IZA DP No. 10266

Correcting for Misreporting of Government Benefits

Nikolas Mittag

October 2016

Correcting for Misreporting of Government Benefits

Nikolas Mittag
CERGE-EI
and IZA

Discussion Paper No. 10266
October 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 10266
October 2016

ABSTRACT

Correcting for Misreporting of Government Benefits*

Recent validation studies show that survey misreporting is pervasive and biases common analyses. Addressing this problem is further complicated, because validation data are usually convenience samples and access is restricted, making them more suitable to document than to solve the problem. I first use administrative SNAP records linked to survey data to evaluate corrections for misreporting that have been applied to survey data. Second, I develop a method that combines public use data with an estimated conditional distribution from the validation data. It does not require access to the validation data, is simple to implement and applicable to a wide range of econometric models. Using the validation data, I show that this method improves upon both the survey data and the other corrections, particularly for multivariate analyses. Some survey-based corrections also yield large error reductions, which makes them attractive alternatives when validation data do not exist. Finally, I examine whether estimates can be improved based on similar validation data, to mitigate that the population of interest is rarely validated. For SNAP, I provide evidence that extrapolation using the method developed here improves over survey data and corrections without validation data. Deviations from the geographic distribution of program spending are often reduced by a factor of 5 or more. The results suggest substantial differences in program effects, such as reducing the poverty rate by almost one percentage point more, a 75 percent increase over the survey estimate.

JEL Classification: C15, C81, I32, I38

Keywords: measurement error, survey errors, misreporting, food stamps, poverty

Corresponding author:

Nikolas Mittag
CERGE-EI
Politických vězňů 7
Prague 1, 111 21
Czech Republic
E-mail: nikolas.mittag@cerge-ei.cz

* This research was supported by the US Census Bureau, the Population Center at NORC, the Czech Science Foundation (through grant no. 16-07603Y) and the Czech Academy of Sciences (through institutional support RVO 67985998). A previous version was titled “A Method of Correction Applied to the Food Stamp Program”. Any opinions and conclusions expressed are mine and do not necessarily represent the views of the US Census Bureau, the New York Office of Temporary and Disability Assistance (OTDA) or the NORC. I would like to thank Robert Moffit for help with the implementation of the SMC method. 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. The analyses in this paper were conducted at a secure facility at the US Census Bureau and the Chicago Census Research Data Center as part of a project of the US Census Bureau Center for Administrative Records Research and Applications and the results were reviewed to prevent the disclosure of confidential information. 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 Lupe Bedoya, Dan Black, Kerwin Charles, Pavel Čížek, Ryan Gallagher, Patrick Gaulé, Štěpán Jurajda, Jan Kabátek, Bruce Meyer, Laura Wheaton and seminar participants at Bilkent University, the 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. Marin Drlje provided excellent research assistance. All errors are mine.

1 Introduction

Recent studies use comparisons to administrative aggregates and validation data to show that government benefits are severely underreported in survey data (e.g. Lynch et al., 2007, Marquis and Moore, 1990, Meyer, Mok and Sullivan, 2015*a*,*b*, Meyer, Goerge and Mittag, 2016, Taeuber et al., 2004). A growing literature links administrative records to survey data and finds that these survey errors severely bias common analyses for which both policy makers and researchers frequently use these surveys. For example, Meyer and Mittag (2015) demonstrate that inaccurate survey data badly distort our understanding of the income distribution, poverty and program effects. While these record linkages improve our understanding of the nature, consequences and reasons for survey errors, most applied studies rely on either survey or administrative data only. The validation data are rarely used to answer substantive questions. A likely reason is that access to the linked data is usually restricted for reasons of confidentiality. In addition, validation data often only cover short time periods or small geographic areas, rather than the population of interest. Therefore, what can be done to solve or alleviate the problem of misreporting and how linked data can help to do so is still an open question.

In this paper, I examine how the contaminated survey data can still be used to study government transfer programs and how we can improve survey based estimates by incorporating information from validation data. I use SNAP (food stamps) records for New York State (NY) linked to the American Community Survey (ACS) to evaluate to what extent recently applied corrections for underreporting improve estimates. In addition, I develop a simple method that combines survey data with information from the linked data. Rather then requiring access to the confidential data, it uses an estimate of the conditional distribution of the correctly measured variable given reports and other covariates from the linked data. I show that it accurately recovers estimates from the validation data for NY, so that it can solve the problem of data access. Finally, I examine how we can learn more about the population of interest from the convenience samples that are usually validated. I provide evidence that extrapolation from NY to the entire US using the conditional distribution works well for SNAP. While corrections that do not make use of linked data already improve substantially over the survey reports, incorporating information from the linked data further reduces bias, particularly in multivariate analyses. Substantively, the results show that misreporting causes severe bias in the ACS in NY as well and that the bias from misreporting may well be even larger in the entire US than in NY. From a methodological point of view, the results show that linked data can be put to better use by employing methods such as the one developed here to mitigate the problem of data access and to extrapolate from convenience samples. Even though other corrections do not perform as well, some of them still substantively improve over the survey estimates. Thus, carefully applying such corrections is important when validation data are not even available for similar populations.

In particular, I use administrative records on SNAP receipt and amounts received from NY linked to the 2008-2010 ACS as validation data. With more than a quarter of true SNAP recipients not reporting receipt in the NY ACS, misreporting is a severe problem that affects analyses substantively. For example, survey reports understate program participation by 23 percent and the survey suggests that the program is targeted at a much poorer population. However, addressing the problem of misreporting by directly substituting the survey reports with the accurate administrative measures is currently not feasible for two reasons. First, linked data are subject to strong confidentiality requirements and rarely accessible to the interested researcher. Second, validation data are often based on a convenience sample for which a measure of truth is available, such as one state, rather than the entire survey sample.

This raises the question of how to analyze government benefits when validation data are not available. Most papers at most acknowledge underreporting, but a few corrections have been applied. They change estimates, but it is not clear to what extent they improve estimate accuracy in light of the evidence that the underlying assumptions are violated (Bollinger and David, 2005, Meyer, Goerge and Mittag, 2016). The linked data allow me to evaluate whether and when such survey-based corrections[1] improve estimates. I focus on the microsimulations based on program rules from TRIM (Zedlewski and Giannarelli, 2015) and the method first used in Scholz, Moffitt and Cowan (2009), henceforth the SMC method, which imputes government program receipt and amounts received based on models estimated from the survey reports.

In addition, I explore whether information from the linked data that can be disclosed to the public can further improve estimate accuracy. To do so, I develop a two-step method that allows researchers without access to validation data to obtain consistent estimates from public use data. The first step requires someone with access to the validation data to estimate the distribution of the accurate measures conditional on reports and other covariates. The estimated conditional distribution can be disclosed to the public. I show that combining it with the public use data in the second step sharply improves estimate accuracy. This conditional distribution method is simple to implement and can be used to correct a large class of econometric models.

I first use the NY validation data to evaluate corrections for misreporting. The conditional distribution method reproduces results from the linked data accurately, mitigating the problem of access. If validation data do not exist at all, corrections based on survey data only can still substantively improve estimates. However, even the methods that perform best are not as accurate and their performance varies. For example, the differences are larger for bivariate statistics where survey based corrections even amplify some of the biases. The results provide guidance in choosing a method of correction when validation data are not available. Then, I examine the extent to which the conditional distribution method can be used to extrapolate across geography in order to alleviate the problem that validation data often only cover a specific geographic area. I first provide evidence that extrapolating from NY to the US is promising from two comparisons. First, I extrapolate within NY and find that the method I propose is still more accurate than the survey-based alternatives. Second,

[1] I refer to them as survey-based corrections throughout, even though they sometimes make use of other data sources such as administrative aggregates as well.

I compare the geographic distribution of program expenditures from extrapolation to official statistics. While other corrections also reduce deviations from the official statistics, the extrapolation yields even larger error reductions, often reducing errors by a factor between 5 and 10. This shows that correcting based on validation data can be more accurate than the survey-based corrections even if the validation data are a convenience sample and not available for the area of interest. Substantively, the survey data seem to understate program effects even more severely at the national level than in NY. For example, the corrections suggest that survey data understate the poverty reduction due to SNAP by 75 percent of the estimated effect according to survey reports. Thus, they overstate the poverty rate (including SNAP) by about one percentage point or about twice the difference I find for NY. This difference due to misreporting of SNAP is on the order of magnitude of the bias due to non-response to the income question (Hokayem, Bollinger and Ziliak, 2015).

Overall, this study emphasizes that misreporting in survey data is severe and should not be ignored. The performance of the corrections are encouraging. While the results are for SNAP, the problem and potentially the solutions are of importance more generally, because misreporting in survey data is by no means confined to government programs. Among others, measurement error is known to affect reports of income, education, employment status and health insurance coverage. Errors have been show to be related to both true values and important covariates, which presents researchers with similar challenges as those encountered with SNAP here and suggests that similar solutions may work well. If validation data can be created, for example by linking surveys to tax forms or other records, the same methods can be applied. The conditional distribution can also be used to add variables that are missing entirely, such as SNAP amounts in the ACS and to impute missing data due to non-response. Thereby, data linkage and the methods discussed here can provide key tools to address the problem of survey data quality.

Section 2 introduces the linked data. Section 3 discusses misreporting of government benefits. Section 4 describes the corrections I examine and the conditional distribution method. Section 5 uses the linked data to analyze SNAP in NY and to evaluate corrections. Finally, in section 6 I examine extrapolation to the entire US. Section 7 concludes.

2 Linking Administrative Records to Survey Data

Measurement error severely affects variables of key importance in economic models, such as income (e.g. Bound and Krueger, 1991, Bollinger, 1998, Dahl and Schwabish, 2011, Abowd and Stinson, 2013), education (Black, Sanders and Taylor, 2003), employment status (Poterba and Summers, 1986, Chua and Fuller, 1987) and health insurance coverage (e.g. Davern et al., 2008). Bound, Brown and Mathiowetz (2001) provide an overview that emphasizes the pervasive nature of misreporting in survey data. For government benefits, comparing aggregate spending to survey totals points to severe underreporting in the survey data (Meyer, Mok and Sullivan, 2015*b*). This prompts many researchers to use administrative records instead of survey data. However, many questions cannot be answered with administrative data only. For example, they lack the required covariates to analyze subpopulations such as the elderly or single-parents. They also do not include non-recipients and therefore cannot be used to analyze program take-up. The advantages of both data sources can be combined by linking the survey data to administrative records (see for example, Marquis and Moore, 1990, Bound and Krueger, 1991, Meyer, Goerge and Mittag, 2016, and the overview in Bound, Brown and Mathiowetz, 2001). This combines the accurate administrative measure with the detail of the survey data and thereby addresses the measurement error problem.

The data I use were created by linking administrative records on SNAP benefits from the NY Office of Temporary and Disability Assistance (OTDA) to the NY sample from the 2008-2010 ACS. The ACS is the largest US household survey, including approximately 2.5 percent of the population each year. The large sample size makes it one of the most important sources of information for sub-state areas and an increasingly used alternative to the Current Population Survey Annual Social and Economic Supplement (CPS). It is representative below the state level, which allows me to examine extrapolation within NY. The ACS asks for SNAP receipt, but not for amounts received. Therefore, I also report results from the CPS in order to assess the accuracy of analyses that require program amounts, such as the poverty reduction and total amounts received. Comparability may be a concern, but the CPS is a sample from the same population[2] and the main data source for such analyses in practice and therefore is an important point of comparison.[3]

The administrative data are records of all monthly payments to SNAP recipients in NY from 2007 through 2012. They contain identifiers that have been checked by OTDA against social security records for every individual on a SNAP case. The numbers appear to be accurate: The overall total from these records differs from actual aggregate outlays by less than one percent in each year. The administrative and survey data are linked using the Person Identification Validation System (PVS) of the US Census Bureau. See NORC (2011) and Wagner and Layne (2014) for detailed descriptions. In short, the PVS uses the person data (such as address, name, gender, and date of birth) from the administrative records and the survey data 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 transformed social security number (PIK) of the record from the reference file 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 in this paper are conducted 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.[4]

[2]The CPS does not include group quarters, but I exclude them from the ACS as well.

[3]To match the reference period of the ACS, I pool the two corresponding years of the CPS and aggregate official numbers over the two calendar years. The SIPP would be another alternative at the national level, it is not representative for NY.

[4]Defining SNAP receipt and poverty status at the household level provides a cleaner comparison of estimate accuracy by abstracting from the difficult problem of defining SNAP units in the survey data. See Meyer, Goerge and Mittag (2016) for a discussion. However, the difference in definitions should be kept in mind when comparing the results to official poverty and program statistics.

Consequently, I correctly classify households as long as I do not fail to match *all* true recipients to the administrative records. More than 94 percent of the households in the survey contain at least one PIKed member and are therefore likely to be linked correctly.[5] Following Meyer, Goerge and Mittag (2016), I take the imperfect match rate into account by multiplying the weights by the inverse of the predicted probability of any household member having a PIK (Wooldridge, 2007). Appendix table A1 reports the parameter estimates used in this correction, appendix table A2 provides summary statistics for the linked data. These data have been used in several papers that discuss the linking process and data accuracy further (among others Meyer and Mittag, 2015, Celhay, Meyer and Mittag, 2016*a*,*b*, Cerf Harris, 2014, Scherpf, Newman and Prell, 2014). Given the high match rate and the quality of the administrative records, I consider the administrative data to be "truth" even though there may be small inaccuracies due to mismatches or erroneous payments. This enables me to evaluate the corrections for misreporting by comparing the estimates to this standard set by the linked data.[6]

While linked data can solve many data issues, they face two key problems as pointed out by Bound, Brown and Mathiowetz (2001): First, the data are usually only available to the researchers involved in the original data collection due to confidentiality and access rules. Second, validation data commonly do not exist for the exact population of interest. They are often only available for a few years or a small geographic area. Like most previous validations of program data, the administrative SNAP records are obtained from state agencies and therefore only cover one state. Studies on the impact of government benefits at the national level or in years where validation data are not available need to either extrapolate from the existing data or rely on reports. The method developed in this paper mitigates these two problems, because it does not require access to the validated data and does well at extrapolation across time and geography.

