Downloaded from by guest on October 16, 2026. Copyright 2020

Errors in Survey Reporting and Imputation and Their Effects on Estimates of Food Stamp Program Participation

Bruce D. Meyer
Nikolas Mittag
Robert M. Goerge

ABSTRACT

Accurately measuring government benefit receipt in household surveys is necessary when studying disadvantaged populations and welfare programs. The Food Stamp Program is especially important given its size and recent growth. To validate survey reports, we link administrative data on participation in two states to three key household surveys. We find that between 23 and 50 percent of true food stamp recipient households do not report receipt. A substantial number of true nonrecipients are also recorded as recipients. We examine reasons for these errors, including imputation, an important source of error. Error rates vary with household characteristics, implying complicated biases in multivariate analyses, such as regressions. We directly examine biases in common survey-based estimates of program receipt by comparing them to estimates from our linked data. We find that the survey estimates understate participation among single parents, nonwhites, and low-income households and also lead to errors in multiple program receipt and time and age patterns of receipt.

Bruce D. Meyer is at Harris School of Public Policy, University of Chicago (bdmeyer@uchicago.edu). Nikolas Mittag is at CERGE-EI, joint workplace of Charles University Prague and the Economics Institute of the Academy of Sciences of the Czech Republic (nikolasmittag@posteo.de). Robert M. Goerge is at Chapin Hall, University of Chicago (rgoerge@chapinhall.org). *The authors thank David Johnson, John Kirlin, Gayatri Koolwal, Alan Krueger, Cathleen Li, Wallace Mok, Daniel Schroeder, James Spletzer, Jane*

(continued on next page)

[Submitted August 2018; accepted May 2020]; doi:10.3368/jhr.58.1.0818-9704R2
JEL Classification: C81, D31, I32, and I38
ISSN 0022-166X E-ISSN 1548-8004 © 2022 by the Board of Regents of the University of Wisconsin System
Supplementary materials are freely available online at: http://uwpress.wisc.edu/journals/journals/jhr-supplementary.html
This open access article is distributed under the terms of the CC-BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0) and is freely available online at: http://jhr.uwpress.org
Nikolas Mittag https://orcid.org/0000-0003-3894-0703

Downloaded from by guest on October 16, 2026. Copyright 2020

I. Introduction

Accurately measuring government benefit receipt in household surveys is important to assess the economic circumstances of disadvantaged populations, program take-up, the distributional effects of government programs, and other program effects. The Food Stamp Program (or FSP, now the Supplemental Nutrition Assistance Program, or SNAP) is especially important given its large and growing size and findings of its effects on health, labor supply, food security, consumption, and other outcomes.[1] Recognizing that surveys may have errors, this study examines the misreporting of Food Stamp Program receipt using a new linkage of administrative microdata from two states to three major survey data sets. We argue that our administrative measure of program receipt is sufficiently accurate to study error at the household level, allowing us to examine rates of misreporting and imputation error. We study how survey errors vary with household characteristics to assess their likely determinants and consequences. Program participation is often used as either a dependent or independent variable in multivariate models, such as regression analyses.[2] We examine how survey error affects such estimates of program receipt, which are commonly obtained from the error-prone survey data we evaluate.

There is growing evidence that program receipt is badly reported in household surveys. The most extensive and frequently cited evidence compares weighted totals of dollars or recipients in household surveys to analogous figures provided by government agencies. The most comprehensive research of this form in terms of programs and surveys covered is Meyer, Mok, and Sullivan (2015a,b), which references many earlier studies.[3] It finds net underreporting of program receipt that is substantial, widespread, and steadily growing over time. A common criticism of these aggregate studies is that

Stavely, Shelly Ver Ploeg, Derek Wu, two anonymous referees, and audiences at the American Economic Association Meetings, American Statistical Association Meetings, Baylor University, European Congress of Methodology, ITSEW, Harvard University, USDA, Yale University, and ZEW for beneficial comments. The authors are grateful for the assistance of many Census Bureau employees, including David Johnson, Amy O'Hara, Lynn Riggs, and Frank Limehouse. Lucy Bilaver, Kerry Franzetta, and Janna Johnson provided excellent research assistance. This research was supported by the Economic Research Service of the USDA, the Russell Sage Foundation, Alfred P. Sloan Foundation, Charles Koch Foundation, the Menard Family, the Czech Science Foundation (through grant no. 16-07603Y), and the Czech Academy of Sciences (through institutional support RVO 67985998). Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the USDA or the U.S. Census Bureau. The data analysis was conducted at the Chicago RDC and was screened to avoid revealing confidential data. Due to confidentiality restrictions, the linked data used in this paper can only be accessed at secure facilities of the U.S. Census Bureau by researchers with Special Sworn Status. The authors cannot make the data or intermediate results publicly available, but researchers can request access to the data by writing a proposal to the U.S. Census Bureau. Further information is available at https://www.census.gov/about/adrm/linkage/guidance.html and from the authors.

