of Government Benefits†

By Nikolas Mittag*

Data linkage studies often document, but do not remedy, severe survey errors. To improve survey estimates despite restricted linked data access, this paper develops a convenient and general estimation method that combines public use data with conditional distribution parameters estimated from linked data. Analyses using linked SNAP data show that this method sharply improves estimates and consistently outperforms corrections that mainly rely on survey data. Yet, some univariate corrections perform well when linked data do not exist. For SNAP, extrapolating from linked data across time and geography still improves upon estimates using survey data only, even after survey-based corrections. (*JEL* C81, C83, H75, I18, I38)

Household surveys are crucial for many academic and policy analyses. Official estimates of employment and poverty rates (US Census Bureau, 2015) are based on the Current Population Survey (CPS). The Congressional Budget Office uses survey data to score legislation and document the distributional impacts of taxes and transfers (e.g., Congressional Budget Office 2013, 2016). The American Community Survey (ACS) used here is an important source of information on sub-state areas. Unfortunately, government benefits are severely underreported in many surveys (Meyer, Mok, and Sullivan 2015). A growing literature validates survey data by linkage to administrative records, which shows that survey errors

* Mittag: CERGE-EI, joint workplace of Charles University Prague and the Economics Institute of the Academy of Sciences of the Czech Republic, Politických vězňů 7, Prague 1, 111 21, Czech Republic (email: nikolasmittag@posteo.de). Matthew Shapiro was editor for this article. This research was supported by the US Census Bureau, the Population Center at NORC, the Czech Science Foundation (16-07603Y), Charles University (UNCE/HUM/035), and the Czech Academy of Sciences (RVO 67985998). A previous version was titled "A Method of Correction Applied to the Food Stamp Program." Any opinions and conclusions are mine and do not necessarily represent the views of the US Census Bureau, the New York OTDA, or NORC. While this work was not subject to formal Census Bureau content review, Census staff reviewed all statistical output to ensure that no confidential information was disclosed. Information presented here is derived in part from the Transfer Income Model, Version 3 (TRIM) and associated databases. TRIM requires users to input assumptions and/or interpretations about economic behavior and the rules governing federal programs. Therefore, the conclusions presented here are attributable only to the authors of this report. I am grateful for the assistance of OTDA, particularly Dave Dlugolecki and George Falco, and many Census Bureau employees including Mike Berning, Frank Limehouse, Graton Gathright, Joey Morales, and Amy O'Hara for their assistance with the data. I would like to thank Robert Moffitt for help with implementation of the SMC method and Lupe Bedoya, Dan Black, Kerwin Charles, Pavel Čižek, Ryan Gallagher, Patrick Gaulé, Štěpán Jurajda, Jan Kabátek, Bruce Meyer, Laura Wheaton, and seminar participants at Bilkent University, the US Census Bureau, CERGE-EI, EALE, the ES Summer Meeting, the Illinois Economic Association, IAAE, JSM, the Harris School of Public Policy, RES, Sabanci University, SUNY Albany, Tilburg University, UCSB, and the University of Alicante for comments. Suggestions from two anonymous referees substantially improved this article. Marin Drlje, Sam Mayr, and Derek Wu provided excellent research assistance. All errors are mine.
† Go to https://doi.org/10.1257/pol.20160618 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

severely bias common analyses including measures of poverty and the distribution of income, as well as analyses of program targeting and effects (Meyer and Mittag 2019; Bee and Mitchell 2018). Directly substituting survey reports with accurate administrative measures would improve estimates, but is currently not feasible for two reasons. First, linked data are rarely accessible, usually due to confidentiality rules. Second, linked data are often based on a convenience sample for which a measure of truth is available, such as one state. Consequently, linked data are rarely used to answer substantive questions. Instead, most applied studies and official statistics rely on either survey or administrative data only.

In this paper, I examine how contaminated survey data can still be used to study transfer programs, and how information from the linked data can improve survey estimates. I first explore whether information from the linked data that can be disclosed to the public can improve estimates. To do so, I develop a two-step method that allows researchers without access to linked data to obtain consistent estimates, including regression coefficients, from public use data using simulations based on linked data. In the first step, someone with access to the linked data estimates and publishes the distribution of the accurate measures conditional on reports and other covariates. In the second step, researchers without linked data access use the estimated distribution to back out parameter estimates from the public use data. This conditional distribution can also be used to add missing variables to surveys and to impute data missing due to nonresponse. I use administrative records on the Supplemental Nutrition Assistance Program (SNAP) from New York State (NY) linked to the ACS to show that the method accurately reproduces results from the linked data. The method is simple to implement and can correct a large class of estimators, so it can solve the problem of data access.

The conditional distribution method requires linked data to exist and an estimate of the conditional distribution to be available. Often, at best, similar data have been linked, which raises the question of how to analyze transfer programs without linked data. I use the linked data to evaluate whether using survey information to correct for underreporting or extrapolating from similar linked data improves estimates. Specifically, I first examine whether recently applied corrections for underreporting improve estimates. I evaluate two survey-based approaches, imputations based on program rules and imputations based on models of receipt estimated from survey reports. The results show that carefully applying corrections solely based on survey data can still improve estimates, which makes them attractive alternatives when linked data do not exist. Contrary to corrections using linked data, survey-based corrections do not incorporate information that could separate how true program receipt varies with other variables from how misreporting varies with these variables. Therefore, survey-based corrections at best partially correct the bias and can worsen it, particularly when analyzing multivariate relationships like regressions with mismeasured independent variables.

Second, I study whether extrapolation from similar linked data still improves estimates. Linked data often only exist for a similar population or a different time period, but not for the population of interest. For transfer programs, most current linkage studies only cover one state and a few years. Can we use these linked convenience samples to learn more about the population of interest?

I provide evidence from two sources that extrapolation from NY to the entire United States using the conditional distribution works well for SNAP: First, extrapolation across geography within NY is still more accurate than the survey-based corrections. Second, extrapolation to the entire US reduces average error to official statistics in the geographic distribution of program spending by up to 90 percent. The findings on extrapolation are specific to SNAP, but they suggest that corrections based on linked data can be more accurate than survey-based corrections even if linked data are not available for the population of interest.

Overall, the results show that survey estimates can be improved and that information from linked data can help. The method developed here mitigates two key problems of linked data, access and coverage, that limit the usefulness of current linkage studies. Substantively, the analysis of SNAP in NY shows that survey reports overstate the poverty rate (when including SNAP benefits as income), and miss a large share of program benefits higher up in the distribution of reported annual incomes. Extrapolation suggests similar or larger bias at the national level. For example, the survey reports overstate the poverty rate by about 1 percentage point, about twice the difference I find for NY.

Section I reviews misreporting of government benefits. Section II describes the data. Section III describes the corrections and the conditional distribution method. Section IV evaluates corrections for analyses of SNAP in NY. Section V examines extrapolation. Section VI concludes. Online Appendix A provides additional results mentioned in the paper. Online Appendix B and C further discuss data linkage and the methods used in this paper. Online Appendix D summarizes definitions and abbreviations.

I. Using Data Combination to Address Survey Error

Recent studies demonstrate severe problems of data accuracy in many major household surveys. Survey error severely affects important variables, such as income (Bound and Krueger 1991, Bollinger 1998, Dahl, DeLeire, and Schwabish 2011, Abowd and Stinson 2013), education (Black, Sanders, and Taylor 2003) and employment status (Poterba and Summers 1986, Chua and Fuller 1987). The overview by Bound, Brown, and Mathiowetz (2001) emphasizes that misreporting is pervasive. For government benefits, Meyer, Mok, and Sullivan (2015) demonstrates severe underreporting for many programs in several surveys. Celhay, Meyer, and Mittag (2018a) documents high error rates of SNAP and public assistance receipt in the ACS data used here; 26 percent of SNAP recipients do not report receipt. The rate of non-recipients reporting receipt is much lower at 1.2 percent. Previous linkage studies have found similar error rates in the Survey of Income and Program Participation (SIPP) and in the predecessor of the ACS, as well as even higher rates in the CPS (Marquis and Moore 1990; Meyer, Mittag, and Goerge 2018; Taeuber et al., 2004). All studies find errors to be related to other important variables, such as income and household composition.