[5]Slightly more than 90 percent of the individuals in the survey are assigned a PIK.

[6]The assumption that they represent "truth" may not be warranted with other linked records, such as tax records, which are also based on reports. They may still be a better measure than survey reports and both measures can be combined as in Abowd and Stinson (2013) or Kapteyn and Ypma (2007).

3 Misreporting of Government Benefits

Celhay, Meyer and Mittag (2016*a*) use the same linked ACS data to examine misreporting and find high error rates that are systematically related to other variables. 26 percent of households receiving SNAP according to the administrative measure do not report receipt in the survey (false negatives). The rate of true non-recipients reporting receipt (false positives), is much lower at 1.2 percent. Previous validation studies found similar and high rates of false negatives in the Survey of Income and Program Participation (SIPP), the predecessor of the ACS and even higher rates in the CPS (Marquis and Moore, 1990, Meyer, Goerge and Mittag, 2016, Taeuber et al., 2004). In line with my results, false positives rates are low in all surveys. Less is known about misreporting of program amounts. Meyer, Mok and Sullivan (2015*b*) and Celhay, Meyer and Mittag (2016*a*) provide some evidence that SNAP amounts, conditional on reporting receipt, are reported correctly on average, but the absolute values of the errors are high. In order to improve responses on SNAP receipt, the ACS does not collect information on SNAP amounts (Hisnanick, Loveless and Chesnut, 2007). However, there is considerable interest in program amounts in the ACS, for example, to compute poverty measures for sub-state areas (e.g., Bohn et al., 2013, D'Onofrio et al., 2015, Levitan and Renwick, 2010). Data linkage and the conditional distribution method can also be used when a variable is missing entirely.

The problem of misreporting program receipt is not confined to SNAP. Meyer, Mok and Sullivan (2015*a*) demonstrate that a similar extent of misreporting plagues survey reports of other government programs. All studies find misreporting to be related to important covariates in common analyses of government programs, such as income and household composition. Thereby, these studies confirm that measurement error is not classical. Thus, arguments regarding its consequences often invoked in applied research (e.g., that it leads to attenuation or causes no bias when the dependent variable is affected) are likely to be wrong, for they only hold in the classical framework.

This extent and systematic nature of misreporting is a serious problem even for simple analyses of government programs. The high underreporting rate makes analyses based on survey data only understate program take-up. The relationship to other variables affects our understanding of who benefits from the program. For example, reporting declines as income increases, which makes survey data understate receipt above the poverty line. Misreporting changes over time, causing analyses based on survey data to confound changes in the program with changes in reporting. The effects of misreporting in more complex econometric models are even more complicated. Binary choice models, which are often used to study program take-up (e.g., Currie and Grogger, 2001, Haider, Jacknowitz and Schoeni, 2003, Currie, 2006) are one of the few cases where the bias has been examined in theory and in practice (Bollinger and David, 1997, Meyer and Mittag, 2014, Meyer, Goerge and Mittag, 2016, Nguimkeu, Denteh and Tchernis, 2016). These biases limit or skew what we learn about government programs, such as to what extent the low take-up among the elderly is due to a decline in take-up or an increase in underreporting with age. Few results exist for other econometric models, so it is hard to assess the impact of misreporting on studies, such as those of program effects on nutritional intake (Butler and Raymond, 1996, Hoynes and Schanzenbach, 2009) or labor supply (Fraker and Moffitt, 1988, Hagstrom, 1996, Keane and Moffitt, 1998).

4 Methods of Correcting for Misreporting

Despite its well-known extent, only few studies attempt to correct for misreporting. In this section, I first describe corrections for underreporting from recent applied work, focusing on the SMC method used in Scholz, Moffitt and Cowan (2009), Moffitt and Scholz (2010) and Ben-Shalom, Moffitt and Scholz (2012), a modification of it and TRIM. Then, I discuss the conditional distribution method that I use to incorporate information from the linked data.

Most corrections in applied research obtain a more reliable measure of total recipients or amounts received from sources such as the program budget and make survey totals match this number. A simple way to do so is to calculate the reporting rate and then scale up survey estimates. Only adjusting the final estimates is simple and transparent, but only feasible when the parameters of interest are affected by misreporting in a simple way. This is unlikely to be true and scaling up estimates does not work for many parameters of interest such as correlations or regression coefficients. For parameters that cannot simply be scaled up, one may still be able to use the difference between survey and administrative totals in the correction. However, this usually requires adjusting the microdata rather then just the final estimates, for example by assigning program receipt to some non-recipients. See Meyer and Sullivan (2006) for a discussion and an early correction method based on program rules. The two survey based methods I evaluate below differ in the way they decide which non-recipients to assign receipt to. The SMC method imputes additional program receipt using the probability of receipt and amounts as predicted by models estimated from the survey data. TRIM predicts eligibility and amounts based on program rules.

Specifically, the SMC method uses the survey data to estimate a probit model of program receipt,[7] and to predict the probability of program receipt for each non-recipient household. SMC then assign receipt to the households with the highest probability of receipt until the number of recipients match administrative aggregates. Amounts for the imputed recipients are assigned based on the prediction and a randomly drawn residual from a regression of amounts on basic demographics among those reporting receipt. Finally, they scale up amounts for all recipients to match total spending from official statistics. See Scholz, Moffitt and Cowan (2009) p. 218-219 for further detail. They develop this method using the SIPP, while I apply it to the ACS, which does not include program amounts. Therefore, I impute program amounts for both reported and imputed recipients based on regressions that predict reported amounts among recipients in the CPS. The ACS and the CPS are representative of the same populations and the definitions of the covariates are comparable in the two surveys. Thus, contrary to the original method, I impute amounts for all recipients and based on a

[7]Including the following covariates: income, education, marital status, number of children, race/ethnicity, gender of the family reference person, region, age of the family reference person, age of children, and participation in other programs.

different survey.

A shortcoming of this method is that it assigns receipt to the most likely recipients. This clearly leads to overimputation for likely recipients (such as the very poor) and underimputation for less likely recipients. I avoid this problem in a modified SMC method by assigning receipt with probabilities proportional to the predicted probabilities.[8] Another problem of the SMC method is that the parameter estimates used to predict the probability of receipt (and amounts) are biased, because they are estimated using the misreported data. This tends to overstate receipt by those likely to report, while still understating receipt by groups that are unlikely to report. Thereby, one further overimputes among accurate reporters and underimputes among those likely to misreport. This will preserve or even amplify the bias in estimates of the correlations with characteristics that predict misreporting. Consequently, the SMC method can be expected to improve the level of underreporting, but it is not consistent[9] and likely to preserve or amplify some biases.

Another approach is to use program rules to determine eligibility and amounts received among participants. Thereby, they avoid relying on the contaminated reports of program receipt and rely on the survey reports of determinants of eligibility and amounts instead. A problem of imputation based on program rules is that the rules determine eligibility, but not all eligible individuals indeed take up program benefits. This leads to overimputation unless one assigns receipt to a subset of those eligible only. Another problem of such methods is that determining eligibility from survey reports is very noisy (Newman and Scherpf, 2013). Often many true recipients appear ineligible according to their other reported characteristics. Whether this is due to key information such as monthly income not being reported in the survey or due to other variables such as income being misreported (or imputed) in the survey as well is unclear. The TRIM model, which I evaluate below, uses very detailed microsimulation based on program rules to address many shortcomings of this approach.

[8] A more recent implementation of the SMC method also imputes probabilistically (but in a slightly different way, I was not aware of this work in progress).

[9] Unless the survey-based estimates of the model of take-up and amounts are consistent, in which case the correction is likely to be unnecessary.

For example, it simulates monthly income to determine eligibility and assigns receipt among the eligible to make the recipient population match USDA quality control data. TRIM makes use of reported recipiency status by automatically including eligible households that report receipt of SNAP in the simulated caseload, but those with imputed receipt are not treated as reporters for this purpose. See Zedlewski and Giannarelli (2015) for further detail on TRIM. TRIM is developed and maintained by the Urban Institute and commonly used in academic and policy research (see e.g. Giannarelli, Wheaton and Morton, 2015, Sherman, 2009, 2015, U.S. Government Accountability Office, 2015). The TRIM simulations are based on the CPS, which is not representative for sub-state areas, so I cannot use TRIM in all analyses below.

Overall, on one hand, there is ample evidence that corrections based on survey data only will not fix the problem of misreporting. They are problematic if misreporting is related to other covariates, because they do not incorporate information that could separate how true program receipt varies with other variables from how misreporting varies with these variables. This makes it unlikely that they improve bias in estimates of how the variable of interest varies with other variables that also predict misreporting. This is particularly problematic, because validation studies clearly show that key variables that predict receipt also predict misreporting (Meyer, Goerge and Mittag, 2016). On the other hand, the corrections discussed above are likely to improve the *level* of underreporting by incorporating information on aggregate receipt. Consequently, they may improve estimates, but are unlikely to yield consistent estimates. This makes it important to evaluate their performance and the conditions under which they work well in order to choose among the available options when other, potentially consistent, methods are infeasible.

The literature on measurement error offers a few other corrections based on survey data only that require additional assumptions, such as restrictions on the distribution of the measurement error (e.g., Poterba and Summers, 1986, Hausman, Abrevaya and Scott-Morton, 1998, Hong and Tamer, 2003) independence assumptions that allow for an instrumental variable approach (Hausman et al., 1991, Hu and Schennach, 2008) or repeated measures with independent errors (Li, 2002, Schennach, 2004). Validation studies provide strong evidence that these assumptions do not hold, limiting their use for the case of government programs. See for example Almada, McCarthy and Tchernis (2016) for an application to SNAP where an IV approach leads to large bias. Bounds instead of point estimation avoid these assumptions (e.g. Horowitz and Manski, 1995), but tend to be too wide to be informative (Gundersen and Kreider, 2008). Others use validation data to consistently estimate models in the presence of arbitrary measurement error (e.g. Hsiao, 1989, Sepanski and Carroll, 1993, Lee and Sepanski, 1995, Li, 2002, Chen, Hong and Tamer, 2005, Ichimura and Martinez-Sanchis, 2009). For different estimators, they show that the law of total probability can be used to express an estimator that contains unobserved (e.g. accurately measured) variables in terms of observed variables and a function of the conditional distribution of the unobserved variables given the observed ones. They then propose estimators that jointly estimate this function and the parameters of interest. These estimators avoid assumptions on the measurement error and its relation to covariates, but require the validation data to be available to the researcher.[10] The estimator I use to combine the public use data with information from the validation data builds on the same insight, but I estimate the conditional distribution rather than a function of it and use it in a two-step estimator to avoid the problem of data access.

To introduce notation, consider a model including variables only available in the administrative data X^A and (potentially) other covariates Z. In the application below, X^A includes SNAP receipt and amounts received and Z contains income relative to the poverty line. The public use data only contain Z and error-prone reported versions X^R of some or all of the administrative variables. For the ACS, X^R includes reported receipt only, because amounts are not reported. The conditional distribution method can also be applied to use such variables that are missing from the public use data entirely or partially due to item

[10]Another advantage of these methods and the one I describe below over the survey-based methods discussed before is that the asymptotic properties of the latter are unknown. Thus, inference and tests usually ignore the correction, and are therefore flawed. However, given the large biases, correct inference may be considered a second-order issue.

non-response. The linked data contain (X^A, X^R, Z), so they can be used to estimate the model of interest. However, there are many cases in which that is not possible, for reasons of confidentiality or coverage as discussed above.

The intuition behind estimation from the public use data 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 of interest from the public use data. The objective function of estimators containing unobserved variables X^A can be expressed in terms of X^R, Z and $f_{X^A|X^R,Z}$ using the law of total probability. For example, a (non-linear) regression function $Y = g(X^A, \beta)$, implies

$$\mathbb{E}(Y|X^R, Z) = \int g(x^A, \beta) \cdot f_{X^A|X^R,Z}(x^A|X^R, Z)dx^A \tag{1}$$

The expression on the right only depends on variables in the public use data, which has been used to derive consistent estimators for a large class of models including maximum likelihood, minimum distance (Hsiao, 1989) and moment based models (Chen, Hong and Tamer, 2005). These models can also be estimated by the conditional distribution method. It is closely related to Chen, Hong and Tamer (2005), who derive moment conditions from a generalization of (1). They propose semi-parametric estimators that depend on a model-specific function and use both data sources simultaneously. Instead, I propose to estimate the parameters from (a potentially generalized version of) equation (1) directly using a two-step estimator based on a parametric estimate of the conditional distribution. The first step estimates the conditional distribution and requires the linked data, the second step estimates the outcome model and does not require the linked data. The advantages are practical, it solves the problem of access and is simple to implement.

To illustrate the estimation procedure, assume for now that an estimate of the conditional distribution, $\hat{f}_{X^A|X^R,Z}$, is available. As equation (1) suggests, the parameters of interest can be estimated by integrating X^A out of the objective function of the estimator. I do this by simulation, i.e. by generating D draws from $\hat{f}_{X^A|X^R,Z}(X^A|X^R = x_i^s, Z = z_i)$ for every observation $i = 1...N$ in the public use data.[11] This yields a data set containing $D \cdot N$ observations of (X_{id}^A, X_i^R, Z_i). Consistent parameter estimates can be obtained from this data set in the same way as if X^A were available in the original data, i.e. by solving or optimizing the original objective function on the expanded data. Standard errors need to be adjusted for the estimation of the first stage parameters as in Newey and McFadden (1994) and simulation error as in McFadden (1989) or Li (2000). However, the correction for simulation error can be made arbitrarily small and the correction for estimated first stage parameters is small with common sample sizes. In summary, estimation from the public use data can usually be done in three simple steps if an estimate of the conditional distribution is available: First, generate D draws for every observation from the estimated conditional distribution. Second, run the original program on the expanded data set. Third, adjust standard errors.