1. See Hoynes and Schanzenbach (2009); Almond, Hoynes, and Schanzenbach (2011); and Schmidt, Shore-Sheppard, and Watson (2016), for example.
2. Examples where program participation is the main dependent variable include Blank and Ruggles (1996); Haider, Jacknowitz, and Schoeni (2003); Figlio, Gundersen, and Ziliak (2000); Currie and Grogger (2001); and Ziliak, Gundersen, and Figlio (2003), while cases where it is an explanatory variable include Schmidt, Shore-Sheppard, and Watson (2016); Gundersen and Ziliak (2003); and Blundell and Pistaferri (2003).
3. Also see Coder and Scoon-Rogers (1996), Roemer (2000), Wheaton (2007), and Rothbaum (2015).

they identify net underreporting only, so they may understate errors by missing false negative reports (failures to report true receipt) that are cancelled out by false positive reports (incorrect reports of receipt). The results are also potentially biased by frame, nonresponse, and weighting errors. Furthermore, such aggregate studies have a limited ability to examine how survey error varies with interview and respondent characteristics. This limitation hinders their ability to study the determinants of errors, their consequences for substantive studies, and potential corrections.

Linking surveys to administrative microdata provides a potential solution to these limitations. By comparing survey values to true values, linked validation data can allow us to uncover the extent of error, its causes, and consequences. Unfortunately, surveys of the literature have noted that there exist very few "complete record check" studies that use validation data for the entire population. Such studies are needed to assess false positive reports and thus net reporting of receipt (Bound, Brown, and Mathiowetz 2001). The few studies that do complete record checks tend to suffer from small sample sizes and often rely on a single state or survey. In addition, they are rarely able to analyze or correct possible biases that might result from linkage problems.

In this study, we link administrative data on food stamp receipt from two states to three of the most important economic surveys: the Current Population Survey (CPS), the American Community Survey (ACS), and the Survey of Income and Program Participation (SIPP). The CPS is the most used labor economics survey and the source of official income and poverty statistics. The ACS replaced the Census long-form data and is the largest general household survey, allowing fine geographic analyses. The SIPP is the most detailed survey of program receipt and commonly thought to have the highest-quality data. The Social Security Numbers on the food stamp records that we use have been verified (compared to SSA records) as a necessary condition for receipt of benefits, so the accuracy of the linkage is very high. We discuss likely remaining biases due to linkage error.

Our recent related papers use linked data to analyze how survey error biases specific estimates, such as the estimated poverty rate (Meyer and Mittag 2019a) and mean dollar transfers (Meyer and Mittag 2021a), as well as how linked data can be used to improve the accuracy of these estimates in the presence of survey error (Mittag 2019; Davern, Meyer, and Mittag 2019). The exceptional data accuracy and sample size in this study allow us to go beyond aggregate statistics and study survey error at the household level. In contrast to prior papers, this accuracy allows us to analyze errors in whether a household reports program receipt (both over- and underreporting) and what predicts these errors. We then return to aggregate survey estimates by examining how errors at the household level skew estimates of the determinants of program receipt. Some recent and ongoing work already builds on our findings in this study. Meyer and Mittag (2017) derives theoretical results to better understand the patterns of bias in multivariate models we find here and evaluates bias corrections. In Meyer and Mittag (2019b), we directly extend this study by examining two aspects of survey errors—how they vary across interviewers and geography—that we cannot examine with the data we use here. In Celhay, Meyer, and Mittag (2021), we use a larger linked sample from New York to extend the analyses of misreporting in this paper to imputation error and the reporting of multiple programs. Celhay, Meyer, and Mittag (2022) uses the same data to analyze potential causes of the reporting errors we document here. See Meyer and Mittag (2021b) for an overview that puts these studies in the context of the broader literature.