This severe and systematic misreporting is a serious problem even for simple survey-based analyses of government programs, income, and poverty (Meyer and Mittag 2019). The effects of measurement error in more complex econometric models are usually intractable. Several studies (e.g., Bollinger and David 1997; Meyer and Mittag 2017; Nguimkeu, Denteh, and Tchernis 2019) document complicated biases that limit or skew what we learn about government programs. Even worse, important variables can be entirely missing from survey data. For example, the ACS does not ask for SNAP amounts and neither the CPS nor the ACS contain reports of housing assistance amounts. This lack of program amounts severely limits studies of poverty measures, such as the supplemental poverty measure, that include non-monetary benefits (e.g., Bohn et al. 2013, Office of the Mayor 2015, Levitan and Renwick 2010).

Accurate measures of the variables of interest are sometimes available from other sources, such as administrative records. Yet, many questions cannot be answered with administrative data only, because they lack the required covariates. Administrative records are also rarely representative of the population of interest. For example, records from transfer programs do not include non-recipients and therefore cannot be used to analyze program take-up. Linking survey data to administrative records combines the advantages of both data sources by adding the accurate administrative measure to the detail of the survey. Thereby data linkage can address both the problem of missing and mismeasured variables.

II. Data Sources and Linkage

This study uses three data sources: survey data from the ACS and the CPS Annual Social and Economic Supplement, as well as administrative SNAP records from the NY Office of Temporary and Disability Assistance (OTDA). I use the CPS to assess the accuracy of analyses that require benefit amounts because the ACS does not contain amounts received. The administrative records contain all monthly SNAP payments in NY from 2007 through 2010 for every individual on a SNAP case. Individual identifiers have been checked by OTDA against social security records. The records appear to be accurate; their overall total differs from actual aggregate outlays by less than 1 percent in each year.

The administrative and survey data are linked using the Person Identification Validation System (PVS) of the US Census Bureau (Wagner and Layne, 2014). In short, the PVS uses personal data including address, name, gender, and date of birth to find a matching record in a reference file that contains all transactions recorded against a social security number. If a matching record is found, the anonymized social security number (the PIK) is attached to the corresponding observation. Matches are made based on this PIK, which is obtained for 99 percent of the administrative records. The data are linked at the individual level, but all analyses here are at the household level. I consider a household to receive SNAP if any member received SNAP in the reference period of the ACS according to the administrative records. More than 94 percent of the surveyed households contain at least one PIKed member and are therefore likely to be classified correctly. Following Meyer, Mittag, and Goerge (2018), I account for incomplete linkage by multiplying the weights by the inverse of the predicted probability of any household member having a PIK (Wooldridge 2007). Online Appendix Table A1 reports the parameter estimates used in this correction. Online Appendix B provides additional information on data sources and linkage. Meyer and Mittag (2019) and Celhay, Meyer, and Mittag (2018a,b) discuss data accuracy further. Summary statistics are in Tables A2–A4.

Errors in the administrative records and linkage errors likely cause inaccuracies in the linked data. Given the high-match rate and the quality of the administrative records, such errors should be negligible compared to survey error here. Therefore, I can evaluate corrections for misreporting by comparing estimates to those from the linked data. Assuming that the linked data are accurate enough to do so may not be warranted with other linked administrative records.

While linked data can solve many data issues, they face two key problems (Bound, Brown, and Mathiowetz 2001). First, the data are usually only available to a few researchers due to confidentiality and access rules. Second, linked data rarely exist for the population of interest. They are often only available for a few years or for a small area. Studies at the national level or of other years need to either extrapolate from existing linked data or rely on survey reports. The method in this paper mitigates these problems because it does not require direct access to the linked data and does well at extrapolation.

III. Methods of Correcting for Misreporting

Despite its well-known extent, few studies attempt to correct for misreporting. Simple statistics can be corrected by scaling survey estimates up by the ratio of official and survey totals. This benchmarking does not work for many parameters, such as correlations or regression coefficients, and assumes that misreporting is not related to any covariates. Therefore, I focus on methods that adjust the microdata. I first describe two corrections that impute transfer receipt based on information in the survey data. Then, I discuss the conditional distribution method that I use to incorporate information from the linked data. Detailed instructions on the implementation of each method are provided in online Appendix C.

A. *Survey-Based Corrections*

One approach to impute additional recipients uses program rules to assign receipt and amounts received to some eligible households that do not report receipt. The TRIM model uses the CPS data and microsimulations to determine eligibility based on program rules. TRIM assigns receipt among those predicted to be eligible, such that the recipient population matches USDA quality control data. See Zedlewski and Giannarelli (2015) for a description of TRIM, which is commonly used in academic and policy research (see e.g., Sherman 2009; Ginnarelli, Wheaton, and Morton 2015; Sherman and Trisi 2015). One problem with this approach is that program rules determine eligibility, but not all eligible individuals take up benefits, which leads to overimputation unless one assigns receipt to some eligibles only. Also, predicting eligibility from survey reports of the determinants of eligibility is very noisy (Newman and Scherpf 2013). Often many true recipients appear ineligible, likely because key determinants of eligibility, such as monthly income, are missing or misreported in the survey.

The second approach (Scholz, Moffitt, and Cowan 2009; Moffitt and Scholz 2010; Ben-Shalom, Moffitt, and Scholz 2012), imputes additional program receipt using models estimated from the survey data. Specifically, Scholz, Moffitt, and Cowan (2009, 218–219) uses SIPP survey data to estimate a probit model of program receipt. They then assign receipt to the households with the highest predicted probabilities of receipt until the number of recipients matches administrative aggregates. They predict amounts received for the imputed recipients from a regression of amounts on demographics among those reporting receipt and add a randomly drawn residual. Finally, they scale up amounts of all recipients to match total program spending. I adapt the approach of Scholz, Moffitt, and Cowan (2009) to the ACS and refer to it as the SMC method. The ACS does not include benefit amounts, so I impute amounts for all recipients based on regressions in the CPS. The ACS and the CPS are representative of the same populations and the covariates are comparable.

The SMC method assigns receipt to the most likely recipients, which leads to overimputation for likely recipients (such as the poor) and underimputation for less likely recipients. To avoid this problem, I also implement a modified SMC method that assigns receipt probabilistically according to the predicted probabilities, until the number of recipients matches administrative aggregates.[1] A more difficult problem is that the parameter estimates used to predict the probability of receipt (and amounts) are biased because they are estimated using the misreported data. Kimball, Sahm, and Shapiro (2008) demonstrates and analyzes this problem for a similar imputation method. In consequence, the estimated model will tend to overpredict receipt by those likely to report receipt, while still predicting too low receipt rates for those unlikely to report receipt. Thereby, one continues to underestimate receipt rates among respondents or demographic groups with low reporting rates. At the same time, overpredicting receipt for households that report more accurately further overstates receipt by demographic groups with high reporting rates. Therefore, the imputed data will preserve or even amplify differences in receipt rates between groups with different reporting rates. Thereby, it will preserve or amplify the bias in estimated (partial) correlations of program receipt with variables that predict misreporting.

B. *Conditional Distribution Method*

The corrections above are unlikely to yield consistent estimates because they do not incorporate information that could separate how true program receipt varies with other variables from how misreporting varies with these variables. Many variables predict *both* receipt and misreporting (Meyer, Mittag, and Goerge 2018). Using information on how misreporting varies with covariates can remove this bias. Such information can be obtained from linked data. A convenient way to enable researchers to correct estimates is to provide an estimate of the distribution of the accurate measures conditional on survey reports and other covariates.

[1] A recent implementation of this approach also imputes probabilistically (Moffitt and Pauley 2018).

Consider a model including variables only available in the linked data X^A and (potentially) other covariates Z. In the application below, X^A includes SNAP receipt and amounts received and Z contains variables such as income. The public use data only contain Z and error-prone reports X^R of some or all administrative variables. Let $f_{X^R,Z}$ be their joint distribution. For the ACS, X^R is reported receipt only, because amounts are not reported. The linked data contain $\left(X^A, X^R, Z\right)$, so they can be used to estimate the model of interest.