Consider correcting some prototypical analyses. First, the effects of SNAP are often analyzed using statistics such as poverty rates or amounts received by demographic subgroups (e.g. Scholz, Moffitt and Cowan, 2009, Sherman, 2011). After drawing D program amounts for each observation from the conditional distribution, $(D \cdot N)^{-1} \sum_{d=1}^{D} \sum_{i=1}^{N} \mathbf{1}(y_{id} < PL_i)$ consistently estimates the poverty rate, where 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 simulating D draws for every observation and computing the OLS slopes from the resulting simulated data set with $N \cdot D$ observations. Third, the binary choice models in analyses of program take-up such as Haider, Jacknowitz and Schoeni (2003) can be estimated from simulated draws of the program receipt indicator.[12] Even complex models, such as Keane and Moffitt (1998), are simple to estimate consistently by applying the routine that maximizes the original likelihood function to data that replaces the mismeasured

[11] Programs to simulate samples from common conditional distributions are available from my website.

[12] A more efficient estimator for linear regressions is to impute the conditional mean of X^A from the conditional distribution for each observation (Schafer and Schenker, 2000). For binary choice models, one can also compute or simulate predicted probabilities of misreporting for each observation and use them in the estimators discussed in Bollinger and David (1997) and Meyer and Mittag (2014).

variables by repeated draws from the conditional distribution. Thus, obtaining consistent estimates from public use data is simple if the original estimation problem is solved.

If the researcher has access to both data sets, the standard advantages of joint estimators may make the estimators from the measurement error and data combination literature (e.g., Chen, Hong and Tamer, 2005, Ichimura and Martinez-Sanchis, 2009) preferable, but obtaining access to both data sets is often not an option. In addition to using a two-step estimator to mitigate the problem of data access, I also use a parametric estimate of the conditional distribution.[13] Again, this choice is based on practical considerations. First, it mitigates disclosure problems, because a parametric estimate of the conditional distribution can be summarized by a vector of parameters and published without infringing confidentiality. Second, it facilitates conditioning on many covariates, because parametric models avoid the curse of dimensionality. Several papers have shown that conditioning on many variables is crucial to avoid bias (Bollinger and Hirsch, 2006, Hirsch and Schumacher, 2004, Steuerle-Schofield et al., 2015). Third, parametric models facilitate incorporating further information after estimation, such as adjusting the parameters to make the expectation of the distribution match known population totals. This can be used to relax the assumption that the conditional distribution does not change when extrapolating. A simple way to test the parametric assumptions is to simulate draws of X^A from the conditional distribution and test whether they come from the same distribution as X^A in the linked data using a Kolmogorov-Smirnov test. This only tests the marginal distribution, the tests proposed in Andrews (1997) and Rothe and Wied (2013) can be used to test the specification of the entire conditional distribution.[14]

Another approach to solve the problem of access are synthetic data, i.e. draws from the estimated distribution of the entire data (Rubin, 1993, Little, 1993, Abowd, Stinson and

[13] Chen, Hong and Tamer (2005) also consider, but do not implement, a closely related joint estimator based on a non-parametric estimate of the conditional distribution rather than a function of it. Brownstone and Valletta (1996) use a similar parametric two-step estimator that only applies to linear outcome models.

[14] If the assumptions fail even for flexible parametric models, one could opt for a non-parametric estimator that can be summarized by a finite set of parameters, such as a sieve estimator (Chen, 2007). However, if the convergence rate of the estimator is slow, the resulting two-step estimator may not be consistent.

Benedetto, 2006, Drechsler and Reiter, 2010). See Little (2012) for a discussion of the advantages of making an estimate of the distribution available instead as well as why a model-based approach such as the one I propose would improve over current practice. However, the main disadvantage of synthetic data is practical: Statistical agencies appear to be more reluctant to release an entire data set than a vector of parameter estimates. Contrary to synthetic data, the parameter estimates also allow the researcher to extrapolate, and choose the number of draws or to impose constraints based on additional information. More generally, the conditional distribution method can also be seen as a multiple imputation method (Rubin, 1987, 1996): SNAP receipt and amounts received are imputed multiple times and estimation is performed on the (partly) simulated data. However, I estimate the outcome model from the stacked repeated draws, rather than estimating it from each draw separately and then averaging over the D estimates as in standard multiple imputation. The two approaches are equivalent for estimators such as means that are consistent even for a single draw ($D = 1$). However, common estimators such as non-linear models or linear models with a misreported independent variable are only consistent as *both* N and D go to infinity, because the simulation error in every single draw causes bias similar to classical measurement error. For these models, the parameters estimates in each replication of standard multiple imputation and hence their average are inconsistent. In line with this, Schenker, Raghunathan and Bondarenko (2010) find an improved, but attenuated relationship when using multiple imputation to correct a non-linear model. The conditional distribution method remains consistent for these estimators by simulating the integral in the objective function. However, this advantage can easily be extended to any multiple imputation estimator by solving the estimation problem on the stacked data rather than repeatedly.

I estimate the conditional distribution by maximum likelihood, because my data contain (X^A, X^R, Z).[15] Estimation is simplified by transforming SNAP receipt and amounts received

[15] In the absence of validation data, the conditional distribution may be identified based on repeated measures (e.g Abowd and Stinson, 2013, Bonhomme and Robin, 2010, Kapteyn and Ypma, 2007, Schennach, 2004) or independence assumptions (e.g Hu and Schennach, 2008).

into a univariate problem: an amount of zero is equivalent to no receipt. 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. The parametric framework makes it simple to preserve this mixed nature of the variable. I use a standard Probit for the probability of receiving an amount of zero. For non-zero amounts, a left truncated normal density in which the mean is a function of the covariates fits the data well.[16] This yields the following conditional distribution:

$$f_{X^A|X^R,Z}(X^A|X^R = x^R, Z = z) = \begin{cases} \Phi(x^R\alpha + z\beta; 0, 1) & \text{if } X^A = 0 \\ \left[1 - \Phi(x^R\alpha + z\beta; 0, 1)\right] \frac{\phi(X^A; x^R\gamma + z\delta, \sigma)}{1 - \Phi(\tau; x^R\gamma + z\delta, \sigma)} & \text{else} \end{cases} \tag{2}$$

$\alpha, \beta, \gamma, \delta, \sigma$ and τ are estimated from the data. This model is similar to a Tobit type II model with an additional truncation point τ. Thus, the interpretation of the parameters is similar: The parameters of the mass point, α and β, are similar to Probit parameters, with the slight difference that they determine the probability that a household does *not* receive SNAP. The parameters of the truncated normal, γ, δ, σ and τ, determine the mean and variance of amounts received if the household participates. The choice of conditioning variables Z is a standard specification issue, additional guidance for similar cases is given in Steuerle-Schofield et al. (2015). Parameter estimates from the 2009 and 2010 linked ACS are in appendix table A5.[17] The estimates are not surprising: Reporting status, including non-response, is a strong predictor of receipt and amount received. Household composition and income capture most of the remaining variation. The results in section 5 and 6 are robust to reasonable specification changes.

As pointed out above, an advantage of a parametric model is that it is simple to incorporate information by modifying or adding parameters. For government transfers, total

[16]Note that while this restricts the density to be a truncated normal at every value of the covariates, it does not force the marginal density of SNAP amounts to be a truncated normal. Neither a right truncation point nor mixture models improve the model, nor does relaxing the rate of decay of the tails by using a t-distribution. Programs to estimate these conditional distributions are available from my website.

[17]All parameter estimates and variance matrices for 2008-2010 can be downloaded from my website.

spending and recipients are often available from official statistics. One can adjust the intercepts for the mass point and amounts to make the expected value of SNAP recipients and amounts match the official numbers. If the official numbers are available for geographic or demographic subpopulations, one can also add separate intercepts for these subpopulations to β to make the expected number of participants match and to δ to make total amounts match. In the application to SNAP below, I use total dollars received from the Bureau of Economic Analysis (BEA) and the number of recipient households from the US Department of Agriculture (USDA).[18] The BEA numbers are available at the county, MSA and state level. However, the smallest geographic areas in the ACS public use data are PUMAs, so I define county groups as the smallest geographic areas that can be identified in both data sources.[19] I use the aggregates from the BEA and USDA in the extrapolation to other states in section 6.2, but use aggregates calculated from the linked data for the extrapolation within NY to isolate the difference in the results that is due to extrapolation.

5 Re-assessing SNAP in New York State

In this section I compare results from the linked data to those from unlinked survey data and corrections for misreporting. The survey reports underestimate the poverty reduction, suggest that SNAP serves a poorer population and miss a large share of program benefits higher up in the distribution of reported annual incomes. The conditional distribution method reproduces results from the linked data accurately, so it can solve the problem of data access. Corrections based on survey data only can still substantively improve the problem. However, they are not as accurate as the conditional distribution method, particularly for bivariate

[18] I pool weighted fiscal years to match the time period of the ACS as in Meyer, Mok and Sullivan (2015*b*). The USDA uses a different household definition and publishes average monthly participation instead of annual participation. For NY, the number of participating households is 8-10 percent lower in the linked data and I use this factor to make the official aggregates for other states comparable to the survey estimates. The model is non-linear, so I calculate the intercepts using a Newton-Rhapson procedure that adjusts each intercept in β iteratively until the expected number of recipients match. The intercepts in δ are calculated the same way, but take the adjustment of β into account.

[19] The US Census Bureau defines PUMAs as statistical geographic areas containing at least 100,000 people. I construct county groups as the smallest combination of counties that does not cut PUMA boundaries.

analyses, where survey based corrections even amplify some of the biases.

Table 1 provides a first impression of how descriptive statistics of SNAP in NY vary across data sources and methods. Column 1 contains the results from the linked ACS data, which are the point of reference to which I compare the other results. Column 2 uses the ACS Public Use Micro Data (PUMS).[20] The ACS does not ask for SNAP amounts, so I also provide results from CPS reports in column 3. All years refer to ACS survey years, summary statistics are in tables A2-A4. Comparing the results from survey reports in columns 2 and 3 to those from the linked data in column 1 confirms that misreporting causes substantial bias. As previous studies show for other surveys and states, survey reports fail to capture a large fraction of receipt. For example, the first row shows that the CPS underestimates the total amount of SNAP received in NY by an impressive $1.4 billion, missing a third of dollars received. As documented in Meyer and Mittag (2015), this leads the survey to substantially understate program effects. The receipt rates in row 2 are also higher than we are lead to believe by the reports, by 29 percent in the ACS and by almost 50 percent in the CPS.

Table 1: SNAP in New York State by Data Source, 2010

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Reports		Cond. Distribution		SMC Method		
	Linked Data	ACS PUMS	CPS	current par.	lagged par. adj.	orig.	modified	TRIM CPS
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 file. Columns 2 and 3 contain survey reports from the ACS and CPS. Columns 4 uses the conditional distribution estimated using the 2010 linked ACS, column 5 uses parameters from 2009 adjusted to make total amounts and recipient households match administrative numbers in expectation. Column 6 uses imputations in the ACS PUMS as in Scholz, Moffitt and Cowan (2009), column 7 uses a similar approach, but assigns receipt probabilistically. Column 8 uses CPS microsimulations from the TRIM model of the Urban Institute. All analyses use household weights (adjusted for PIK probability in column 1).

The net understatement of transfers in the previous paragraph can also be inferred from comparisons to administrative aggregates, but linking data at the household level is necessary

[20] The PUMS data have been edited for confidentiality and contain about 75% of households the restricted internal ACS. The linked data were created from the internal ACS, but the PUMS data would be used in practice. Results using the internal ACS file are very similar and available upon request. Due to the large sample sizes, standard errors are too small to affect the results substantively throughout.

to assess how well the survey captures multivariate relationships such as program targeting, take-up and its effects on poverty. The last two rows of table 1 examine the relation of program receipt to income relative to the poverty line by examining the poverty rate among recipients and a simple measure of the poverty reduction.[21] Row 3 shows that misreporting affects analyses of program targeting, because the reports in both surveys overstate the fraction of recipient households below the poverty line. Recipients above the poverty line are less likely to report SNAP receipt, which makes the program look more focused on those with very low annual income than it really is. The ACS overstates the fraction of recipients in poverty by 15 percent, which is sizable, but less pronounced than the overstatement by 45 percent in the CPS. The last line of Table 1 shows that the survey data understate program effects. The poverty reduction due to SNAP is 23 percent higher according to the linked ACS than what the CPS suggests. Similar to what Meyer and Mittag (2015) find when analyzing linked CPS data for 2008-2012, the CPS misses a reduction of poverty by half a percentage point. These findings extend previous results on survey errors to the ACS and underline that the substantive differences they find are not confined to the CPS.

The key question in this paper is whether and how researchers can improve survey based estimates. Columns 4 to 8 of Table 1 examine the performance of such corrections. Columns 4 and 5 use the conditional distribution method, columns 6 and 7 use the SMC method and the revised version of it that imputes probabilistically. Column 8 contains results from the TRIM simulations. All corrections make receipt rates and total amounts match administrative totals. Thus, the first two rows of the table are close to the linked data by construction and not informative about the performance of the methods.[22]

In contrast to the meaningful differences between the linked data and the reports, the

[21] Poverty status is determined using household income and poverty lines. The poverty reduction is calculated as the difference in the poverty rate when using cash income and when adding SNAP, i.e. it only has a causal interpretation if one assumes that there are no behavioral effects.