Downloaded from by guest on October 16, 2026. Copyright 2020

We find substantial underreporting of food stamp receipt, with a quarter to half of true recipient households not recorded as such, depending on the survey. A much smaller share of nonrecipients are recorded as receiving food stamps. Since most households are nonrecipients, these false positives can nonetheless have a substantial effect on net reporting. A large share of these false positives, though not a majority, are imputed observations. The large differences across the three surveys in false negative and false positive rates suggest that survey design plays an important role in survey accuracy.

We show that both false negative and false positive reports are associated with a variety of household and interview characteristics, including income and race. From a methodological perspective, this shows that survey errors are not random. Thus, the errors also lead to complicated biases in multivariate analyses that are difficult to correct. In addition, instrumental variable methods will be inconsistent.[4] Our evidence on the determinants of errors also sheds light on theories of misreporting. We briefly examine the role of comprehension, salience, recall, and stigma, which the literature has suggested as causes of misreporting. Since there are few situations where we have independent and accurate measures to evaluate survey quality, this evidence on program receipt should also aid the improvement of household surveys.

Finally, we examine how survey error affects studies of the use of government programs, a large literature that often relies on error-ridden self-reports of program receipt. For example, the surveys we examine were used in several recent studies of Food Stamp Program participation (Blank and Ruggles 1996; Gundersen and Oliveira 2001; Haider, Jacknowitz, and Schoeni 2003; Wu 2010; Ganong and Liebman 2018). Similar binary choice models with reported program participation as the dependent variable are also frequently estimated for other programs; see Bitler, Currie, Scholz (2003) for an example and Currie (2006) for an overview. These estimates likely suffer from bias due to underreporting as well. Models of program receipt are also used to increase take-up and better target programs to the most needy, an issue that has long concerned policymakers (see U.S. General Accounting Office 2004 for efforts to raise food stamp participation). However, as Bound, Brown, and Mathiowetz (2001) note, little work has examined the consequences of program receipt errors for substantive analyses.[5] Despite the large literature on the distributional consequences of welfare and social insurance programs, few studies attempt to correct for misreporting.[6]

Using our data with both error-ridden and true measures of program receipt, we analyze the consequences of this nonclassical measurement error for a prototypical application: binary choice models with program receipt as the dependent variable. Such models are often used to study program take-up, typically showing that participation rates among eligibles are well below one.[7] Given the extent of underreporting, a major part of what appears to be nonparticipation may actually be recipients whose receipt is not recorded in the survey. Our linked data indicate that the survey data understate

4. Nguimkeu, Denteh, and Tchernis (2019) provide a formal treatment showing that the bias with instrumental variable methods can be severe.

5. Notable exceptions include Bollinger and David (1997, 2001), Pierret (2001), and Gundersen and Kreider (2008); see Section II for further discussion and references.

6. See, for example, Wheaton (2007); Scholz, Moffitt, and Cowan (2009); and Meyer (2010) for exceptions.

7. For excellent reviews of research on take-up of food stamps and other programs, see Remler and Glied (2003) and Currie (2006).

Downloaded from by guest on October 16, 2026. Copyright 2020

participation by single parents, nonwhites, and the elderly, as well as the extent to which participation declines as incomes rise. Maybe surprisingly, we also find that the sign of the association of most variables, such as age, education, and family type, is correct.

In the next section, we review the literature on misreporting of government transfers. Section III describes our data sources and linkage. Section IV provides our main evidence on the extent of survey error and discusses likely biases from linkage errors. Section V analyzes how survey error varies with household characteristics. Section VI examines the bias from survey error and how it affects our understanding of program receipt. Section VII offers conclusions.

Downloaded from by guest on October 16, 2026. Copyright 2020

II. Misreporting in Survey Data

Several studies document significant misreporting of transfer program income in survey data. Bound, Brown, and Mathiowetz (2001) and Moore, Stinson, and Welniak (2000) provide reviews of the literature, so we focus our summary on their main conclusions and newer studies. We examine reporting of whether a program was received rather than the amount received. The evidence on reporting of amounts is scant, but there is some evidence that the main determinant of underreported dollars is whether receipt is reported at all (Moore, Marquis, and Bogen 1996; Moore, Stinson, and Welniak 2000; Meyer, Mok, and Sullivan 2015a).