The key idea is that the conditional distribution of X^A given X^R and Z, $f_{X^A|X^R,Z}$, is enough to back out the parameters from the public use data. The objective function of estimators containing X^A can be expressed in terms of $\left(X^R, Z\right)$ and $f_{X^A|X^R,Z}$ using the law of iterated expectations (Hsiao 1989; Chen, Hong, and Tamer 2005). Consider a parameter θ defined by the moment condition:

$$E\left[m\left(X^A, Z, \theta\right)\right] = 0. \tag{1}$$

This setup covers most common estimators and allows for mismeasured dependent and independent variables. If θ is the mean of X^A, then $m\left(x^A, z, \theta\right) = x^A - \theta$. Note, θ could also be the coefficient vector from regressing X^A on Z, so $m\left(x^A, z, \theta\right) = z'\left(x^A - z\theta\right)$, or from a maximum likelihood model, so $m(\cdot) = \partial\ell/\partial\theta$ are the derivatives of the log likelihood function. The function $m(\cdot)$ depends on X^A, so one cannot use it to estimate θ from the public use data. Nevertheless, the law of iterated expectations implies moment conditions in terms of the observed $\left(X^R, Z\right)$ that are uniquely solved by θ. The conditional expectation of $m(\cdot)$ at given values of the survey reports and the covariates is $\tilde{m}\left(x^R, z, \theta\right) = E\left[m\left(X^A, z, \theta\right) | X^R = x^R, Z = z\right] = \int m\left(x^A, z, \theta\right) f_{X^A|X^R,Z}\left(x^A, x^R, z\right) dx^A$. Taking the expectation over the public use sample yields a modified moment condition:

$$E\left[\tilde{m}\left(X^R, Z, \theta\right)\right] = \int\int \tilde{m}\left(x^R, z, \theta\right) f_{X^R,Z}\left(x^R, z\right) dx^R dz = 0. \tag{2}$$

By the law of iterated expectation and (1), this moment condition is uniquely solved by the true value of θ. Equation (2) only depends on variables in the public use data and the conditional distribution $f_{X^A|X^R,Z}$.

The method in this paper replaces $f_{X^A|X^R,Z}$ by a parametric estimate from the linked data and then simulates the integral in (2) to estimate θ. The advantages of estimation in two steps are practical: it solves the problem of access and is simple to implement.[2] In the first step, a researcher with access to the linked data estimates the conditional distribution. I use a parametric estimate of the conditional distribution for practical reasons discussed further below. Most importantly, a parametric estimate mitigates disclosure problems because it is a vector of parameters that can be published without infringing confidentiality.

[2] Programs are available from my website (currently http://home.cerge-ei.cz/mittag/).

In the second step, which does not require access to the linked data, researchers use this estimated conditional distribution and the public use data to estimate their model of interest. As equation (2) shows, the parameters can be estimated by using the estimated conditional distribution to integrate X^A out of the objective function of the estimator. Solving the integral can be computationally burdensome, but is simple by simulation. The researcher generates D draws from $\hat{f}_{X^A|X^R,Z}\left(X^A, X^R = x_i^R, Z = z_i\right)$ for every observation $i = 1, \ldots, N$ in the public use data, which yields $D \cdot N$ observations of $\left(X_{id}^A, X_i^R, Z_i\right)$. Consistent parameter estimates can be obtained from this dataset as if X^A were available, i.e., by solving or optimizing the original objective function using the expanded data.

This strategy rests on two key assumptions. First, it requires the conditional distribution $f_{X^A|X^R,Z}$ to be the same in the linked data and the public use data. This assumption is trivially true if the two samples are the same, but may not hold if they differ. Second, the estimate of the conditional distribution needs to be consistent, which may not be the case if the parametric assumptions do not hold. Parametric assumptions are testable and can be relaxed as discussed in online Appendix C. Under these assumptions, the simulated sample equivalent of (2) converges to its true expectation as the size of the linked and the public use sample, as well as D, go to infinity. Thereby, the estimator is a standard simulated method of moments estimator, so consistency and asymptotic normality follow from McFadden (1989).

Standard errors need to be adjusted for the estimation of the parameters of the conditional distribution (Newey and McFadden 1994) and simulation error (McFadden 1989). The correction for simulation error can be made arbitrarily small by choosing D to be large. Unfortunately, little can be said about the required number of draws in general. For a specific case, one can simply increase D until the estimates stabilize. In the application below, I use 10 draws to estimate means and 200 draws to estimate probits. Due to the large linked sample, the correction for the estimation of the parameters of the conditional distribution is negligible in the application below. The correction likely matters when the parameters of the conditional distribution are less precisely estimated. When the linked sample is small relative to the public use sample, standard errors can remain large even with large public use samples. See online Appendix C for detail.

A key advantage of surveys is that they often collect many covariates. Many variables are known to predict survey errors. The conditional distribution method allows errors to be arbitrarily correlated with the variable of interest as well as the covariates Z. Thereby, it can accommodate covariates related to both the true and the reported values of the variable of interest. To do so, the conditional distribution has to condition on all variables in the outcome model that are not independent of X^A (conditional on X^R and the remainder of Z). Otherwise the estimator suffers from the imputation bias discussed in Bollinger and Hirsch (2006) and Hirsch and Schumacher (2004). Researchers who estimate the conditional distribution in the linked data often do not know which variables the outcome models contain. Consequently, it is important to condition on many covariates, which is an advantage of a parametric conditional distribution.[3]

Note that X^A can be a matrix, so the idea extends to multiple mismeasured variables, which requires estimating a joint conditional distribution. If one is willing to assume a parametric joint distribution, so that estimation remains $\sqrt{N}$-consistent, the extension is straightforward. It only requires simulating D draws of multiple variables from this joint distribution. The parametric assumptions can be relaxed by estimating the joint distribution semi-parametrically, or by estimating multiple conditional distributions. The convergence rate of the estimator will be the rate of the conditional distribution, which quickly decreases with more variables. So the idea easily extends to the multivariate case, but whether it works well in practice for more than one or two variables is an open question.

The literature on measurement error contains several related estimators. For linear regressions, the estimator simulates conditional mean imputation (Schafer and Schenker 2000). For binary choice models with a misclassified dependent variable, the estimator is a simulated version of the estimator in Bollinger and David (1997). Chen, Hong, and Tamer (2005); Hsiao (1989) and Ichimura and Martinez-Sanchis (2009) develop similar semi-parametric one-step estimators. If the researcher has access to the linked data, the advantages of joint and semi-parametric estimators likely make these estimators preferable. The conditional distribution method can also be seen as a multiple imputation method (Rubin 1987, 1996). A key difference is that I estimate the outcome model from the data that stacks the D draws, rather than averaging estimates based on one draw each, which makes the estimator consistent for statistics that require *both* N and D to go to infinity. See online Appendix C for further discussion.

The key advantage of the conditional distribution method over the survey-based methods above is the source of the imputation model. Using information on how misreporting is related to other variables allows the conditional distribution method to disentangle variation in X^A and variation in misreporting. The information from linked data can also reduce the variance of the imputation error, yielding a more precise estimate. Even so, this at best improves estimates to the same extent as using the linked data. If the assumptions of the conditional distribution method hold, it estimates the same parameter as the linked data. Any bias from errors in the linked data also affects estimates from the conditional distribution method. Nevertheless, given the extent of survey error, even estimates from linked data with some error are often an improvement over the survey estimates.

To illustrate implementation, consider the case of transfer receipt. Online Appendix C provides additional details. The first step requires access to the

[3] As in standard regressions, conditioning on irrelevant variables does not cause bias, but conditioning on variables that are endogenous, i.e., predict the error term of the outcome model, can cause bias. In practice, this can often be avoided by including these variables in the outcome model as well. See Steuerle-Schofield et al. (2015) for a discussion of the consequences of conditioning on the dependent variable. For the purpose of extrapolating to other surveys, one may (also) want to estimate a more parsimonious conditional distribution to make sure comparable variables are available. Both problems can, in principle, be solved without access to the linked data by integrating variables out of the estimated conditional distribution.

linked data. They contain $\left(X^A, X^R, Z\right)$, so estimating the conditional distribution by maximum likelihood is straightforward. Consider estimating the conditional distribution of SNAP receipt and amounts received that I use in the application below. Estimation can be simplified to a univariate problem. I estimate the distribution of SNAP amounts conditional on reported receipt and other covariates, allowing for a mass point at zero that implies no receipt. I use a probit model for the probability of receiving an amount of zero. For nonzero amounts, a left truncated normal density in which the mean is a function of the covariates fits the data well.[4] Thus, I estimate the following conditional distribution:

$$(3) \qquad f_{X^A|X^R,Z}\left(X^A, X^R = x^R, Z = z\right)$$
$$= \begin{cases} \Phi\left(x^R\alpha + z\beta; 0, 1\right) & \text{if } X^A = 0 \\ \left[1 - \Phi\left(x^R\alpha + z\beta; 0, 1\right)\right] \dfrac{\phi\left(X^A; x^R\gamma + z\delta, \sigma\right)}{1 - \Phi\left(\tau; x^R\gamma + z\delta, \sigma\right)} & \text{if } X^A \neq 0; \end{cases}$$

α, β, γ, δ, σ, and τ are estimated from the data. The parameters of the mass point, α and β, determine the probability that a household does *not* receive SNAP. The parameters of the truncated normal, γ, δ, σ, and τ, determine the distribution of amounts received by participants. The sample includes observations for which SNAP receipt is imputed, so Z contains an indicator for imputation status. Usually, one may want to allow the models of reporting and imputation errors to differ by more than a constant, but only 1 percent of observations in the linked data are imputed. Parameter estimates from the 2009 and 2010 linked ACS are in online Appendix Table A5.[5] The estimates are not surprising: Reported receipt is a strong predictor of receipt. Household composition and income capture most of the remaining variation. The results below are robust to reasonable specification changes.