[22] There are minor differences: The conditional distribution method matches the numbers in expectation rather than exactly within sample. The SMC method matches amounts exactly, but makes the number of recipient individuals rather than households match. Thus, the number of recipient households is slightly overstated (the conditional distribution method matches households and conversely understates recipient individuals). TRIM matches slightly different numbers.

results from the ACS PUMS data and the estimated conditional distribution in column 4 are virtually identical to those in column 1. This shows that the conditional distribution method can recover the correct estimates from the public use data, allowing researchers without access to the confidential validation data to correct their estimates if current estimates of the conditional distribution are available. Parameter estimates of the conditional distribution are often not available from contemporaneous data, either because data linkage takes time or because validation data are only available for some years. This raises the question whether validation data from previous years can be used to improve over the survey results, as in, for example, Bollinger and David (1997), Davern et al. (2009) and Davern, Meyer and Mittag (2016). It assumes that the relationship between X^A and (X^S, Z) remains the same over time, which is at best an approximation.[23] Multiple years of validation data allow me to assess whether this approximation still improves over the survey data by using the conditional distribution from the 2009 ACS to correct for misreporting in the 2010 ACS. I adjust the conditional distribution to make the number of recipient households and total amounts received match the 2010 linked data in each county group as described in section 4.[24] The performance of extrapolation to other years is application-specific, but the results in column 5 are promising: Extrapolation closely replicates the poverty reduction and the poverty rate among recipients. This shows that even though the conditional distribution changes over time, a rich set of conditioning variables and incorporating additional information can still improve over survey reports.

In summary, columns 4 and 5 show that combining estimates from the linked data with public use data performs almost as well as the linked data and much better than the survey data. However, often neither validation data nor parameter estimates are available. This raises the question to what extent the methods to correct for underreporting from previous

[23]A formal test rejects that the parameters of the conditional distribution are the same for 2009 and 2010, which is not surprising given the large sample. While most parameter estimates are qualitatively similar, there are some differences that likely reflect changes in economic conditions.

[24]The unadjusted results (available upon request) show that allowing the level of receipt to change is important. They still improve considerably over the survey data, but underestimate program effects since they only partly capture the expansion of the program from 2009 to 2010.

studies without validation data improve over the survey data and by how much they fall short of methods that incorporate information from the linked data. The remaining columns of table 1 evaluate the survey-based corrections described in section 4.

The original SMC method in column 6 improves over the CPS reports, yet it now overstates the poverty reduction and further overstates the fraction of the recipients below the poverty line. Assigning receipt to the most likely recipients rather than probabilistically clearly overimputes receipt for the poorest and underimputes for less likely recipients. As column 7 shows, this problem can be fixed by assigning receipt probabilistically based on the predicted probabilities, as discussed in section 4. The results are closer to the linked data than both the survey data and the original SMC method. Thus, the SMC method substantially improves over the survey data, but probabilistic assignment of receipt is important. However, the modified SMC method still slightly understates the poverty reduction and overstates poverty among recipients. Yet both flavors of the SMC method produce statistics that are closer to the linked data than the TRIM results for the NY sample from the CPS in column 8. While TRIM improves the understatement of recipients and amounts received, it overstates the poverty reduction relative to the linked data almost as much as the CPS understates it. It leaves the fraction of recipients below the poverty line almost unchanged and thereby substantially overstates it.

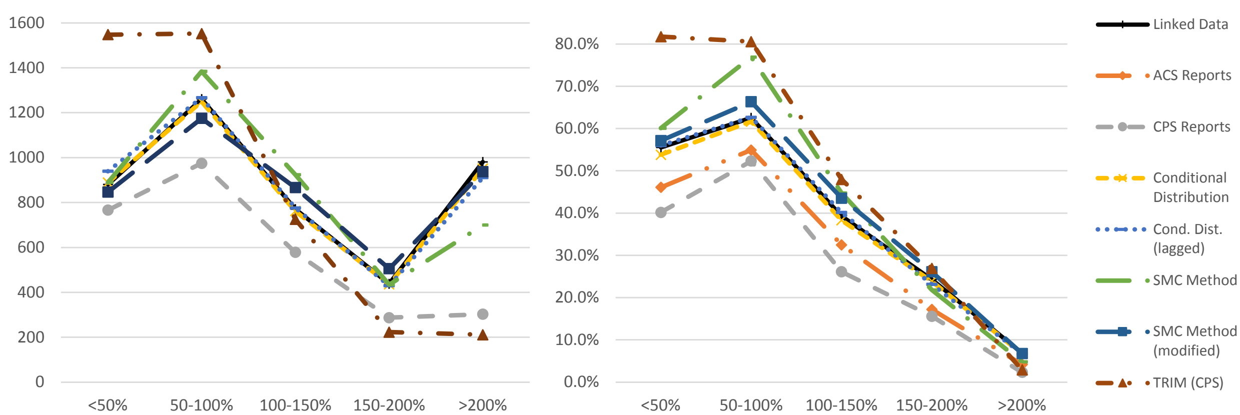
Figure 1 further analyzes the consequences of misreporting and the corrections. It plots estimates of total amount received (left panel) and receipt rates (right panel) for bins formed by annual reported income divided by the poverty line.[25] Such analyses are important to evaluate program targeting and take-up. They are often extended to the multivariate case using binary choice models, but I focus on the bivariate case here for ease of exposition. As I argue below, the patterns I find for bivariate analyses likely extend to the multivariate case. The differences between reports and the linked data in the left panel reveal that

[25]Note that program eligibility depends on a different income measure (monthly SNAP unit income). Also, reported income may contain substantial error as well, but would be used for such analyses in the absence of a better measure. While these two facts may help to make sense of the patterns in the linked data, they do not help to explain why they are so different in the survey data.

households above the poverty line benefit from SNAP more than the survey suggests. The CPS understates total amounts received at all income levels and the difference increases with income. Of the 1.4 billion dollars missing in the survey, one billion dollars are received by households with annual income above the poverty line. While most differences due to underreporting are larger below the poverty line in absolute terms, the survey misses a larger fraction of amounts received above the poverty line and the non-monotonic relation to income. While most studies focus on receipt close to or below the poverty line, this makes correcting for underreporting particularly important when examining the effect of transfer programs higher up in the distribution of annual incomes.

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

Total Amount (in Million $) Fraction of Recipient Households



The right panel shows that receipt rates are higher throughout the income distribution according to the administrative data, but contrary to amounts, the survey correctly captures the shape of the income gradient. Survey underreporting partly, but not fully, explains low take-up among the very poor. Both surveys suggest that less than half of the households below 50 percent of the poverty line participate in SNAP, even though these households have very low resources and are likely to be eligible. However, the participation rate in the linked data is 55.6 percent. Almost a fifth of these households only appear not to take up benefits in the survey, as they are recipients according to the linked data, but do not report receipt in the survey. However, receipt rates are still far from complete, so the question of low take-up among the poorest households remains open.

Turning to the corrections, figure 1 shows that the conditional distribution method, both based on current and lagged parameters, reproduces the income gradient in take-up and amounts received well. While the results from the SMC method are substantially closer to the linked data than the survey reports, it amplifies the bias in the gradient. For example, receipt rates rise even steeper up to the poverty line than in the ACS and fall even quicker with income. This is particularly pronounced for the original SMC method, which overstates receipt between 50 and 100 percent of the poverty line more than the survey data understate it. This makes the gradient much steeper on either side, which overall underlines that assignment of receipt should not be deterministic. The modified SMC method greatly mitigates this problem, but does not fix it entirely. TRIM overstates how steep the decline of receipt rates and amounts is with income even further. It also does not reproduce the non-monotonic shape, suggesting instead that both receipt rates and amounts decline with income monotonically. This shows that attempting to correct for underreporting can overturn patterns that the contaminated survey data capture correctly, such as that participation rates rise with income below the poverty line.

The differences between survey and linked data in figure 1 are driven by the fact that reporting is most accurate among households with income between 50 and 100 percent of the poverty line. Table 2 shows the underlying numbers from the linked data as well as reporting rates[26] for all other methods. Reporting rates vary with income, so misreporting leads to non-classical measurement error. While the conditional distribution method closely replicates the results from the linked data, as shown by reporting rates close to 100 percent, the non-monotone reporting pattern presents a challenge for survey-based corrections. Simple corrections, such as scaling up program amounts or participation uniformly, understate participation and amounts above the poverty line and overstate the impact below the poverty

[26] I define reporting rates as the respective estimate divided by the true number from the linked data.

line. More generally, all survey-based corrections have to rely on a model of *reported* receipt. The measurement error from underreporting will bias the parameters of this model (Meyer and Mittag, 2014) and hence lead to systematic errors. The reporting pattern observed here implies that households close to the poverty line, who are more likely to report, 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. Consequently, imputing receipt for those more likely to report receipt will reinforce how the survey data misrepresent such multivariate relationships. While the correlations and hence the biases likely differ in other applications, this makes the advantages of incorporating information from the linked data likely to generalize to other cases.

The results from the corrections based on survey data only in table 2 confirm this problem. The modified SMC method overimputes where reporting is best and underimputes in both tails of the income distribution. This leads to a much wider range of reporting rates between 93 and 115 percent. The original SMC procedure additionally suffers from the problem that it overimputes among the poor as discussed above. It overstates amounts received (by 23 percent) and receipt rates (by 10 percent) for the most accurate reporters between 50 and 100 percent of the poverty line and in return understates it for those above 150 percent of the poverty line (by up to 30 percent). This again emphasizes the importance of assigning receipt probabilistically, but also shows that the direction of the bias in studies that assign receipt deterministically is predictable. The TRIM results in the last row of each panel forcefully underline that it is difficult to recover the true relationship between SNAP receipt and income from survey reports only. TRIM substantially overcorrects both participation and amounts in the lowest income bins with reporting rates up to 176 percent. This suggests that while the survey data understate program effects for households below the poverty line, analyses based on TRIM (e.g. Sherman, 2009, 2015, U.S. Government Accountability Office, 2015) are likely to overstate them.

The fact that survey-based corrections do not reproduce bivariate relationships well, even

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%
Total Amount Received (in Million $)					
Linked Data	880	1261	771	438	981
Reporting Rates					
CPS Reports	87%	77%	75%	66%	31%
Cond. Distribution	101%	99%	99%	99%	97%
Cond. 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%
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%
Cond. Distribution	97%	99%	97%	98%	98%
Cond. 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%
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. The rows for each measure contain the same methods as the columns in table 1, see the notes there. Reporting rates are the estimates from the respective method divided by the same statistic from the linked data (in percent). All analyses conducted using household weights (adjusted for PIK probability in the linked data)

though they improve levels, is not surprising and likely to extend to multivariate analyses and other applications. The survey-based corrections incorporate additional information on net underreporting by using administrative aggregates, which greatly improves receipt levels. However, they do not introduce any information regarding *which* households underreport, so they cannot correct for systematic misreporting. The conditional distribution estimates the relationship of misreporting with household characteristics from the linked data and is therefore able to correct for non-classical measurement error. Consequently, it is likely to produce better results particularly when interested in multivariate relationships where some of the variables of interest are related to misreporting.

Overall, the results for NY underline that misreporting in surveys substantively skews our understanding of the targeting, take-up and effects of anti-poverty programs. From a methodological point of view, comparing the corrections emphasizes three points. First, one can substantively improve over the survey data even if no validation data are available. Particularly differences in levels, such as the receipt rate in deep poverty, are much closer to truth for the modified SMC method. Second, the accuracy of corrections varies, so choosing a good method is crucial. That the modified SMC method works well suggests that it is important to use the additional information the survey provides and to assign receipt probabilistically. Third, even the best methods that rely only on survey data just partially correct the bias from underreporting. None of the survey based corrections captures the income gradient well, so methods that use validation data seem preferable when examining correlations or multivariate models. The conditional distribution can solve the problem that even when linked data exist, access is often restricted. The extrapolation across time cautiously suggests that corrections based on past validation data can still improve estimates when current validation data are not available.

6 Extrapolation Across Geography

Validation data often do not exist for the exact population of interest. When the data of interest have not been validated, the analyst needs to choose between using the survey reports, corrections that do not require validation data or corrections based on similar validation studies. For government benefits, most studies validate only one state. Thus, whether the bias from misreporting is similar in NY and the US and how much we can learn about the entire population from a geographically confined subsample are important, but open questions. I examine how well extrapolating across geography using the conditional distribution method works and compare it to the survey-based corrections. Extrapolation is simple in practice: Implementation is the same as in section 5, but requires that the conditional distribution is the same in the validated and the target data. In lieu of a direct test of this key assumption, I provide encouraging evidence from extrapolation within NY and from comparing extrapolated to administrative totals. Extrapolation to the entire US shows that the conditional distribution method reproduces the geographic distribution of SNAP spending more accurately than survey reports and survey-based corrections. This provides evidence that despite often being convenience samples, current validation studies can be put to further use by improving analyses of the population of interest. The other corrections also improve over the survey data. All corrections suggest that the survey data understate program effects even more severely at the national level than in NY. For example, the poverty reducing effect of SNAP is understated by about one percentage point or 75 percent of the reduction according to survey reports.

6.1 Testing Extrapolation within New York

SNAP receipt varies between urban and rural areas as well as across neighborhoods, which makes extrapolation across geography challenging. An advantage of parametric models of the conditional distribution for extrapolation is that they simplify conditioning on many covariates. Rather than the extent of unconditional heterogeneity, this makes parameter heterogeneity and the residual variation the crucial determinants of the precision of the extrapolation. Based on simple variance decompositions, I find that 98 percent of the variation in receipt rates and 60 percent of the variation in average amounts received across PUMAs are explained by the covariates I use in the conditional distribution. These results are encouraging, indicating that a substantial share of the geographic variation is captured by the covariates. To investigate this issue in more detail, I split the linked data into two samples based on geography and extrapolate from the eastern to the western subsample.[27] The coefficient estimates for the main covariates in appendix table A6 are similar, but a test of parameter equality rejects that the distributions are the same.

Table 3: Extrapolation from East to West New York State, 2010

	(1)	(2)	(3)	(4)	(5)
	Linked Data	ACS Reports	Conditional Distribution		SMC Method modified
Parameters	-	-	East NY	...adj.	NY
Total amount (in million \$)	2,014	-	1,987	2,013	2,127
Fraction of recipient HH	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 HH (×1000)	-	1.8	1.1	0.8	1.2
poverty reduction	-	-	0.49%	0.51%	0.58%

Notes: All results are for the western PUMAs. Column 1 uses the administrative measures from the linked ACS data, column 2 reports from the ACS PUMS. Column 3 uses the parameter estimates of the conditional distribution from the eastern PUMAs and the ACS PUMS from the western PUMAs. Column 4 adjusts the estimated distribution to make total amounts and recipient households match the linked data in expectation in each county group. Column 5 uses the same imputations as column 7 in table 1, but only the western PUMAs. Mean absolute deviations are the average absolute value of the difference between PUMA totals in the linked data and the method in the respective column. All analyses conducted using household weights (adjusted for PIK probability in column 1).