Three main approaches are used to assess the validity of survey reports: comparisons of survey aggregates to administrative totals, partial validation studies, and full validation studies. Comparisons of estimated totals from survey reports to administrative totals show that the survey reports generally fall substantially short of actual program spending. See Meyer, Mok, and Sullivan (2015b) and the many earlier studies that they cite. The rate of net underreporting differs sharply across programs and surveys and has tended to rise over time. Comparing survey totals to official statistics points to severe data quality issues, but this approach leaves many important questions open. If weighting does not correct for undercoverage or nonresponse, the difference between survey and administrative totals estimates the combined bias from misreporting and other sources of survey error. Aggregate comparisons also cannot provide information on the extent to which false negatives are counterbalanced by false positive reports. In addition, they can only provide very limited information about the factors that are associated with survey error. Finally, aggregate data cannot be used to assess bias in applications using multivariate data or to devise and evaluate corrections for the bias in such analyses. Consequently, aggregate studies provide an important indicator of survey problems, but an accurate measure of receipt at the individual or household level is needed to determine the causes and consequences of survey error.

Linking survey and administrative data can provide this accurate measure to validate survey responses.[8] Most early linkage studies are partial record check studies that only

8. We do not mean to argue that administrative or linked data are more accurate in general. Obtaining an accurate measure via data linkage requires high-quality administrative records and linkages, as discussed in Meyer and Mittag (2021b). Administrative data can contain substantial amount of error; see Niehaus and Sukhtankar (2013) for an extreme case. See Courtemanche, Denteh, and Tchernis (2019) and Meyer and Mittag (2019b) for a discussion of a specific linked data source. We discuss likely inaccuracies and their consequences

examine the survey response of known program recipients. Past food stamp validation studies have found substantial rates of false negative reports that differ considerably across studies. For example 20 percent of true recipients are not recorded as such in the 1984 SIPP (Marquis and Moore 1990), and 40 percent are not in the Maryland sample of the 2001 predecessor to the American Community Survey (Taeuber et al. 2004). There are large differences in the false negative rates across these studies. While these studies can provide evidence on false negative rates and characteristics associated with failure to report receipt, they cannot examine false positive reporting. Consequently, they only allow inference about net reporting rates under the assumption that the effect of false positives is negligible. Marquis et al. (1981) and Moore, Stinson, and Welniak (2000) review the findings of this literature. Both reviews document substantial false negative rates for many transfer programs but also argue that the literature overemphasized underreporting because that is what the existing partial record check studies are able to capture.

This line of argument leads both Moore, Stinson, and Welniak (2000) and Bound, Brown, and Mathiowetz (2001) to call for more complete record check studies that validate the reports of both recipients and nonrecipients to examine both types of error. To advance research in this way requires linking the survey to the universe of program recipients, so that not being included in the administrative data confirms not receiving the program. Unfortunately, such linked data are rarely available. If they are, they typically only cover a short time period and a small subset of the survey respondents, such as those from a single state. Yet complete record check studies provide important additional insights about survey error.

The few existing complete validation studies agree on the finding that false positive rates are much lower than false negative rates. However, there is some variation in the rates of false positives across studies. For the Food Stamp Program, false positive rates range from 0.3 percent in Bollinger and David (1997) to 2–3 percent in Moore, Marquis, and Bogen (1996). As there are far more nonrecipients than recipients, even such low rates of false positives lead to high error counts. Early complete record check studies point towards substantial counts of errors in both directions, leading to slight net overreporting. While this result challenges the notion that the net effect of misreporting is to understate total program receipt (Marquis et al. 1981), more recent validation studies tend to find net underreporting of food stamp receipt (Marquis and Moore 1990; Moore, Marquis, and Bogen 1996; Taeuber et al. 2004; Nicholas and Wiseman 2009, 2010; Kirlin and Wiseman 2014; Meyer and Mittag 2019a,b). Given that most studies focus on a single survey or state, it is unclear whether the differences between studies are due to state, survey, or other study-specific factors. Consequently, important questions remain on the sign of the net bias and the extent to which it depends on the type of survey.

Even in the most favorable case of small or no net survey error, the substantial error rates these studies find at the household level are likely to bias analyses of subpopulations and multivariate models, especially if errors are correlated with individual and household characteristics (see, for example, Meyer and Mittag 2019a; Nguimkeu,

in Section IV.C. If the linked data contain substantial error, other methods are required (for example, Abowd and Stinson 2013; Kapteyn and Ypma 2007; Oberski et al. 2017; and Meijer, Rohwedder, and Wansbeek 2012).

Downloaded from by guest on October 16, 2026. Copyright 2020

Denteh, and Tchernis 2019). It is common to assume that errors are independent of other variables in order to provide a simple summary measure of the degree of survey error (for example, Moore, Stinson, and Welniak 2000) or to correct the bias due to survey error (for example, Hausman, Abrevaya, and Scott-Morton 1998). In light of the importance of this assumption, it is surprising that few studies examine whether misreporting is indeed unrelated to other variables. Notable exceptions are Bollinger and David (1997, 2001, 2005), who reject this assumption by showing that reporting of food stamp receipt is related to income, gender, education, and household structure, as well as later survey attrition.