A parametric distribution makes it simple to incorporate additional information in the second step. For government benefits, total spending and recipients are often available from official statistics. One can adjust the intercepts for the mass point and amounts to make total recipients and amounts match the official numbers in expectation. If the official numbers are available for geographic or demographic subpopulations, such as states, one can also add separate intercepts for these subpopulations to β and δ.

If the estimated parameters of the conditional distribution are made available, other researchers can use them to correct analyses based on the public use data. To illustrate this second step, consider some prototypical analyses. First, transfer programs are often analyzed using statistics such as receipt or poverty rates. After drawing D benefit amounts for each observation from the conditional

[4] Restricting the density to be a truncated normal at every value of the covariates does not force the marginal distribution of SNAP amounts to be a truncated normal. Using a right truncation point, a t-distribution, or mixture models does not improve model fit.

[5] Additional parameter estimates and variance matrices can be downloaded from my website.

distribution, $(D \cdot N)^{-1}\sum_{d=1}^{D}\sum_{i=1}^{N}\mathbf{1}(y_{id} < PL_i)$ consistently estimates the poverty rate. The term PL_i is the poverty line for household i and y_{id} is their income including the simulated program amount from draw d. Second, linear regressions can be estimated by computing the OLS slopes from the simulated data with $N \cdot D$ observations. Third, maximum likelihood models, such as binary choice models of program take-up, can be estimated by maximizing the likelihood function over the simulated draws of the program receipt variable. More generally, obtaining consistent estimates from public use data only requires solving the original estimation problem for a dataset that is D times larger and adjusting standard errors as discussed in online Appendix C.

IV. Reassessing SNAP in New York State

Table 1 shows how descriptive statistics of SNAP in NY vary across data sources and corrections. Comparing the results from ACS Public Use Micro Data (PUMS) and CPS survey reports in columns 2 and 3 to those from the linked data in column 1 confirms that survey reports fail to capture a large fraction of receipt.[6] The first row shows that the CPS misses a third of SNAP dollars in NY, an impressive \$1.4 billion. The receipt rates in row 2 are also higher than the reports suggest, by 29 percent in the ACS and by almost 50 percent in the CPS. Row 3 shows that the survey reports overstate the fraction of recipients in poverty by 15 (ACS) and 45 (CPS) percent. This error skews analyses of program targeting, making the program look more focused on those with very low annual income than it really is. The last row contains the difference in the poverty rate when SNAP is added to the income definition. This simple measure of poverty reduction due to SNAP is 0.5 percentage points (23 percent) higher than what the CPS suggests. In line with Meyer and Mittag (2019), this shows that the survey data understate program effects.

Columns 4 to 8 of Table 1 examine whether and how researchers can improve survey estimates. All corrections make receipt rates and total amounts match administrative totals, so the first two rows are not informative about the performance of the methods.[7] The results from the ACS PUMS data and the estimated conditional distribution in column 4 are virtually identical to those in column 1. Thus, the conditional distribution method can recover the correct estimates from the public use data.

Parameter estimates of the conditional distribution are often not available from current data. Using linked data from previous years (Bollinger and David 1997; Davern, Meyer, and Mittag forthcoming) requires the conditional distribution to remain constant, which is at best an approximation. In column 5, I use the conditional distribution from linked 2009 ACS data to correct for misreporting in the 2010 ACS. I adjust the conditional distribution to make the number of recipient

[6] Due to the large sample sizes, standard errors are too small to affect the results substantively throughout.

[7] There are minor differences: the conditional distribution method matches the numbers in expectation. SMC and TRIM match exactly within sample. The conditional distribution method matches recipient households (and understates recipient individuals). The SMC method matches recipient individuals (and overstates recipient households). TRIM matches slightly different numbers.

Table 1—SNAP Statistics According to Different Data Sources and Corrections, NY 2010

		Survey reports		Conditional distribution		SMC method		
	Linked data (1)	ACS PUMS (2)	CPS (3)	Current par. (4)	Lagged par. adj. (5)	Original (6)	Modified (7)	TRIM CPS (8)
Total amount (in billion $)	4.33	—	2.91	4.29	4.32	4.33	4.33	4.26
Fraction of recipient HH	17.9%	13.8%	11.6%	17.6%	17.8%	18.1%	18.8%	19.6%
Poverty rate among recipients	38.6%	44.3%	55.9%	39.4%	40.1%	46.3%	40.0%	55.3%
Poverty reduction	2.4%	—	1.9%	2.3%	2.3%	2.7%	2.2%	2.8%

Notes: Column 1 uses the administrative measures from the linked internal ACS. Columns 2 and 3 use ACS and CPS survey reports. Column 4 uses the 2010 ACS PUMS and the conditional distribution estimated from the 2010 linked ACS; column 5 uses the estimated distribution from 2009, adjusted to make SNAP dollars and recipient households match administrative numbers in expectation. Column 6 uses the correction of Scholz, Moffitt, and Cowan (2009) adapted to the ACS. Column 7 uses the same approach, but assigns receipt probabilistically. Column 8 uses the (CPS-based) TRIM simulations. All analyses use household weights (adjusted for incomplete linkage in column 1).

households and total amounts received match the 2010 linked data in each group of counties identified in the PUMS data as described in online Appendix C. The performance of extrapolation is application-specific, but the results in column 5 are promising. Extrapolation closely replicates the results from the linked data. Even though the conditional distribution changes over time, a rich set of conditioning variables and incorporating additional information can still improve survey estimates.

Often, neither linked data nor parameter estimates are available, so researchers have to rely on survey data only. The remaining columns of Table 1 evaluate survey-based corrections. The SMC method in column 6 improves over the CPS reports, but it overstates the poverty reduction and the fraction of recipients in poverty. Assigning receipt to the most likely recipients clearly overimputes for those with high-predicted probabilities of receipt (such as the poorest) and underimputes for less likely recipients. Column 7 shows that this problem can be improved by assigning receipt probabilistically. This modified SMC method substantially improves over both ACS and CPS survey data. It still slightly understates the poverty reduction and overstates poverty among recipients, but produces program statistics and poverty rates similar to the conditional distribution method. Both versions of the SMC method are more accurate than the TRIM results in column 8. TRIM improves the understatement of recipients and amounts received, but the bias in the poverty rate among recipients and the poverty reduction are almost as large as with the CPS reports.