However, the results in table 3 indicate that extrapolation still substantially improves over the survey data and performs slightly better than the best other correction. As in table 1, the first two columns report results from the linked data and the ACS survey data, but for the western subsample only. I do not report any results from the CPS, because it is not representative at the sub-state level. Columns 3 and 4 use the conditional distribution from the eastern PUMAs to predict SNAP receipt and amounts in the western subsample. Column 4 adjusts the parameters to make total amounts and the number of recipient households match in expectation for each county group as discussed in section 4. Adjusting the parameters does not change the results much, which underlines that most of the geographic

[27] To avoid extrapolating from NYC to Upstate NY or vice versa, I split both NYC and Upstate NY into eastern and western PUMAs. See Appendix figure A1 for a map and table A2 for summary statistics.

variation is captured by the covariates. Column 5 reports results from the modified SMC method.[28] It does not rely on validation data, so there is no need to extrapolate. Therefore, I use the imputations from the previous section, but restrict the sample to the western PUMAs. The results of all three corrections are similar to the true results obtained from the linked data and certainly improve over the survey data. For all statistics, the adjusted conditional distribution performs slightly better then the unadjusted distribution, which in turn performs marginally better than the SMC method. However, the differences certainly pale compared to the bias in the survey data.

Next, I examine how well the extrapolation captures the geographic variation in SNAP spending, recipients and the poverty reduction. Such regional analyses are often of interest for policy (see e.g. Bohn et al., 2013, D'Onofrio et al., 2015, Newman and Scherpf, 2013). Statistics on spending and participation are available for small geographic areas, so they also provide an important way to evaluate the extrapolation to the entire US, where "truth" from the linked data is not available. The lower part of table 3 reports the mean absolute deviation from the PUMA totals according to the linked data for each estimation method. Comparing the error in the estimated number of recipients in the second row further underlines that all three methods perform better than simply relying on the survey reports. The error reductions are substantial, ranging from 37 percent in column 5 to 59 percent in column 4. Note that column 4 and 5 makes the number of recipient households and total amounts match the numbers from the linked data for larger geographic areas (county groups for column 4 and the entire state for column 5). Thus, the numbers are not forced to match exactly, but one may be reluctant to compare the numbers in column 4 and 5. However, even the estimates from the unadjusted conditional distribution show that extrapolation across geography based on the linked data can not only improve upon the survey data, but can also perform better than survey-based corrections. The results in column 5 suggest that adjusting the conditional distribution to match county group totals may further improve

[28]Results from the original SMC method are available upon request. The modified, probabilistic, version performs better as in the previous section.

the extrapolation. The fact that it recovers the geographic distribution well is particularly encouraging, because the conditional distribution does not contain geographic indicators in its conditioning set. Thus, regional variation is only captured by the covariates, which makes it particularly important to include a large set of covariates (Hirsch and Schumacher, 2004, Bollinger and Hirsch, 2006).

Table 4: Extrap. From East to West NY: SNAP by Income in % of the Poverty Line, 2010

Income in % of HH Poverty Line	≤50%	50-100%	100-150%	150-200%	≥200%
Total Amount Received (in Million \$)					
Linked Data	379	575	372	191	496
Reporting Rates					
Conditional Distribution	103%	98%	98%	100%	97%
Cond. Distribution (adjusted)	105%	99%	99%	101%	97%
SMC Method (modified)	101%	98%	111%	134%	103%
Percentage of Households Receiving SNAP					
Linked Data	55.8%	61.1%	39.4%	23.1%	7.1%
Reporting Rates					
ACS Reports	88%	89%	83%	74%	61%
Conditional Distribution	98%	99%	97%	99%	97%
Cond. Distribution (adjusted)	101%	100%	99%	101%	99%
SMC Method (modified)	105%	108%	109%	115%	100%
Share of Households					
Linked Data	4.4%	7.8%	8.3%	8.2%	71.4%
ACS (all other methods)	4.5%	7.2%	8.0%	8.3%	72.0%

Notes: Columns are defined based on annual reported household cash income divided by the household poverty line. The rows for each measure contain the same methods as the columns in table 3, see the notes there. Reporting rates are the estimates from the respective method divided by the same statistic from the linked data (in percent). All analyses conducted using household weights (adjusted for PIK probability in the linked data).

Table 4 replicates table 2 to examine whether extrapolation reproduces the relation between SNAP and income relative to the poverty line well. As before, the corrections all improve bivariate analyses, but the differences between the methods are amplified. Both the unadjusted and the adjusted extrapolation based on the conditional distribution deviate from the linked data by at most 5 percent. Reporting rates are almost constant and close to 100 percent for both versions of the method, so both capture the income gradient well. As in table 2, the SMC method improves the income gradient compared to the survey estimates. However, it overstates both receipt and amounts received between 100-200 percent of the poverty line with reporting rates up to 134 percent. Overall, the results in table 4 further underline that using the information from the linked data in corrections such as the conditional distribution method is particularly important when examining bivariate (and likely also multivariate) statistics.

In summary, extrapolation within NY works well for SNAP. The extrapolation error in this sample application is smaller than the error due to misreporting even after applying the correction that works best in the previous section. The analyses provide evidence that incorporating information from validation data can outperform survey-based corrections even if validation data are only available for another geographic area.

6.2 Extrapolation to the Entire US

Extrapolating to the entire US requires the conditional distribution, i.e. misreporting, to be identical across states. This is difficult to verify and likely only an approximation. There is no variation in state institutions within NY and differences between regions may be more pronounced when extrapolating to other states. However, NY is a large and demographically diverse state with variation in the administration of the program by county.[29] Consequently, the evidence in the previous section that extrapolation works well within NY suggests that extrapolation to the entire US may be promising.

The results on the geographic distribution of spending in table 5 provide further supportive evidence. Column 1 and 2 use the ACS and CPS reports. Column 3 uses the parameter estimates from NY to predict SNAP receipt and amounts in the 2010 ACS public use data for the entire US. The results in column 4 adjust these parameter estimates as discussed in section 4. The adjustment makes the expected number of recipient households match USDA state aggregates and total dollars spent by state match state totals from the BEA. Column 5

[29]Note that SNAP is administered by the states, but it is a federal program with uniform rules.

reports the results from applying the modified SMC method to the ACS state by state. That is, both the models for receipt and amounts received are estimated for each state and total recipients and amounts received are adjusted to match state totals.[30] Column 6 contains results from the TRIM microsimulations for the CPS US sample.

Contrary to the analyses for NY, linked data are not available to assess the accuracy of the extrapolation. However, I can still evaluate the methods by comparing aggregates to official statistics in the lower part of table 5. The BEA publishes total dollars spent according to administrative sources by county and for large metropolitan statistical areas (MSAs). In practice, one may prefer to make the correction methods match these BEA totals. Instead, I use the sub-state BEA numbers to measure the accuracy of the corrections for misreporting by computing mean absolute deviations of total dollars spent by county group to the administrative numbers. In order to evaluate the CPS reports and TRIM, which are only representative for large MSAs below the state level, I also compute mean absolute deviations of total dollars spent by MSA.[31]

The mean absolute deviations show drastic improvements over survey reports for all methods. Even for the unadjusted extrapolation, the mean absolute deviation in the survey data is almost 5 times as large at the state level and nearly 4 times as large at the MSA level. Adjusting the parameters to match state totals further improves extrapolation. Column 4 shows that the average error in MSA totals for the adjusted extrapolation is only 11 percent of the average error according to survey reports. Thus, the amount by which one gets total spending wrong is almost 10 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 and they certainly favor extrapolation over survey reports. The mean absolute deviations also provide evidence that even when extrapolating across geography, using the conditional distribution can be more accurate than the SMC

[30]Results from applying the method to the entire US are similar and available upon request.

[31]I only use MSAs with more than half a million inhabitants and exclude MSAs with more than 1 percent of the population in rural PUMAs to ensure comparability to the ACS.

method and TRIM. Both methods still reduce error rates compared to the survey data. The SMC method captures the distribution of funds over county groups almost as accurately as the unadjusted conditional distribution. TRIM even performs a bit better than the unadjusted extrapolation for state totals and on par with it for MSAs. This makes both approaches attractive for studies of the geographic distribution of funds when validation data are not available or extrapolation is not feasible. However, the results also suggest that making use of validation data further improves accuracy. Average errors are between another 29 and 83 percent lower for the adjusted conditional distribution than for the corrections in column 5 and 6.[32] The relatively large error of the SMC method for MSAs further underlines the importance of caution when correcting multivariate statistics. It is likely that the larger errors for MSAs stem from differences between rural and urban areas that are difficult to correct without additional information.

In summary, both the evidence from extrapolating within NY and from comparisons to administrative aggregates suggest that extrapolation across geography works well for SNAP. While there is no way to assess the accuracy of the substantive findings in the remainder of the paper, this suggests that the results from the corrections are more reliable than those from the survey data and that among the corrections, the results from the adjusted conditional distribution are likely to be the most accurate ones.

The upper part of table 5 shows large and important differences between the survey reports and the corrections. The differences between the corrections, particularly the conditional distribution method and the SMC method, are much smaller. The poverty rate among recipients according to the SMC method and both versions of the conditional distribution method differ slightly. However, at somewhat above 40 percent, they agree that SNAP serves a less poor population than the reports suggest. While TRIM improved the geographic distribution of SNAP spending, the poverty rate among recipients according to TRIM is similar to the survey results. Given that both TRIM and the survey data overstated the poverty

[32]This does not include state totals, because they are adjusted to match the aggregate numbers.

Table 5: Extrapolation to the Entire U.S., 2010

	(1)	(2)	(3)	(4)	(5)	(6)
	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	-
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	-

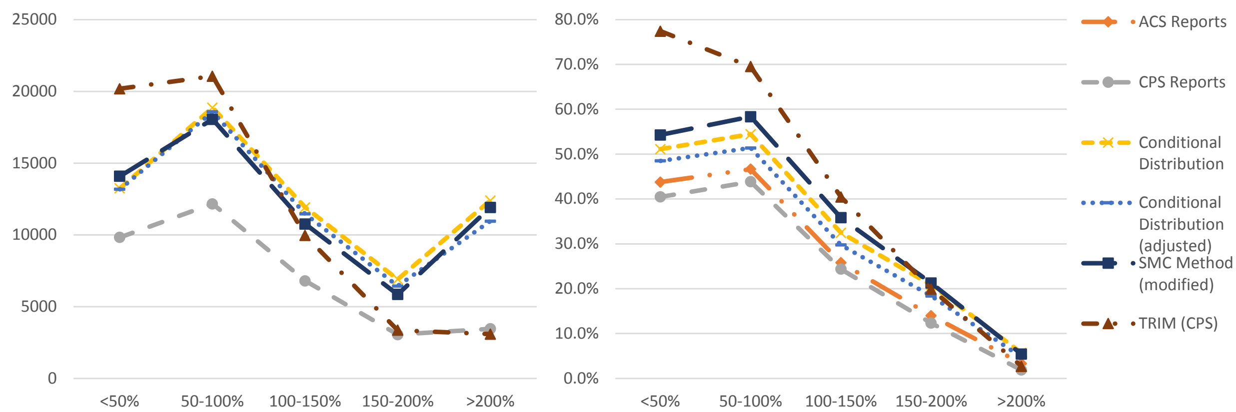
Notes: Column 1 and 2 contain ACS and CPS survey reports. Column 3 uses the 2010 U.S. ACS PUMS data and the conditional distribution from NY, column 4 adjusts the estimated distribution to make total amounts and recipient households match administrative totals in expectation in each state. Column 5 uses the same method as column 7 in table 1 applied state by state, i.e. using state specific models and adjusting to state totals. Column 6 uses CPS microsimulations from the TRIM model of the Urban Institute. Mean absolute deviations are the average absolute value of the difference between administrative totals from the BEA and the method in the respective column. All analyses conducted using household weights.

rate among SNAP recipients in NY, one may find the numbers in column 3-5 that suggest that less than 50 percent of SNAP recipients have annual income below the poverty line more credible. All corrections agree that the poverty rate including SNAP benefits is between 2.1 and 2.4 percentage points. This would mean that the CPS reports understate the poverty reduction by about one percentage point or about 75 percent relative to the 1.3 percentage points suggested by the reports. This would make the bias in estimates of poverty rates from misreporting of SNAP alone similar in magnitude to bias from non-response to the income question (Hokayem, Bollinger and Ziliak, 2015). The difference is larger than the understatement of the poverty reduction that Meyer and Mittag (2015) find for NY, cautiously suggesting that the consequences of underreporting they demonstrate for NY may be even larger at the national level.

Figure 2 shows that there are pronounced differences between the survey estimates and the corrections in how receipt rates and amounts received change with reported annual income relative to the poverty line as well. As in NY, the corrections yield similar results

Figure 2: SNAP by Income Relative to the Poverty Line, NY 2010

Total Amount (in Million $) Fraction of Recipient Households



for univariate statistics, but they diverge more when analyzing the relation between income and program receipt. Yet, the pattern in the differences is similar to the findings for NY: TRIM suggests a much steeper decline of both amounts and receipt rates. This pronounced difference is likely due to substantial overimputation in the left tail and underimputation further up in the distribution of annual incomes. The results using the conditional distribution and the SMC method are substantively similar, particularly for total amounts in the left panel. As expected, the SMC method yields higher receipt rates especially around the poverty line, where reporting is likely to be most accurate. That the differences between the methods are similar in the US and NY, further suggest that the results from the adjusted conditional distribution method are the most accurate estimates. However, the SMC method also delivers plausible results and they agree on several substantive findings. Both suggest that take-up in poverty, particularly in deep poverty, is higher. Receipt rates are higher throughout, but decline faster with income than the survey reports suggest. On the other hand, total amounts spent are non-monotonic according to both methods and higher further up in the distribution of annual incomes. Thus, overall, SNAP receipt seems to be spread out more throughout the distribution of annual incomes than according to the survey reports.