As Bound, Brown, and Mathiowetz (2001) point out in their review, there are only few analytic results on the consequences of such nonclassical measurement error, with the biases often being intractable and model-specific. The literature on identification and bias of treatment effects in the presence of misclassification of the treatment variable provides ample examples of the complex, but often severe consequences of nonclassical measurement error. Kreider (2010) provides an extreme example of the consequences in a specific case. Other recent papers on the estimation of treatment effect with misclassification examine the consequences of nonclassical measurement errors (for example, Millimet 2011; Almada, McCarthy, and Tchernis 2016), as well as partial identification (for example, Gundersen and Kreider 2009; Gundersen, Kreider, and Pepper 2012; Kreider et al. 2012; Jensen, Kreider, and Zhylyevskyy 2019) and point identification (Nguimkeu, Denteh, and Tchernis 2019).

In the absence of formal results, complete record check studies offer a unique opportunity to analyze the biases in specific cases by comparing models relying on a validated variable to those using a survey variable. For example, Bollinger and David (1997, 2001) and Meyer and Mittag (2017) use validation data to analyze the impact of survey error on multivariate models with a misclassified dependent variable. A few recent and ongoing studies use a similar approach to examine program receipt rates (Cerf Harris 2014; Meyer and Mittag 2019b), poverty statistics (Meyer and Mittag 2019a; Nicholas and Wiseman 2009, 2010), and program effects (Kang and Moffitt 2019). This validation data approach makes it possible to examine directly whether survey error can explain surprising empirical findings, such as the low take-up of government programs among the elderly (Haider, Jacknowitz, and Schoeni 2003) and among households in extreme poverty (Tiehen, Jolliffe, and Gundersen 2012).

High error rates also raise the question why people misreport in surveys. In their review of the literature, Sudman and Bradburn (1974) point out the lack of a general theory of reasons for survey errors. Along the same lines, Bound, Brown, and Mathiowetz (2001) note that few fundamental principles have been established in the literature. They divide reasons for misreporting into three areas: cognitive processes, social desirability, and essential survey conditions or survey design.

The cognitive process of answering a question involves comprehension of the question, recalling information from memory, and communicating the result. Cognitive factors may lead to misreporting because of, among others, difficulties understanding questions, difficulties recalling information, and information that is not salient (for a review, see Sudman, Bradburn, and Schwarz 1996). Much of the empirical literature focuses on recall and retrieval problems. The research provides some evidence that a longer recall period leads to more errors, but the evidence is mixed and far from conclusive. For example, Meyer and

Downloaded from by guest on October 16, 2026. Copyright 2020

Mittag (2019b) provide evidence of recall effects for SNAP. Marquis and Moore (1990) find no effect of recall when analyzing the receipt of food stamps and other programs. Bound, Brown, and Mathiowetz (2001) suggest that rather than the mere passage of time, the complexity of the experience over time is related to misreporting. Thus, households with irregular or infrequent receipt should be more likely to fail to report. Complex patterns of receipt may also lead respondents to confuse government programs and fail to report a program they receive, while reporting receipt of another program. Recall periods could affect survey accuracy in more complex ways because respondents often misstate the timing of events. They tend to report events that occurred before or after the reference period as having happened in the reference period. Such "telescoping" of events can lead to both false negative and false positive errors. Another precept in the literature on how cognitive processes affect survey errors is that more salient events are more easily remembered. Sometimes it has been found that high salience can lead to overreporting.

Another important reason for misreporting is social desirability, which refers to a tendency of respondents to report socially desirable answers whether or not they are true. See Bound, Brown, and Mathiowetz (2001) for a comprehensive discussion. The economic literature focuses on the social stigma associated with dependence on government programs as a reason not to report receipt. This idea suggests underreporting among those with higher income and education for whom welfare receipt seems more out of place. We would also expect underreporting due to stigma to be more prevalent among those who may seem less needy, such as the elderly, two-parent families, and the childless. More generally, social desirability may affect respondent cooperativeness, which Bollinger and David (2001) emphasize as a determinant of accurate reporting.