Figure 1 plots estimates of total amounts received and receipt rates for bins formed by annual reported income divided by the poverty line. Such analyses are important to evaluate program targeting and take-up. Reported annual income would be used in practice, but it should be kept in mind that eligibility depends on monthly income and that income likely is reported with error as well. These two facts may help to explain the patterns in the linked data, but not why the survey

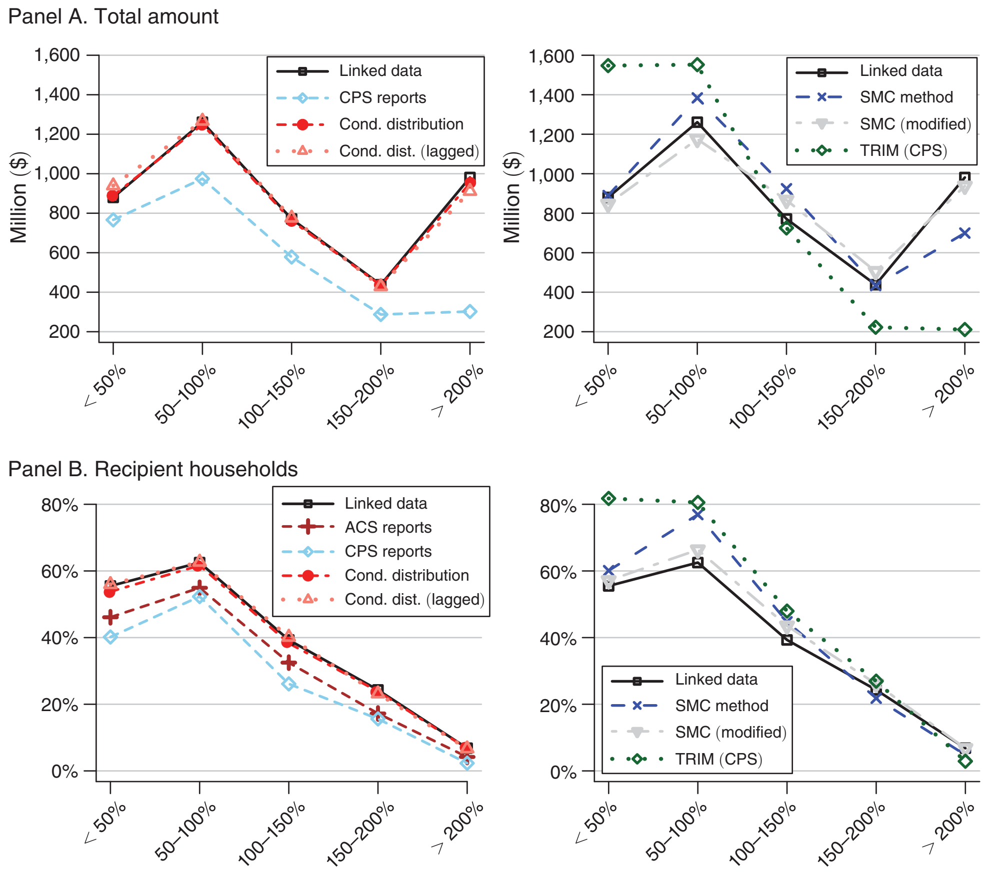

Figure 1. SNAP by Income Relative to the Poverty Line, NY 2010

results are so different. The surveys understate both amounts received and receipt rates throughout the income distribution. For amounts, the difference is particularly pronounced above the poverty line. Of the $1.4 billion missing in the CPS, $1 billion is received by households with an annual income above the poverty line. The CPS also misses that total dollars received increase for the highest income category. For receipt rates, the difference is more pronounced among the very poor, which partly explains the apparent low take-up among the very poor. Both surveys suggest that more than half of the households below 50 percent of the poverty line do not participate in SNAP, though they are likely eligible and indigent. The linked data show that almost a fifth of these households actually receive benefits, but do not report receipt in the survey.

Turning to the corrections, Figure 1 shows that the conditional distribution method, both using current and lagged parameters, reproduces the income gradient in take-up and amounts received well. Although the results from the SMC method are much closer to the linked data than the survey reports, it amplifies the bias in the gradient. Receipt rates rise more steeply up to the poverty line than in the ACS and then fall more quickly with income. This bias is particularly pronounced for

TABLE 2—SNAP BY INCOME RELATIVE TO THE POVERTY LINE, NY 2010

Income in % of HH poverty line	≤50%	50–100%	100–150%	150–200%	>200%
Panel A. Total amount received (in million $)					
Linked data	880	1,261	771	438	981
Reporting rates					
CPS reports	87%	77%	75%	66%	31%
Conditional distribution	101%	99%	99%	99%	97%
Conditional distribution (lagged)	107%	100%	101%	98%	93%
SMC method	101%	110%	120%	99%	71%
SMC method (modified)	96%	93%	112%	115%	96%
TRIM (CPS)	176%	123%	94%	51%	22%
Panel B. Percentage of households receiving SNAP					
Linked data	55.5%	62.5%	39.3%	24.3%	6.8%
Reporting rates					
ACS reports	83%	88%	83%	71%	61%
CPS reports	72%	84%	66%	64%	34%
Conditional distribution	97%	99%	97%	98%	98%
Conditional distribution (lagged)	101%	100%	102.2%	95%	98%
SMC method	108%	123%	114%	90%	70%
SMC method (modified)	103%	106%	111%	107%	99%
TRIM (CPS)	147%	129%	122%	111%	43%
Panel C. Share of households					
CPS	5.9%	7.5%	9.2%	8.2%	69.1%
Linked data	4.9%	7.7%	8.4%	8.4%	70.6%
ACS (all other methods)	5.1%	8.0%	8.5%	8.3%	70.2%

Notes: Columns are defined based on annual reported household cash income divided by the household poverty line. Rows contain the same methods as the columns in Table 1; see the notes there. The first row of each panel contains the estimate from the linked data. The remaining rows contain the estimates from the respective method divided by the same statistic from the linked data. All analyses are conducted using household weights (adjusted for incomplete linkage in the linked data).

the original SMC method, which overstates receipt between 50 and 100 percent of the poverty line more than the survey data understate it. TRIM overstates how rapidly receipt rates and amounts decline with income even further. The TRIM results show that attempting to correct for underreporting can overturn patterns that the contaminated survey data capture correctly, such as that participation rates rise with reported income below the poverty line.

The differences between survey and linked data in Figure 1 are due to reporting being most accurate between 50 and 100 percent of the poverty line. Table 2 shows the underlying numbers from the linked data and reporting rates (the ratio of survey and linked data estimates) for all other methods. Survey reporting rates vary with income non-monotonically, which presents a problem for survey-based corrections because they rely on a model of *reported* receipt. As discussed in Section IIIA, measurement error biases the parameters of this model (Meyer and Mittag 2017; Kimball, Sahm, and Shapiro 2008) and, hence, leads to systematic errors. Here, households close to the poverty line are more likely to report, so they are too likely to have receipt assigned to them. Those higher up in the income distribution are less likely to report, so the corrections still understate their receipt.

TABLE 3—COEFFICIENTS ON SNAP RECEIPT IN PROBIT MODELS OF EMPLOYMENT

			Conditional distribution		SMC method		
	Linked data (1)	CPS reports (2)	Current par. (3)	Lagged par. adj. (4)	Original (5)	Modified (6)	TRIM CPS (7)
SNAP receipt	−0.773	−0.893	−0.764	−0.759	−1.003	−0.854	−0.817
SNAP amount ($1,000)	−0.009	−0.004	−0.010	−0.011	0.011	0.008	−0.041

Notes: The dependent variable is an indicator of whether the household head is employed. See Table 1 for notes on the methods. All models control for family type, age and education categories, and race and gender of the household head. Full results including ACS reports and coefficients on the controls are in online Appendix Table A6. All analyses use household weights (adjusted for incomplete linkage in column 1).

Table 2 confirms this problem: the modified SMC method overimputes where reporting is best and underimputes in both tails of the income distribution. The problem is clearly amplified for the original SMC method. Consequently, imputing receipt to those more likely to report receipt reinforces how the survey data misrepresent multivariate relationships. TRIM overcorrects both participation and amounts in the lowest income bins with reporting rates up to 176 percent. While the survey data understate program effects below the poverty line, analyses based on TRIM (e.g., Sherman 2009; Sherman and Trisi 2015) are likely to overstate them. The conditional distribution method closely replicates the results from the linked data, which underlines the advantage of incorporating information from linkage studies.

To evaluate corrections for models with a mismeasured independent variable, Table 3 presents coefficient estimates of probit models. The models regress whether the household head is employed on SNAP receipt and amounts received, controlling for basic demographics. Similar models are used to examine labor supply effects, but program receipt is clearly endogenous here. The purpose of Table 3 is to evaluate the corrections by comparing estimates to those from the linked data in column 1. The conditional distribution method in column 3 reproduces the results from the linked data well. Using lagged instead of current parameters mainly affects the estimated coefficient on amounts received, which is small according to all methods. Contrary to the simple statistics in Table 1, there are meaningful differences between survey-based corrections and the conditional distribution method when the mismeasured variable is a regressor. The SMC method substantially overstates the difference in employment rates between recipients and non-recipients. Probabilistic imputation reduces the bias, but it remains more than ten times larger than the bias of the conditional distribution method. The TRIM estimates are closer to the linked data estimates for receipt than the SMC method and the CPS reports, but further off for amounts received. TRIM is based on the CPS, so it is not clear whether the estimates from the ACS and CPS are comparable. The coefficients on the covariates in online Appendix Table A6 differ, suggesting that the models may not be comparable. The coefficients on the covariates also show that only the conditional distribution prevents the misreporting bias from spreading to other coefficient estimates.