Overall, the survey-based corrections seem to improve over the survey reports, indicating that correcting for underreporting of government transfers can sharply improve estimates even when no validation data are available. However, extrapolation based on the conditional distribution can yield more accurate estimates. Thus, if validation data are available for a similar population, extrapolating from them can further improve accuracy. Substantively, the results provide evidence that survey data severely understate the poverty reduction of transfer programs and that SNAP receipt is spread more through the income distribution. This provides some evidence that the biases in common analyses due to misreporting, that Meyer and Mittag (2015) find for NY, may be even larger at the national level.

7 Conclusion

Recent validation studies establish that the substantial misreporting in survey data severely affects what we learn about government benefits. Most analyses based on survey data do not address the problem and those that attempt to correct for misreporting provide little evidence that the corrections improve their estimates. This paper uses validation data to address both substantive and methodological questions that this raises.

Substantively, the linked administrative data from NY show that while the ACS is more accurate than the CPS, using the ACS instead does not solve the problem of misreporting. With 26 percent of true recipients failing to report receipt in NY, the ACS reports still misrepresent the impact of SNAP. The program seems to be targeted at a poorer population according to the survey data. The reports misrepresent the non-monotonic relation between receipt and income relative to the poverty line as well, missing how much receipt is spread out across the distribution of annual incomes. If we had survey data only, we would also underestimate the importance of SNAP for households in deep poverty and by how much it reduces poverty. The linked data only cover one state, but section 6 suggests similar or larger differences at the national level. For example, correcting for misreporting increases the estimated poverty reduction by about one percentage point or 75 percent of the impact measured by the survey data. The results suggest that the sizable biases in common analyses of low income populations Meyer and Mittag (2015) find for NY may well be even larger in studies of the entire US. This underlines the importance of more accurate sources of information and methods to address the problem of misreporting in survey data.

From a methodological perspective, this study shows that data linkage can play a key role in addressing the problem of survey errors in several ways. First, it confirms that misreporting is severe and that it is systematically related to important variables such as income, which leads to non-classical measurement error.[33] This makes estimation from the contaminated survey data difficult, because it violates the assumptions of common corrections for measurement error. Second, I evaluate whether such corrections improve or just change estimates, using the validation data as a measure of truth. I show that even corrections that mainly rely on survey data can improve over the survey reports. Corrections can also increase bias, but the results provide guidance in choosing a method. TRIM improves receipt rates, amounts and their geographic distribution, but sharply overcorrects among households below the poverty line. The comparison between the original and the modified SMC method shows that assigning receipt probabilistically is important and points to the likely direction of the bias in studies that assign it deterministically. The modified SMC method substantially improves over the survey data throughout and over TRIM for most analyses. These substantial improvements for descriptive statistics such as receipt rates are encouraging for policy analyses that often rely on such estimates. However, corrections that incorporate little or no information besides the original survey data are likely to amplify the bias in multivariate analyses with variables that are related to misreporting, such as income. Thus, researchers interested in causal inference should exercise more caution when applying survey-based corrections in multivariate analyses.

[33] The patterns of misreporting in this study also raise interesting questions about the nature, consequences and reasons for misreporting, which are beyond the scope of this paper. We examine some of them in Celhay, Meyer and Mittag (2016*a*,*b*).

Finally, I describe a method to combine information from validation data with public use data and show that it can further improve estimate accuracy. It can make validation studies more useful by mitigating two of their key problems. First, access to validation data is often restricted. The conditional distribution method can solve the problem of access by offering a simple way to sharply improve estimates from public use data. Compared to other correction methods, it particularly improves multivariate analyses. Second, validation data are often only available for a convenience sample, such as one state, rather than the population of interest. If the reporting structure is similar in the validated sample and the population of interest, we can still use the validation data to learn more about the population of interest. I provide evidence that the conditional distribution method can improve this problem by showing that using it to extrapolate across time and geography clearly improves upon survey reports. For example, it reduces the estimation error in the geographic distribution of receipt and amounts spent, often by a factor of 5 or more and almost 10 fold in some cases.

More generally, the findings in this paper further underline that researchers need to take measurement error into account when using survey data. Misreporting and its consequences are severe for government benefits in all major household surveys. Similar problems are likely to affect studies that use other variables known to suffer from non-classical measurement error, such as income or education. Consequently, improving corrections for measurement error is important. This study suggests that corrections similar to those used in the literature on government programs may also improve over survey reports for other variables. Validation data could be constructed for other variables by record linkage as well, for example by linking tax records to survey data. Even if the validation data are not accessible for reasons of confidentiality or only available for another year or geographic area, the correction based on validation data applied in this paper may still improve the results. If the extent of geographic heterogeneity in (mis-)reporting is sufficiently low for other variables as well, they could allow us to learn more from the linked data about the entire US despite being geographically restricted. This is particularly important for multivariate models, where survey-based corrections are less likely to work well. Whether the findings in this study apply to other variables is an open question, but they suggest that validation data and approaches like the conditional distribution method have the potential to mitigate the pervasive problem of misreporting.

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): 1451–1467.

Abowd, John M., M. Stinson, and G. Benedetto. 2006. "Final report to the social security administration on the SIPP/SSA/IRS public use file project." U.S. Census Bureau, Longitudinal Employer-Household Dynamics Program.

Almada, Lorenzo, Ian McCarthy, and Rusty Tchernis. 2016. "What Can We Learn about the Effects of Food Stamps on Obesity in the Presence of Misreporting?" *American Journal of Agricultural Economics*, 98(4): 997–1017.

Andrews, Donald W. K. 1997. "A Conditional Kolmogorov Test." *Econometrica*, 65(5): 1097–1128.

Ben-Shalom, Yonatan, Robert A. 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.* , ed. Philip Jefferson, Chapter 22, 709–749. Oxford:Oxford University Press.

Black, Dan A., Seth Sanders, and Lowell Taylor. 2003. "Measurement of Higher Education in the Census and Current Population Survey." *Journal of the American Statistical Association*, 98: 545–554.

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 Policy Report.

Bollinger, Christopher R. 1998. "Measurement Error in the Current Population Survey: A Nonparametric Look." *Journal of Labor Economics*, 16(3): 576–594.

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): 827–835.

Bollinger, Christopher R., and Martin H. David. 2005. "I didn't tell, and I won't tell: dynamic response error in the SIPP." *Journal of Applied Econometrics*, 20(4): 563–569.

Bonhomme, Stephane, and Jean-Marc Robin. 2010. "Generalized Non-parametric Deconvolution with an Application to Earnings Dynamics." *Review of Economic Studies*, 77(2): 491–533.

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, , ed. James J. Heckman and Edward Leamer, Chapter 59, 3705–3843. Amsterdam:Elsevier.

Brownstone, David, and Robert G. Valletta. 1996. "Modeling Earnings Measurement Error: A Multiple Imputation Approach." *The Review of Economics and Statistics*, 78(4): 705–717.

Butler, J. S., and Jennie E. Raymond. 1996. "The Effect of the Food Stamp Program on Nutrient Intake." *Economic Inquiry*, 34(4): 781–798.

Celhay, Pablo, Bruce D. Meyer, and Nikolas Mittag. 2016*a*. "Errors in Program Reporting and its Implications." Unpublished Manuscript.

Celhay, Pablo, Bruce D. Meyer, and Nikolas Mittag. 2016*b*. "What Leads to Measurement Errors? Evidence from Reports of Program Participation from Multiple Surveys." Unpublished Manuscript.

Cerf Harris, Benjamin. 2014. "Within and Across County Variation in SNAP Misreporting: Evidence from Linked ACS and Administrative Records." U.S. Census Bureau CARRA Working Paper No. 2014-05.

Chen, Xiaohong. 2007. "Large Sample Sieve Estimation of Semi-Nonparametric Models." In *Handbook of Econometrics*. Vol. 6b, , ed. James J. Heckman and Edward Leamer, Chapter 76, 5549–5632. Amsterdam:Elsevier.

Chen, Xiaohong, Han Hong, and Elie Tamer. 2005. "Measurement Error Models with Auxiliary Data." *The Review of Economic Studies*, 72(2): 343–366.

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.

Currie, Janet. 2006. "The Take-up of Social Benefits." In *Public Policy and the Income Distribution*. , ed. Alan J. Auterbach, David Card and John M. Quigley, Chapter 3, 80–148. New York, NY:Russell Sage Foundation.

Currie, Janet, and Jeffrey Grogger. 2001. "Explaining Recent Declines in Food Stamp Program Participation." *Brookings-Wharton Papers on Urban Affairs*, 203–244.

Dahl, Molly, 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): 750–774.

Davern, Michael, Bruce D. Meyer, and Nikolas Mittag. 2016. "Creating Improved Survey Data Products Using Linked Administrative-Survey Data." Federal Committe on Statistical Methodology Research Conference Proceedings.

Davern, Michael, Jacob A. Klerman, Jeanette Ziegenfuss, Victoria Lynch, and George Greenberg. 2009. "A partially corrected estimate of medicaid enrollment and uninsurance: Results from an imputational model developed off linked survey and administrative data." *Journal of Economic & Social Measurement*, 34(4): 219–240.

Davern, Michael, Kathleen Thiede Call, Jeanette Ziegenfuss, Gestur Davidson, Timothy J Beebe, and Lynn Blewett. 2008. "Validating health insurance coverage survey estimates a comparison of self-reported coverage and administrative data records." *Public Opinion Quarterly*, 72(2): 241–259.

D'Onofrio, Christine, John Krampner, Danny Silitonga, Jihyun Shin, and Vicky Virgin. 2015. "The CEO Poverty Measure, 2005-2013." Center for Economic Opportunity Policy Report.

Drechsler, Jörg, and Jerome P. Reiter. 2010. "Sampling With Synthesis: A New Approach for Releasing Public Use Census Microdata." *Journal of the American Statistical Association*, 105(492): 1347–1357.

Fraker, Thomas, and Robert Moffitt. 1988. "The effect of food stamps on labor supply: A bivariate selection model." *Journal of Public Economics*, 35(1): 25–56.

Giannarelli, Linda, Laura Wheaton, and Joyce Morton. 2015. "How Much Could Policy Changes Reduce Poverty in New York City?" The Urban Institute Poverty, Vulnerability, and the Safety Net Research Report.

Gundersen, Craig, and Brent Kreider. 2008. "Food Stamps and Food Insecurity What Can Be Learned in the Presence of Nonclassical Measurement Error?" *Journal of Human Resources*, 43(2): 352–382.

Hagstrom, Paul A. 1996. "The Food Stamp Participation and Labor Supply of Married Couples: An Empirical Analysis of Joint Decisions." *The Journal of Human Resources*, 31(2): 383–403.

Haider, Steven J., Alison Jacknowitz, and Robert F. Schoeni. 2003. "Food Stamps and the Elderly: Why Is Participation so Low?" *The Journal of Human Resources*, 38: 1080–1111.

Hausman, Jerry A., Jason Abrevaya, and Fiona M. Scott-Morton. 1998. "Misclassification of the dependent variable in a discrete-response setting." *Journal of Econometrics*, 87(2): 239–269.

Hausman, Jerry A., Whitney K. Newey, Hidehiko Ichimura, and James L. Powell. 1991. "Identification and estimation of polynomial errors-in-variables models." *Journal of Econometrics*, 50(3): 273–295.

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.

Hisnanick, John J., Tracy Loveless, and John Chesnut. 2007. "Evaluation Report Covering Receipt of Food Stamps." U.S. Census Bureau American Community Survey Content Test Report H.6, Washington, D.C.

Hokayem, Charles, Christopher Bollinger, and James P Ziliak. 2015. "The Role of CPS Nonresponse in the Measurement of Poverty." *Journal of the American Statistical Association*, 110(511): 935–945.

Hong, Han, and Elie Tamer. 2003. "A simple estimator for nonlinear error in variable models." *Journal of Econometrics*, 117(1): 1–19.

Horowitz, Joel L, and Charles F Manski. 1995. "Identification and Robustness with Contaminated and Corrupted Data." *Econometrica*, 63(2): 281–302.

Hoynes, Hilary W., and Diane W. Schanzenbach. 2009. "Consumption Responses to In-Kind Transfers: Evidence from the Introduction of the Food Stamp Program." *American Economic Journal: Applied Economics*, 1(4): 109–39.

Hsiao, Cheng. 1989. "Consistent estimation for some nonlinear errors-in-variables models." *Journal of Econometrics*, 41(1): 159–185.

Hu, Yingyao, and Susanne M. Schennach. 2008. "Instrumental Variable Treatment of Nonclassical Measurement Error Models." *Econometrica*, 76(1): 195–216.

Ichimura, Hidehiko, and Elena Martinez-Sanchis. 2009. "Estimation and Inference of Models with Incomplete Data by Combining Two Data Sets." Unpublished Manuscript.

Kapteyn, Arie, and Jelmer Y. Ypma. 2007. "Measurement Error and Misclassification: A Comparison of Survey and Administrative Data." *Journal of Labor Economics*, 25(3): 513–551.

Keane, Michael, and Robert Moffitt. 1998. "A Structural Model of Multiple Welfare Program Participation and Labor Supply." *International Economic Review*, 39(3): 553–589.

Lee, Lung-fei, and Jungsywan H. Sepanski. 1995. "Estimation of Linear and Nonlinear Errors-in-Variables Models Using Validation Data." *Journal of the American Statistical Association*, 90(429): 130–140.

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.

Li, Tong. 2000. "Estimation of nonlinear errors-in-variables models: a simulated minimum distance estimator." *Statistics & probability letters*, 47(3): 243–248.

Li, Tong. 2002. "Robust and consistent estimation of nonlinear errors-in-variables models." *Journal of Econometrics*, 110(1): 1–26.

Little, Roderick J.A. 1993. "Statistical analysis of masked data." *Journal of Official Statistics*, 9: 407–407.

Little, Roderick J. A. 2012. "Calibrated Bayes, an Alternative Inferential Paradigm for Official Statistics." *Journal of Official Statistics*, 28(3): 309–334.

Lynch, Victoria, Dean M. Resnick, Jane Stavely, and Cynthia M. Taeuber. 2007. "Differences in Estimates of Public Assistance Recipiency Between Surveys and Administrative Records." U.S. Census Bureau, Washington, D.C.

Marquis, Kent H., and Jeffrey C. Moore. 1990. "Measurement Errors in SIPP Program Reports." In *Proceedings of the 1990 Annual Research Conference*. 721–745. Washington, D.C.:U.S. Bureau of the Census.

McFadden, Daniel. 1989. "A method of simulated moments for estimation of discrete response models without numerical integration." *Econometrica*, 995–1026.

Meyer, Bruce D., and James X. Sullivan. 2006. "Consumption, Income, and Material Well-Being After Welfare Reform." NBER Working Paper 11976, Cambridge, MA.

Meyer, Bruce D., and Nikolas Mittag. 2014. "Misclassification in Binary Choice Models." NBER Working Paper 17042, Cambridge, MA.

Meyer, Bruce D., and Nikolas Mittag. 2015. "Using Linked Survey and Administrative Data to Better Measure Income: Implications for Poverty, Program Effectiveness and Holes in the Safety Net." NBER Working Paper 21676, Cambridge, MA.

Meyer, Bruce D., Robert Goerge, and Nikolas Mittag. 2016. "Errors in Survey Reporting and Imputation and Their Effects on Estimates of Food Stamp Program Participation." Unpublished Manuscript.

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

Meyer, Bruce D., Wallace K. C. Mok, and James X. Sullivan. 2015*b*. "The Under-Reporting of Transfers in Household Surveys: Its Nature and Consequences." Harris School of Public Policy Working Paper, Chicago, IL.

Moffitt, Robert, and J. Karl Scholz. 2010. "Trends in the Level and Distribution of Income Support." *Tax Policy and the Economy*, 24: 111–152.

Newey, Whitney K., and Daniel L. McFadden. 1994. "Large Sample Estimation and Hypothesis Testing." In *Handbook of Econometrics*. Vol. 4, , ed. Robert F. Engle and Daniel L. McFadden, Chapter 36, 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." USDA Economic Research Service Working Paper ERR-156.

Nguimkeu, Pierre, Augustine Denteh, and Rusty Tchernis. 2016. "On the Estimation of Treatment Effects with Endogenous Misreporting." Unpublished Manuscript, Georgia State University.

NORC. 2011. "Assessment of the US Census Bureau's Person Identification Validation System." NORC at the University of Chicago Final Report presented to the US Census Bureau.

Poterba, James M., and Lawrence H. Summers. 1986. "Reporting Errors and Labor Market Dynamics." *Econometrica*, 54(6): 1319–1338.

Rothe, Christoph, and Dominik Wied. 2013. "Misspecification Testing in a Class of Conditional Distributional Models." *Journal of the American Statistical Association*, 108(501): 314–324.

Rubin, Donald B. 1987. *Multiple imputation for nonresponse in surveys.* New York:Wiley.

Rubin, Donald B. 1993. "Discussion: Statistical Disclosure Limitation." *Journal of Official Statistics*, 9(2): 461–468.

Rubin, Donald B. 1996. "Multiple Imputation After 18+ Years." *Journal of the American Statistical Association*, 91(434): 473–489.

Schafer, Joseph L., and Nathaniel Schenker. 2000. "Inference with Imputed Conditional Means." *Journal of the American Statistical Association*, 95(449): 144–154.

Schenker, Nathaniel, Trivellore E. Raghunathan, and Irina Bondarenko. 2010. "Improving on analyses of self-reported data in a large-scale health survey by using information from an examination-based survey." *Statistics in Medicine*, 29(5): 533–545.

Schennach, Susanne M. 2004. "Estimation of Nonlinear Models with Measurement Error." *Econometrica*, 72(1): 33–75.

Scherpf, Erik, Constance Newman, and Mark Prell. 2014. "Targeting of Supplemental Nutrition Assistance Program Benefits: Evidence from ACS and NY SNAP Administrative Records." USDA Working Paper.

Scholz, John K., Robert Moffitt, and Benjamin Cowan. 2009. "Trends in income support." In *Changing Poverty, Changing Policies.* , ed. Maria Cancian and Sheldon Danziger, Chapter 8, 203–241. New York, NY:Russell Sage Foundation.

Sepanski, Jungsywan H., and Raymond J. Carroll. 1993. "Semiparametric quasi-likelihood and variance function estimation in measurement error models." *Journal of Econometrics*, 58(1-2): 223–256.

Sherman, Arloc. 2009. "Safety Net Effective at Fighting Poverty But Has Weakened for the Very Poorest." Center on Budget and Policy Priorities Policy Report, Washington, D.C.

Sherman, Arloc. 2011. "Poverty and Financial Distress Would Have Been Substantially Worse in 2010 Without Government Action, New Census Data Show." Center on Budget and Policy Priorities Policy Report, Washington, D.C.

Sherman, Arloc. 2015. "Safety Net More Effective Against Poverty Than Previously Thought." Center on Budget and Policy Priorities Policy Report, Washington, D.C.

Steuerle-Schofield, Lynne, Brian Junker, Lowell J Taylor, and Dan A Black. 2015. "Predictive inference using latent variables with covariates." *Psychometrika*, 80(3): 727–747.

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." U.S. Census Bureau Working Paper.

U.S. Government Accountability Office. 2015. "Federal Low-Income Programs: Multiple Programs Target Diverse Populations and Needs." U.S. Government Accountability Office Report to Congressional Requesters 15-516, Washington, D.C.

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." U.S. Census Bureau Center for Administrative Records Research and Applications Working Paper.

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." The Urban Institute Research Report.

Appendix

Figure A1: Map of Counties in East (dark) and West (light) NY Sample

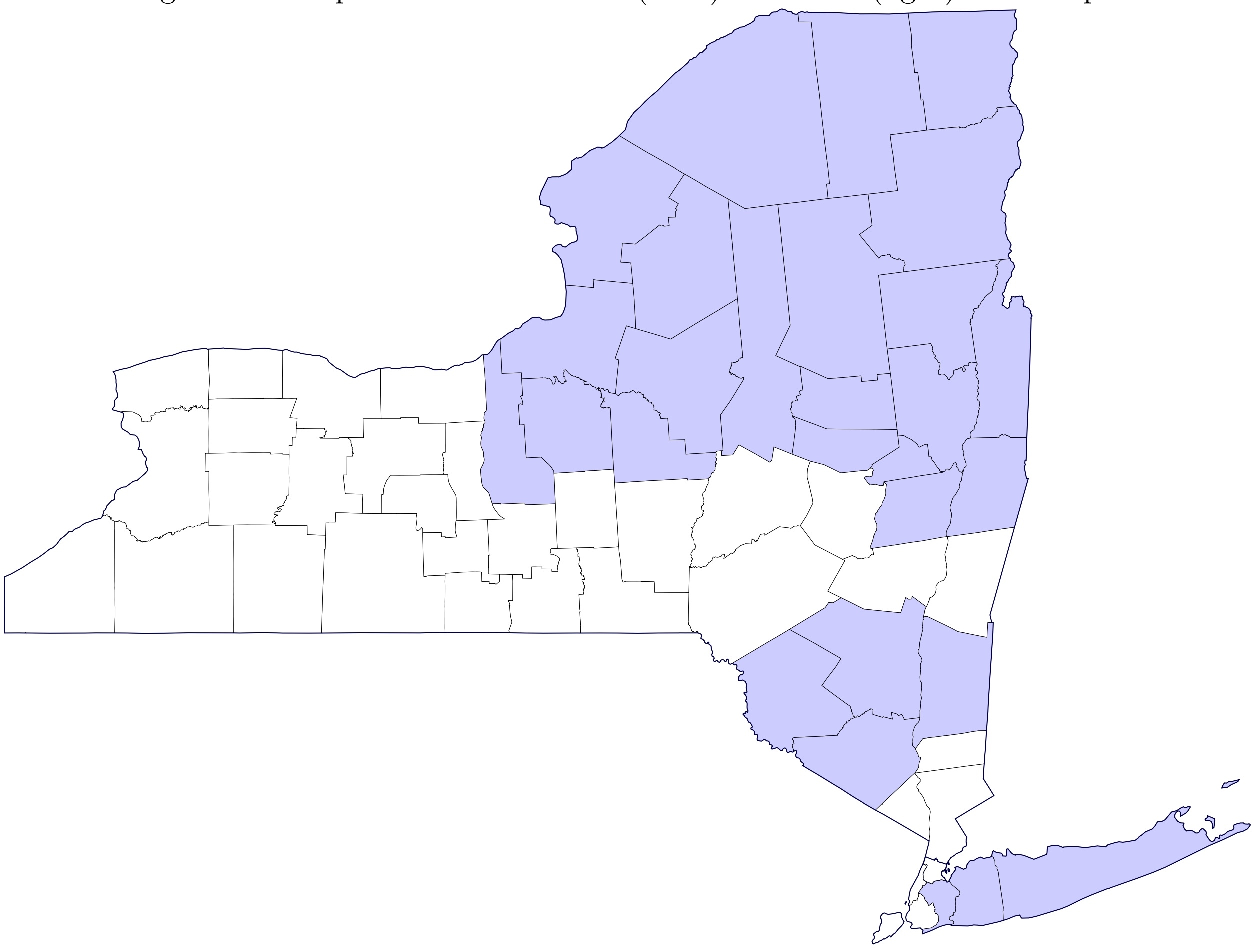


Table A1: The Determinants of a Household having a PIK, Probit Coefficients

	(1)	(2)	(3)	(4)
	2009		2010	
	Coef.	*SE*	*Coef.*	*SE*
Income/Poverty Line	0.015	0.002	0.010	0.002
Age 18-29	-0.128	0.029	-0.023	0.032
Age 30-39	-0.098	0.026	-0.095	0.027
Age 50-59	0.143	0.026	0.089	0.027
Age 60-69	0.275	0.030	0.159	0.031
Age 70 or Older	0.360	0.032	0.218	0.034
Number of Persons in HH	0.050	0.010	0.070	0.011
Number of Persons Under 18	0.019	0.018	-0.017	0.019
Not Married, no Children	-0.133	0.035	-0.103	0.036
Not Married, Children	0.194	0.036	0.045	0.036
Married, no Children	-0.045	0.037	-0.077	0.038
Unemployed	0.030	0.038	0.087	0.040
Not in Labor Force	-0.104	0.023	-0.071	0.024
Not a U.S. Citizen	-0.511	0.024	-0.325	0.027
Hispanic	0.038	0.024	0.060	0.026
White	0.103	0.018	0.071	0.019
Less Than High School	-0.089	0.027	-0.081	0.029
High School Degree	-0.164	0.021	-0.124	0.022
College or More	0.086	0.022	0.060	0.023
Disability	0.070	0.045	0.075	0.047
Disabled, not Working	0.061	0.052	0.056	0.055
Speaks Poor English	-0.273	0.034	-0.180	0.036
Speaks no English	-0.387	0.053	-0.333	0.055
Public Assistance (Reported)	0.065	0.046	0.065	0.049
Constant	1.290	0.046	1.367	0.049
Number of Observations	107,237		106,655	

Notes: All analyses conducted using household weights. Individual characteristics refer to the household head. The omitted family type is married with children, the omitted age category is 40-49 and the omitted education category is some college.

Table A2: Summary Statistics Linked Data

	Linked Data 2009		PIKed NY Sample 2010		NY Subsample 2010	
Variable	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Admin. SNAP Receipt	0.156	0.363	0.179	0.383	0.174	0.379
Admin. SNAP Amount	435.7	1,320.9	601.9	1,662.7	575.1	1,592.1
NY Subsample	0.487	0.500	0.487	0.500	0.000	0.000
SNAP Receipt Reported	0.128	0.334	0.141	0.348	0.139	0.346
SNAP Receipt Imputed	0.009	0.093	0.013	0.113	0.014	0.116
Income/Poverty Line	5.042	6.715	4.809	6.249	4.494	4.922
Age 18-29	0.095	0.293	0.102	0.302	0.087	0.281
Age 30-39	0.170	0.376	0.168	0.374	0.163	0.369
Age 50-59	0.209	0.407	0.210	0.407	0.217	0.412
Age 60-69	0.150	0.357	0.153	0.360	0.154	0.361
Age 70 or Older	0.157	0.364	0.158	0.364	0.160	0.366
Any Income From Capital	0.246	0.431	0.222	0.416	0.215	0.410
# of Persons in HH	2.512	1.507	2.544	1.538	2.662	1.563
# of Children in HH	0.597	1.036	0.596	1.037	0.642	1.047
Not Married, no Children	0.455	0.498	0.455	0.498	0.417	0.493
Not Married, Children	0.098	0.297	0.099	0.298	0.105	0.306
Married, no Children	0.254	0.435	0.255	0.436	0.269	0.444
Linguistic Isolation	0.078	0.268	0.078	0.269	0.080	0.271
# of Persons Employed	1.715	1.479	1.710	1.486	1.792	1.523
Anyone in HH Employed	0.770	0.421	0.766	0.423	0.772	0.420
Elderly or Disabled in HH	0.443	0.497	0.442	0.497	0.459	0.498
Single Household	0.294	0.456	0.288	0.453	0.258	0.438
Unemployed	0.045	0.207	0.051	0.219	0.052	0.223
Not in Labor Force	0.317	0.465	0.322	0.467	0.323	0.468
Female	0.501	0.500	0.505	0.500	0.496	0.500
Not a U.S. Citizen	0.092	0.289	0.096	0.294	0.099	0.298
White	0.726	0.446	0.713	0.452	0.706	0.455
Less than High School	0.135	0.342	0.133	0.340	0.138	0.344
High School Degree	0.256	0.437	0.254	0.435	0.270	0.444
College or More	0.347	0.476	0.349	0.477	0.313	0.464
Disabled	0.154	0.361	0.148	0.355	0.151	0.358
Disabled, not Working	0.120	0.325	0.118	0.323	0.120	0.325
Speaks English Poorly	0.015	0.122	0.015	0.123	0.015	0.121
Speaks No English	0.050	0.219	0.051	0.220	0.054	0.226
Number of Observations		101,335		101,683		49,577

Notes: Individual characteristics refer to the household head. All statistics at the household level using household weights adjusted for the probability of having a PIK.