Overall, the estimates show that the conditional distribution method is considerably more accurate for correlations and models with mismeasured independent variables than other corrections. This advantage likely applies to multivariate analyses, especially with mismeasures independent variables, more generally. Survey-based corrections incorporate information on net underreporting from administrative aggregates, which greatly improves estimated receipt levels. Even so, they do not introduce any information regarding *which* households underreport. Consequently, the survey-based corrections cannot separate how covariates are related to misreporting from how they are related to SNAP receipt. The conditional distribution estimates the relationship of misreporting with household characteristics from the linked data and is therefore able to correct bias from systematic misreporting.

V. Extrapolation across Geography

If linked data do not exist for the population of interest, the analyst needs to choose between the survey reports, corrections without linked data or corrections based on similar linkage studies. For government benefits, many studies link only one state. It is not clear to what extent we can learn about the entire population from such geographically confined subsamples and whether using information from these samples improves over the survey-based corrections. Extrapolation will be valid if the conditional distribution, $f_{X^A|X^R,Z}$, is the same in the linked data and the population of interest. Substantively, this requires misreporting to be the same in the linked data and in the entire population. Rather than unconditional differences in misreporting, conditioning on many covariates makes parameter heterogeneity and the unexplained variation of X^A the main threats to the validity of extrapolation. In lieu of a direct test, I provide encouraging evidence from extrapolation within NY and from comparisons of extrapolated to administrative totals.

To examine extrapolation within NY, I split the linked data into two samples, east and west NY. East NY combines the eastern counties of the NYC metro area and the eastern counties of upstate NY. West NY contains the remainder of the state. See online Appendix Figure A1 and online Appendix B for detail. Table 4 replicates Table 1 for west NY.[8] The main difference from Table 1 is that columns 3 and 4 of Table 4 use the conditional distribution from east NY to simulate SNAP receipt and amounts in west NY. Online Appendix Table A7 shows that the coefficients of the conditional distribution in the two samples are similar, but a test of parameter equality rejects that they are the same. Column 4 adjusts the parameters to make total amounts and the number of recipient households match in expectation for each group of counties as discussed in online Appendix C. Adjusting the parameters does not change the results much, which underlines that most of the geographic variation is captured by the covariates. The modified SMC method in column 5 uses the imputations from the previous section for west NY. The SMC method does not rely on linked data, so there is

[8] The CPS and TRIM are not representative below the state level, so I do not report results.

Table 4—Extrapolating SNAP Statistics from East to West New York State, 2010

Parameters	Linked data — (1)	ACS reports — (2)	Conditional distribution East NY (3)	Conditional distribution ...adj. (4)	SMC method modified NY (5)
Total amount (in million \$)	2,014	—	1,987	2,013	2,127
Fraction of recipient households	17.4%	13.5%	17.0%	17.4%	18.5%
Poverty rate among recipients	36.2%	42.8%	37.3%	37.2%	37.8%
Poverty reduction	2.0%	—	2.0%	2.0%	2.1%
Mean absolute deviation to PUMA level...					
Total amount (in million \$)	—	—	4.3	2.7	4.9
Recipient households ($\times$1,000)	—	1.8	1.1	0.8	1.2
Poverty reduction	—	—	0.49%	0.51%	0.58%

Notes: All results are for west NY, which combines the western counties of the NYC metro area and the western counties of upstate NY. Column 1 uses the administrative measures from the linked ACS, column 2 ACS PUMS reports. Column 3 uses the estimated conditional distribution from east NY and the west NY ACS PUMS. Column 4 adjusts the estimated distribution to make total SNAP dollars and recipient households match the linked data in expectation in each county group (the smallest combination of counties that can be identified in the ACS PUMS data). Column 5 uses the same imputations as column 7 of Table 1, but only for west NY. Mean absolute deviations are the average absolute deviation of the respective method from PUMA totals in the linked data. All analyses conducted using household weights (adjusted for incomplete linkage in column 1).

no need to extrapolate. All three corrections clearly improve the survey estimates. For all statistics, the adjusted conditional distribution performs slightly better than the unadjusted distribution, which in turn performs marginally better than the modified SMC method. Yet, the differences between the corrections pale compared to the bias in the survey data.

The lower part of Table 4 examines how well each method captures geographic variation in program receipt. It reports mean absolute deviations from totals according to the linked data for the smallest areas that the PUMS data identify (PUMAs). All three methods perform better than the survey reports of the number of recipients. The error reductions are substantial, ranging from 37 to 59 percent. The conditional distribution method remains more accurate than the modified SMC method even when extrapolating. Adjusting the conditional distribution slightly improves extrapolation, though not for poverty reduction. That the conditional distribution method recovers the geographic distribution well is particularly encouraging because Z does not contain geographic indicators. Thus, the conditional distribution method reproduces regional variation well despite the fact that regional differences are only captured by the covariates, which is an advantage of being able to condition on many covariates (Hirsch and Schumacher 2004, Bollinger and Hirsch 2006). As above, the differences between the methods are amplified when examining the relation between SNAP and reported income. The extrapolated reporting rates in online Appendix Table A8 are almost identical to those in Table 2. Extrapolations using the unadjusted and the adjusted conditional distribution both deviate from the linked data by at most 5 percent. Both capture the income gradient

Table 5—Extrapolating SNAP Statistics to the Entire United States, 2010

	Survey reports		Conditional distribution		SMC method modified	TRIM
Data	ACS US	CPS US	ACS US	ACS US	ACS US	CPS US
Parameters	—	—	NY	NY, adj.	by state	—
	(1)	(2)	(3)	(4)	(5)	(6)
Total amount (in billion $)	—	35.28	63.31	60.61	60.64	57.64
Fraction of recipient HH	11.9%	10.0%	15.7%	14.2%	16.3%	16.5%
Poverty rate among recipients	47.1%	52.2%	41.4%	43.3%	42.4%	52.6%
Poverty reduction	—	1.3%	2.3%	2.2%	2.1%	2.4%
Mean absolute deviation of total $ received (in million $) to administrative totals . . .						
by state	—	497.2	110.4	3.0	0.0	93.9
for large MSAs	—	210.0	54.2	21.8	125.5	55.6
for county groups	—	—	10.7	8.6	12.0	—

Notes: Column 1 and 2 contain ACS and CPS survey reports. Column 3 uses the 2010 US ACS PUMS and the estimated conditional distribution from NY. Column 4 adjusts the estimated distribution to make total SNAP dollars and recipient households match administrative totals in expectation by state. Column 5 applies the method in column 7 of Table 1 by state, i.e., using state-specific models and matching state totals. Column 6 uses the (CPS-based) TRIM simulations. Mean absolute deviations are the average absolute deviation of the respective method from program spending according to the BEA. County groups are the smallest combination of counties that can be identified in the ACS PUMS data. The CPS and TRIM are not representative of county groups. All analyses conducted using household weights.

well. The SMC method improves the survey estimates, but overstates receipt and amounts received where reporting is most accurate.

Extrapolating to the entire US requires the conditional distribution, i.e., misreporting, to be identical across states, which likely is only an approximation. Nevertheless, that extrapolation works well within NY is promising. Linked data are not available to evaluate extrapolation, so I use official statistics from the Bureau of Economic Analysis (BEA) on total dollars spent by county and for large metropolitan statistical areas (MSAs) to examine how closely the corrections reproduce the geographic distribution of program spending. Online Appendix B provides further detail.

Table 5 replicates Table 4 for the entire United States. The table notes provide detail on the implementation of each method. The lower part of the table shows drastic improvements over survey reports for all methods. Even the unadjusted extrapolation reduces the mean absolute deviation to state totals by 78 percent and to MSA totals by 74 percent. Adjusting the parameters to match total recipients and amounts spent by state in column 4 further improves extrapolation. Average error is almost ten times larger when using the survey reports compared to the extrapolation. These comparisons only evaluate one aspect of the extrapolation, but as argued above, a particularly difficult one. They certainly favor extrapolation over survey reports. As for NY, extrapolation based on the conditional distribution reproduces the geographic distribution of spending more accurately than the two survey-based corrections. Still, both the SMC method and TRIM also reduce error compared to the survey data. Therefore, both approaches are attractive for studies of the geographic distribution of funds when linked data are not available or extrapolation is not feasible. Overall, both extrapolation within NY and comparisons to administrative aggregates suggest that extrapolation across geography works well for SNAP. The upper part of Table 5 shows that correcting for misreporting affects substantive conclusions. Most importantly, all corrections agree that SNAP reduces poverty by 2.1 to 2.4 percentage points. Thus, the CPS seems to understate the poverty reduction by about 1 percentage point or about 75 percent of the reduction according to survey reports. This difference in estimates of poverty rates from misreporting of SNAP alone is similar in magnitude to bias from non-response to the income question (Hokayem, Bollinger, and Ziliak 2015). The difference is larger than what Meyer and Mittag (2019) finds for NY, cautiously suggesting that the consequences of underreporting they show for NY may be even larger at the national level.

As in NY, the corrections yield similar results for descriptive statistics. The conditional distribution method, but also the SMC method, produce similar program and poverty statistics. Again, the corrections diverge more when analyzing the relation between reported income and program receipt. Online Appendix Figure A2 shows that the differences are similar to those in NY. TRIM suggests a much steeper decline of receipt with income. The results using the conditional distribution and the SMC method are substantively similar. As expected, the differences are largest where reporting is likely to be most accurate, with the SMC method suggesting higher receipt rates around the poverty line. Overall, both extrapolation based on the conditional distribution and the SMC method suggest that SNAP receipt is more widely spread throughout the distribution of annual incomes than survey reports indicate.

VI. Conclusion

Misreporting in survey data severely affects analyses of transfer programs. Most studies do not address this problem. Those that use corrections for misreporting usually provide little evidence that the corrections improve estimates. I use linked data to evaluate the performance of such survey-based corrections and to examine how one can further improve survey estimates by combining information from linked data with public use data. The findings show that data linkage can play a key role in addressing the problem of survey error.

Substantively, the linked data from NY show that the ACS and CPS miss a large fraction of SNAP receipt. Receipt is more widely spread across the distribution of annual incomes. The survey reports understate the importance of SNAP for households in deep poverty and the degree to which SNAP reduces poverty. Extrapolation suggests similar effects at the national level. The sizable biases Meyer and Mittag (2019) finds for NY may well be even larger for the entire United States. For example, correcting for misreporting increases the estimated poverty reduction by about 1 percentage point, or 75 percent of the survey estimate.

From a methodological perspective, comparing the corrections emphasizes three points. First, consistent estimates can be obtained from contaminated survey data by amending public use data with simulations based on parameter estimates from linked data. The conditional distribution method is almost as accurate as the linked data, even for analyses of mismeasured independent variables and nonlinear models, where the problem of measurement error bias looms particularly large. Thereby, it offers a simple way to sharply improve estimates when linked data exist, but access is restricted.

Second, incorporating information from linked data can outperform survey estimates and survey-based corrections even if linked data are only available for another year or a similar geographic area. Extrapolation to the entire United States reduces the estimation error in the geographic distribution of SNAP spending by up to 90 percent. So despite often being convenience samples, existing linked data can help us to learn more about the population of interest as long as reporting errors are similar.

Third, survey-based corrections can improve estimates, particularly descriptive statistics, even if no linked data exist. As Table 1 shows, the SMC methods produces program statistics and poverty rates similar to the linked data and the conditional distribution method. This accuracy is encouraging for policy analyses that often rely on simple statistics. If corrections based on linked data are feasible, they are likely preferable on grounds of accuracy and reliability. Table 3 shows that even the best survey-based methods only partially correct bias from underreporting with a mismeasured independent variable. Thus, researchers interested in causal inference particularly benefit from incorporating information from linked data and should exercise caution when applying survey-based corrections. In addition, the accuracy of corrections varies and they sometimes increase bias, making it important to evaluate the performance of different corrections and the conditions under which each correction works well. The results here provide guidance in choosing a method when other options are infeasible. The modified SMC method works almost as well as corrections based on linked data for univariate statistics such as the number of program recipients or total amounts received. For multivariate analyses, especially regression coefficients on mismeasured independent variables, it falls further behind corrections that use information from the linked data, but still substantially improves uncorrected survey estimates. The comparisons to the results from the linked data clearly show that assigning receipt probabilistically is important. The results here point to the likely direction of the bias in studies that do not do so. TRIM improves some simple statistics, but with the exception of the distribution of spending across MSAs, the modified SMC method is more accurate. TRIM sharply overcorrects below the poverty line, making it particularly problematic for studies of extreme poverty.

More generally, the findings in this paper further underline that researchers need to take measurement error into account when using survey data. Currently, researchers often have to rely on the survey data only. Carefully applying survey-based corrections can improve estimates, but researchers could do better with more information on the errors. Survey producers could provide such information from linkage studies. They could also produce simulated values or synthetic data to make estimation more convenient for data users, which would allow them to incorporate additional information that cannot be disclosed in the simulations. Ideally, data producers would provide both the estimated parameters of the conditional distribution and simulated values, which would allow data users to choose the best strategy for their specific case. While simulated data are convenient (and could incorporate important additional information), there are several pitfalls that can be avoided by releasing the model that generated the data. For example, it is important to know which variables the imputation model conditioned on to know whether results for a specific outcome model are reliable. The model is also necessary to obtain correct SEs. See Little (2012); Davern, Meyer, and Mittag (forthcoming); and Loong and Rubin (2017) for discussions of the usefulness of parameter estimates and Reiter and Mitra (2009) for the disclosure risk or publishing them.

The results here are for SNAP, but misreporting is known to affect reports of, among other things, income, education and employment status. Errors are related to both true values and important covariates, so the problems and solutions are likely similar to those found here. If validation data can be created, for example, by linking surveys to tax forms, the same methods can be applied. Whether the findings here apply to other variables is an open question, but the analyses show that linked data and approaches like the conditional distribution method can provide key tools to mitigate the pervasive problem of survey error.

REFERENCES

Abowd, John M., and Martha H. Stinson. 2013. "Estimating Measurement Error in Annual Job Earnings: A Comparison of Survey and Administrative Data." *Review of Economics and Statistics* 95 (5): 1451–67.

Bee, C. Adam, and Joshua Mitchell. 2018. "The Hidden Resources of Women Working Longer: Evidence from Linked Survey-Administrative Data." In *Women Working Longer: Increased Employment at Older Ages*, edited by Claudia Goldin and Lawrence F. Katz, 269–96. Chicago: University of Chicago Press.

Ben-Shalom, Yonatan, Robert Moffitt, and John Karl Scholz. 2012. "An Assessment of the Effectiveness of Anti-Poverty Programs in the United States." In *Oxford Handbook of the Economics of Poverty*, edited by Philip Jefferson, 709–49. Oxford: Oxford University Press.

Black, Dan, Seth Sanders, and Lowell Taylor. 2003. "Measurement of Higher Education in the Census and Current Population Survey." *Journal of the American Statistical Association* 98 (463): 545–54.

Bohn, Sarah, Caroline Danielson, Matt Levin, Marybeth Mattingly, and Christopher Wimer. 2013. *The California Poverty Measure: A New Look at the Social Safety Net*. Public Policy Institute of California. San Francisco, October.

Bollinger, Christopher R. 1998. "Measurement Error in the Current Population Survey: A Nonparametric Look." *Journal of Labor Economics* 16 (3): 576–94.

Bollinger, Christopher R., and Barry T. Hirsch. 2006. "Match Bias from Earnings Imputation in the Current Population Survey: The Case of Imperfect Matching." *Journal of Labor Economics* 24 (3): 483–519.

Bollinger, Christopher R., and Martin H. David. 1997. "Modeling Discrete Choice with Response Error: Food Stamp Participation." *Journal of the American Statistical Association* 92 (439): 827–35.

Bound, John, and Alan B. Krueger. 1991. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics* 9 (1): 1–24.

Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. "Measurement Error in Survey Data." In *Handbook of Econometrics*, Vol. 5, edited by James J. Heckman and Edward Leamer, 3705–3843. Amsterdam: Elsevier.

Celhay, Pablo, Bruce D. Meyer, and Nikolas Mittag. 2018a. "Errors in Reporting and Imputation of Government Benefits and Their Implications." Unpublished.

Celhay, Pablo, Bruce D. Meyer, and Nikolas Mittag. 2018b. "What Leads to Measurement Errors? Evidence from Reports of Program Participation from Multiple Surveys." Unpublished.