Table A3: Summary Statistics ACS Public Use Data, 2010

	NY		NY Subsample		US	
Variable	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
NY Subsample	0.489	0.500				
SNAP Receipt Reported	0.138	0.345	0.135	0.342	0.119	0.323
SNAP Receipt Imputed	0.016	0.125	0.016	0.127	0.013	0.114
Income/Poverty Line	4.858	5.703	4.572	4.688	4.251	4.380
Age 18-29	0.102	0.302	0.088	0.283	0.115	0.319
Age 30-39	0.169	0.375	0.164	0.371	0.172	0.377
Age 50-59	0.210	0.407	0.217	0.412	0.205	0.404
Age 60-69	0.152	0.359	0.153	0.360	0.152	0.359
Age 70 or Older	0.157	0.364	0.159	0.366	0.152	0.359
Any Income From Capital	0.221	0.415	0.213	0.409	0.218	0.413
# of Persons in HH	2.539	1.538	2.655	1.561	2.520	1.483
# of Children in HH	0.595	1.037	0.640	1.048	0.624	1.060
Not Married, no Children	0.455	0.498	0.419	0.493	0.417	0.493
Not Married, Children	0.099	0.298	0.106	0.308	0.097	0.296
Married, no Children	0.255	0.436	0.268	0.443	0.286	0.452
Linguistic Isolation	0.083	0.275	0.083	0.276	0.046	0.210
# of Persons Employed	1.704	1.485	1.787	1.516	1.714	1.482
Anyone in HH Employed	0.765	0.424	0.773	0.419	0.767	0.423
Elderly or disabled in HH	0.441	0.497	0.457	0.498	0.444	0.497
Single Household	0.291	0.454	0.260	0.439	0.274	0.446
Unemployed	0.051	0.219	0.052	0.222	0.055	0.228
Not in Labor Force	0.323	0.468	0.322	0.467	0.314	0.464
Female	0.505	0.500	0.494	0.500	0.470	0.499
Not a U.S. Citizen	0.096	0.295	0.100	0.300	0.065	0.246
White	0.707	0.455	0.700	0.458	0.782	0.413
Less Than High School	0.135	0.342	0.140	0.347	0.125	0.331
High School Degree	0.252	0.434	0.265	0.441	0.262	0.440
College or More	0.348	0.476	0.314	0.464	0.303	0.460
Disabled	0.149	0.356	0.152	0.359	0.164	0.370
Disabled, not Working	0.120	0.325	0.122	0.327	0.126	0.332
Speaks English Poorly	0.052	0.222	0.055	0.229	0.030	0.171
Speaks no English	0.015	0.123	0.014	0.119	0.010	0.101
Number of Observations	74,105		36,192		1,203,777	

Notes: Individual characteristics refer to the household head. All statistics at the household level using household weights.

Table A4: Summary Statistics CPS, 2010

	NY		U.S.	
Variable	*Mean*	*SD*	*Mean*	*SD*
SNAP Receipt Reported	0.116	0.320	0.100	0.300
Reported SNAP Amount	388.1	1,470.6	298.7	1,225.5
Income/Poverty Line	4.351	4.763	4.146	4.403
Number of Observations	6,689		151,368	

Notes: Years refer to the ACS survey years, i.e. 2010 pools the 2010 and 2011 CPS ASEC. All statistics at the household level using household weights adjusted for pooling years.

Table A5: Parameter Estimates of the Conditional Distribution, NY, 2009 and 2010

	2009				2010			
	Mass Point		*Amounts*		*Mass Point*		*Amounts*	
	Coef.	*SE*	*Coef.*	*SE*	*Coef.*	*SE*	*Coef.*	*SE*
SNAP Receipt Reported	-2.51	0.002	1,156.9	6.9	-2.50	0.002	1,109.8	6.8
SNAP Receipt Imputed	-0.50	0.006	534.5	20.8	-0.42	0.005	386.6	18.4
Income Relative to Poverty Line Intercepts								
50-100 %	-0.17	0.015	376.0	29.2	0.58	0.015	383.4	30.6
100-150 %	-0.24	0.021	-777.9	53.3	-0.29	0.020	-532.7	54.9
150-200 %	-0.38	0.032	427.5	103.3	-0.62	0.029	-1,188.9	97.6
>200 %	0.56	0.006	-1,629.3	15.4	0.47	0.005	-1,299.1	15.4
Income Relative to Poverty Line Slopes								
≤ 50 %	-0.40	0.018	-190.6	34.5	-0.05	0.017	5.6	36.6
50-100 %	0.12	0.018	-1,264.8	35.6	-0.85	0.018	-934.1	37.1
100-150 %	0.29	0.016	-371.9	42.5	0.37	0.016	-359.7	43.6
150-200 %	0.42	0.018	-1,096.7	59.2	0.52	0.017	-82.8	56.0
>200 %	0.02	0.000	-22.8	2.2	0.04	0.000	-30.7	2.1
Age 18-29	-0.08	0.003	242.5	9.0	0.00	0.003	248.1	9.3
Age 30-39	-0.05	0.003	242.4	8.0	-0.14	0.003	38.6	8.2
Age 50-59	0.01	0.003	-42.4	8.7	0.04	0.003	-210.1	9.0
Age 60-69	0.12	0.004	-171.3	11.2	0.17	0.004	-132.7	11.6
Age 70 or Older	0.36	0.004	-258.5	11.9	0.34	0.004	-315.9	12.4
Any income from capital	0.33	0.003	-158.0	13.2	0.35	0.003	-594.7	13.9
# of Persons in HH	-0.20	0.001	438.2	3.5	-0.22	0.001	538.4	3.4
# of Children in HH	-0.03	0.002	879.3	4.8	-0.02	0.002	1,047.6	4.7
Not Married, no Children	-0.46	0.004	274.1	10.7	-0.55	0.000	475.5	10.6
Not Married, Children	-0.52	0.003	-89.2	8.7	-0.55	0.003	72.6	8.6
Married, no Children	-0.08	0.004	127.8	12.7	-0.22	0.004	481.7	12.6
Linguistic Isolation	-0.03	0.003	55.9	8.9	-0.05	0.003	-1.9	9.2
# of Persons Employed	0.09	0.002	-300.7	4.4	0.09	0.001	-391.2	4.3
Anyone in HH Employed	0.04	0.004	157.5	10.3	0.02	0.003	147.5	10.5
Elderly or Disabled in HH	-0.19	0.003	-141.4	7.6	-0.14	0.003	-405.9	7.8
Single Household	0.19	0.003	-1,487.1	11.0	0.23	0.003	-1,695.8	11.6
Unemployed	-0.05	0.004	-402.4	10.5	-0.10	0.004	-499.8	10.3
Not in Labor Force	-0.05	0.003	-115.4	8.3	-0.09	0.003	-5.7	8.4
Female	-0.06	0.002	-36.3	6.1	-0.06	0.002	56.2	6.2
Not a U.S. Citizen	0.14	0.003	-313.5	8.1	0.27	0.003	-539.7	8.4
White	0.42	0.002	-183.3	5.4	0.41	0.002	-201.5	5.6
Less Than High School	-0.17	0.003	182.0	7.2	-0.10	0.003	202.6	7.6
High School Degree	-0.13	0.002	72.4	7.1	-0.09	0.002	162.2	7.2
College or More	0.17	0.003	-76.1	10.2	0.20	0.003	46.7	10.3
Disabled	0.04	0.005	57.7	15.0	-0.07	0.005	329.4	16.3
Disabled, not Working	-0.17	0.006	-121.9	15.9	-0.07	0.006	-277.0	17.3
Speaks English Poorly	-0.06	0.004	301.8	9.3	-0.16	0.004	381.2	9.5
Speaks no English	-0.15	0.006	510.2	13.1	0.10	0.006	477.2	14.5
Constant	1.61	0.008	651.7	20.7	1.59	0.007	695.2	20.7
Sigma	2,073	2.3			2,358	2.4		
Left Truncation Point	10	0.002			16	0.003		
Number of Observations	101,335				101,683			

Notes: Parameter estimates of truncated normal conditional distribution. The columns labeled "Mass Point" determine the probability that the household does *not* receive SNAP and can be interpreted like probit coefficients. The columns labeled "Amounts" determine amounts conditional on receipt. Individual characteristics refer to the household head. All analyses use household weights adjusted for PIK-probability.

Table A6: Parameter Estimates of the Conditional Distribution, NY Subsamples 2010

	Eastern Counties				Western Counties			
	Mass Point		*Amounts*		*Mass Point*		*Amounts*	
	Coef.	*SE*	*Coef.*	*SE*	*Coef.*	*SE*	*Coef.*	*SE*
SNAP Receipt Reported	-2.49	0.003	1,145.8	9.7	-2.52	0.003	1,048.3	9.3
SNAP Receipt Imputed	-0.31	0.007	257.9	26.3	-0.55	0.007	516.0	25.0
Income Relative to Poverty Line Intercepts								
50-100 %	0.45	0.021	419.9	43.3	0.66	0.022	379.1	42.4
100-150 %	-0.37	0.028	573.7	80.0	-0.20	0.030	-1,584.3	73.4
150-200 %	-0.51	0.041	-909.5	137.7	-0.73	0.042	-1,000.9	135.1
>200 %	0.32	0.007	-1,177.6	21.9	0.57	0.008	-1364.8	21.3
Income Relative to Poverty Line Slopes								
$\leq$ 50 %	-0.12	0.023	-51.6	50.9	-0.01	0.027	51.8	51.9
50-100 %	-0.77	0.025	-881.3	52.4	-0.89	0.026	-1,003.6	51.5
100-150 %	0.36	0.022	-1,030.8	63.4	0.35	0.023	302.7	58.2
150-200 %	0.42	0.023	-240.4	78.7	0.60	0.024	-169.7	77.7
>200 %	0.05	0.000	-11.0	3.0	0.03	0.001	-40.1	2.9
Age 18-29	-0.05	0.004	-35.4	13.5	0.04	0.005	478.4	12.6
Age 30-39	-0.16	0.004	-24.3	12.1	-0.13	0.004	112.7	10.8
Age 50-59	0.02	0.004	-283.4	13.2	0.07	0.004	-123.6	11.9
Age 60-69	0.15	0.005	-284.7	16.8	0.20	0.005	1.5	15.7
Age 70 or Older	0.33	0.006	-489.7	18.0	0.37	0.006	-148.2	16.8
Any Income From Capital	0.31	0.004	-840.8	20.0	0.39	0.004	-339.3	18.9
# of Persons in HH	-0.23	0.002	409.7	4.9	-0.20	0.002	652.3	4.7
# of Children in HH	0.00	0.003	997.6	7.0	-0.04	0.002	1,056.8	6.3
Not Married, no Children	-0.51	0.005	415.1	15.4	-0.58	0.005	550.0	14.2
Not Married, Children	-0.56	0.005	206.9	12.6	-0.54	0.004	41.1	11.5
Married, no Children	-0.19	0.006	328.2	18.2	-0.24	0.005	621.1	17.1
Linguistic Isolation	-0.03	0.005	-220.8	13.5	-0.07	0.005	181.1	12.3
# of Persons Employed	0.09	0.002	-308.4	6.3	0.08	0.002	-447.1	5.8
Anyone in HH Employed	0.05	0.005	81.2	15.0	-0.02	0.005	189.3	14.5
Elderly or Disabled in HH	-0.15	0.004	-140.2	11.7	-0.13	0.004	-612.5	10.2
Single Household	0.16	0.004	-1,615.7	16.3	0.29	0.004	-1,731.4	16.0
Unemployed	-0.13	0.006	-273.8	15.5	-0.08	0.005	-677.0	13.3
Not in Labor Force	-0.08	0.004	-11.5	12.0	-0.10	0.004	15.3	11.5
Female	0.01	0.003	-66.9	9.0	-0.12	0.003	155.2	8.4
Not a U.S. Citizen	0.18	0.004	-488.4	12.4	0.35	0.004	-534.8	10.9
White	0.35	0.003	-349.1	8.0	0.46	0.003	-82.8	7.6
Less Than High School	-0.04	0.004	289.9	10.9	-0.16	0.004	130.0	10.2
High School Degree	-0.04	0.003	159.0	10.4	-0.13	0.003	178.5	9.7
College or More	0.20	0.004	160.8	14.6	0.20	0.004	-75.5	14.0
Disabled	-0.13	0.007	146.2	24.0	0.00	0.007	528.1	21.4
Disabled, not Working	0.05	0.008	-118.4	25.3	-0.21	0.008	-452.5	22.9
Speaks English Poorly	-0.13	0.005	306.1	14.1	-0.18	0.005	453.1	12.5
Speaks no English	0.23	0.009	495.4	20.7	-0.02	0.009	408.7	20.0
Constant	1.70	0.010	1,013.5	29.5	1.53	0.010	370.9	28.6
Sigma	2,221	3.4			2,442	3.1		
Left Truncation Point	16	0.004			16	0.004		
Number of Observations	49,577				52,106			

Notes: Parameter estimates of truncated normal conditional distribution. The columns labeled "Mass Point" determine the probability that the household does *not* receive SNAP and can be interpreted like probit coefficients. The columns labeled "Amounts" determine amounts conditional on receipt. Individual Characteristics refer to the household head. All analyses use household weights adjusted for PIK-probability.