Chen, Xiaohong, Han Hong, and Elie Tamer. 2005. "Measurement Error Models with Auxiliary Data." *Review of Economic Studies* 72 (2): 343–66.

Chua, Tin Chiu, and Wayne A. Fuller. 1987. "A Model for Multinomial Response Error Applied to Labor Flows." *Journal of the American Statistical Association* 82 (397): 46–51.

Congressional Budget Office. 2013. *The Distribution of Federal Spending and Taxes, 2006*. Congressional Budget Office. Washington, DC, November.

Congressional Budget Office. 2016. *The Distribution of Household Income and Federal Taxes, 2013*. Congressional Budget Office. Washington, DC, June.

Dahl, Molly, Thomas DeLeire, and Jonathan A. Schwabish. 2011. "Estimates of Year-to-Year Volatility in Earnings and in Household Incomes from Administrative, Survey, and Matched Data." *Journal of Human Resources* 46 (4): 750–74.

Davern, Michael, Bruce D. Meyer, and Nikolas Mittag. Forthcoming. "Creating Improved Survey Data Products Using Linked Administrative-Survey Data." *Journal of Survey Statistics and Methodology*.

Giannarelli, Linda, Laura Wheaton, and Joyce Morton. 2015. *How Much Could Policy Changes Reduce Poverty in New York City?* Urban Institute. New York, March.

Hirsch, Barry T., and Edward J. Schumacher. 2004. "Match Bias in Wage Gap Estimates Due to Earnings Imputation." *Journal of Labor Economics* 22 (3): 689–722.

Hokayem, Charles, Christopher R. Bollinger, and James P. Ziliak. 2015. "The Role of CPS Nonresponse in the Measurement of Poverty." *Journal of the American Statistical Association* 110 (511): 935–45.

Hsiao, Cheng. 1989. "Consistent Estimation for Some Nonlinear Errors-in-Variables Models." *Journal of Econometrics* 41 (1): 159–85.

Ichimura, Hidehiko, and Elena Martinez-Sanchis. 2009. "Estimation and Inference of Models with Incomplete Data by Combining Two Data Sets." https://www.researchgate.net/publication/228732375_Estimation_and_Inference_of_Models_with_Incomplete_Data_by_Combining_Two_Data_Sets.

Kimball, Miles S., Claudia R. Sahm, and Matthew D. Shapiro. 2008. "Imputing Risk Tolerance from Survey Responses." *Journal of the American Statistical Association* 103 (483): 1028–38.

Levitan, Mark, and Trudi Renwick. 2010. "Using the American Community Survey to Implement a National Academy of Sciences-Style Poverty Measure: A Comparison of Imputation Strategies." *American Statistical Association Social Statistics Section Conference Proceedings* 1–15.

Little, Roderick J. 2012. "Calibrated Bayes, an Alternative Inferential Paradigm for Official Statistics." *Journal of Official Statistics* 28 (3): 309–34.

Loong, Bronwyn, and Donald B. Rubin. 2017. "Multiply-Imputed Synthetic Data: Advice to the Imputer." *Journal of Official Statistics* 33 (4): 1005–19.

Marquis, Kent H., and Jeffrey C. Moore. 1990. "Measurement Errors in SIPP Program Reports." In *Proceedings of the 1990 Annual Research Conference*, edited by Maxine Anderson-Brown, 721–45. Washington, DC: US Bureau of the Census.

McFadden, Daniel. 1989. "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration." *Econometrica* 57 (5): 995–1026.

Meyer, Bruce D., and Nikolas Mittag. 2017. "Misclassification in Binary Choice Models." *Journal of Econometrics* 200 (2): 295–311.

Meyer, Bruce D., and Nikolas Mittag. 2019. "Using Linked Survey and Administrative Data to Better Measure Income: Implications for Poverty, Program Effectiveness, and Holes in the Safety Net." *American Economic Journal: Applied Economics* 11 (2): 176–204.

Meyer, Bruce D., Nikolas Mittag, and Robert M. Goerge. 2018. "Errors in Survey Reporting and Imputation and Their Effects on Estimates of Food Stamp Program Participation." National Bureau of Economic Research (NBER) Working Paper 25143.

Meyer, Bruce D., Wallace K.C. Mok, and James X. Sullivan. 2015. "Household Surveys in Crisis." *Journal of Economic Perspectives* 29 (4): 199–226.

Mittag, Nikolas. 2019. "Correcting for Misreporting of Government Benefits: Dataset." *American Economic Journal: Economic Policy.* https//doi.org/10.1257/pol.20160618.

Moffitt, Robert A., and Gwyn Pauley. 2018. *Trends in the Distribution of Social Safety Net Support after the Great Recession*. Stanford Center on Poverty and Inequality Policy Brief. Stanford, CA, March.

Moffitt, Robert, and John Karl Scholz. 2010. "Trends in the Level and Distribution of Income Support." *Tax Policy and the Economy* 24 (1): 111–52.

Newey, Whitney K., and Daniel L. McFadden. 1994. "Large Sample Estimation and Hypothesis Testing." In *Handbook of Econometrics*, Vol. 4, edited by Robert F. Engle and Daniel L. McFadden, 2111–2245. Amsterdam: Elsevier.

Newman, Constance, and Erik Scherpf. 2013. "Supplemental Nutrition Assistance Program (SNAP) Access at the State and County Levels: Evidence from Texas SNAP Administrative Records and the American Community Survey." United States Department of Agriculture (USDA) Economic Research Service Working Paper ERR-156.

Nguimkeu, Pierre, Augustine Denteh, and Rusty Tchernis. 2019. "On the Estimation of Treatment Effects with Endogenous Misreporting." *Journal of Econometrics* 208 (2): 487–506.

Office of the Mayor. 2015. *The CEO Poverty Measure, 2005–2013: An Annual Report from the Office of the Mayor*. City of New York Office of the Mayor. New York, April.

Poterba, James M., and Lawrence H. Summers. 1986. "Reporting Errors and Labor Market Dynamics." *Econometrica* 54 (6): 1319–38.

Reiter, Jerome P., and Robin Mitra. 2009. "Estimating Risks of Identification Disclosure in Partially Synthetic Data." *Journal of Privacy and Confidentiality* 1 (1): 99–110.

Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.

Rubin, Donald B. 1996. "Multiple Imputation after 18+ Years." *Journal of the American Statistical Association* 91 (434): 473–89.

Schafer, Joseph L., and Nathaniel Schenker. 2000. "Inference with Imputed Conditional Means." *Journal of the American Statistical Association* 95 (449): 144–54.

Scholz, John K., Robert Moffitt, and Benjamin Cowan. 2009. "Trends in Income Support." In *Changing Poverty, Changing Policies*, edited by Maria Cancian and Sheldon Danziger, 203–41. New York: Russell Sage Foundation.

Sherman, Arloc. 2009. *Safety Net Effective at Fighting Poverty But Has Weakened for the Very Poorest*. Center on Budget and Policy Priorities. Washington, DC, May.

Sherman, Arloc, and Danilo Trisi. 2015. *Safety Net More Effective against Poverty Than Previously Thought*. Center on Budget and Policy Priorities Policy Report. Washington, DC, June.

Steuerle-Schofield, Lynne, Brian Junker, Lowell J. Taylor, and Dan A. Black. 2015. "Predictive Inference Using Latent Variables with Covariates." *Psychometrika* 80 (3): 727–47.

Taeuber, Cynthia, Dean M. Resnick, Susan P. Love, Jane Stavely, Parke Wilde, and Richard Larson. 2004. *Differences in Estimates of Food Stamp Program Participation between Surveys and Administrative Records*. US Department of Commerce, Bureau of the Census. Washington, DC, June.

US Census Bureau. 2015. *Income and Poverty in the United States: 2014*. US Census Bureau Current Population Report P60-252. Washington, DC, September.

Wagner, Deborah, and Mary Layne. 2014. *The Person Identification Validation System* (*PVS*)*: Applying the Center for Administrative Records Research and Applications'* (*CARRA*) *Record Linkage Software*. Center for Administrative Records, Research, and Applications, US Census Bureau. Washington, DC, July.

Wooldridge, Jeffrey M. 2007. "Inverse Probability Weighted Estimation for General Missing Data Problems." *Journal of Econometrics* 141 (2): 1281–1301.

Zedlewski, Sheila, and Linda Giannarelli. 2015. TRIM: *A Tool for Social Policy Analysis*. Urban Institute. New York, May.