Finally, features of the survey design, such as the survey mode and method, also affect the accuracy of survey data (see Groves 1989 for a review). Survey design may also affect the accuracy of the data in mechanical ways through the coding and editing process. Given the high rates of item nonresponse in some household surveys, the imputation methods employed by the survey can be another important source of error.

Downloaded from by guest on October 16, 2026. Copyright 2020

III. Data and Linkage

We examine three large and frequently used household surveys: the 2001 ACS,[9] the 2002–2005 CPS, and data from January 2001 to April 2005 from the 2001 and 2004 panels of the SIPP. Our administrative records provide information on food stamp receipt for all recipients in Illinois and Maryland. The monthly records report program receipt, amounts (for some years), and Social Security Numbers (SSN). The source of the Illinois data is the Illinois Department of Human Services (IDHS) client database. From this database, Chapin Hall created the Illinois Longitudinal Public Assistance Research Database (ILPARD), a longitudinal database of public assistance cases. The ILPARD is updated monthly with new cases from the IDHS system and records that IDHS has changed in the past month. The Food Stamp Program records for Illinois contain monthly information on program utilization of all members of the

9. Strictly speaking, we used the 2001 Supplementary Survey or SS01, which is a predecessor of the ACS.

Table 1
Reported Food Stamp Receipt and Demographics, 2002–2005 CPS

	U.S.	IL and MD	IL	MD
Reported food stamp receipt rate	5.89%	4.73%	5.50%	2.97%
Reported average received ($)	108.05	94.66	111.87	54.87
Age under 18	0.17%	0.17%	0.18%	0.14%
Age 18–39	33.02%	33.71%	34.79%	31.20%
Age 40–64	46.33%	45.61%	44.42%	48.36%
Education less than high school	15.37%	13.35%	13.81%	12.28%
Education high school	30.43%	30.01%	31.10%	27.49%
Education some college	26.80%	25.09%	26.32%	22.24%
Black	9.18%	13.51%	10.64%	20.14%
Other nonwhite	28.82%	28.30%	28.08%	28.81%
Hispanic	7.89%	5.63%	6.59%	3.42%
Poverty rate	13.40%	11.80%	12.71%	9.67%

Notes: Receipt and poverty rates are in terms of households, demographics in terms of household heads. Average received is unconditional annual dollars per household. For the demographics, the omitted category is "65 and older" for age, "college degree or more" for education, and "white" for ethnicity. All estimates use sampling weights and are based on the entire 2002–2005 CPS sample (and therefore slightly differ from the estimated receipt rates in Table 1).

Downloaded from by guest on October 16, 2026. Copyright 2020

household. The data supplied to the Census Bureau cover calendar years 1998–2004. The source of the Maryland data is the Client Automated Resource and Eligibility System (CARES) of the Maryland Department of Human Resources. The data provided to the Census Bureau cover the period 1998–2003 and include monthly information on all Maryland residents receiving food stamps during that period.

Analyzing two states raises the question whether the results generalize to the population overall. Table 1 compares our sample to the entire United States. Many demographic and economic characteristics are similar, but a larger share of our sample is Black and a smaller share of Hispanic origin. Our sample is more educated and less poor, so reported program receipt is lower than in the entire population. These differences are mainly driven by the Maryland sample. The net dollar reporting rates (total survey dollars divided by total dollars paid according to the Bureau of Economic Analysis) for Illinois and Maryland combined is 52 percent for 2001–2004, which is lower than in the entire United States (59 percent). Meyer and Mittag (2019b) use a different survey to analyze geographic variation in misreporting. They document some variation between states, but of a magnitude that would be unlikely to overturn our qualitative conclusions.

We link the survey and administrative data using the Protected Identification Key (PIK), which is an anonymized version of the SSN. PIKs are assigned to both the administrative data and the survey data separately. The administrative records contain SSN because an individual must have a validated SSN in order to receive food stamps (their name, gender, and date of birth must match SSA records). The food stamp data are subject to regular audits by the USDA. The validated SSN in the administrative data

Downloaded from by guest on October 16, 2026. Copyright 2020

is converted to a PIK by the Census Bureau. A PIK is obtained for 96 percent of the Illinois food stamp records over the entire period and 98 percent of the Maryland records. To obtain PIKs for the survey data, the Census Bureau uses name, address, and date of birth from the sampling frame and survey records to match each individual in the survey to a PIK/SSN in a reference file that contains all transactions recorded against a SSN. See NORC (2011) and Wagner and Layne (2014) for further discussion. PIKs are assigned to individuals, allowing us to link the administrative records to the surveys at the individual level. We then aggregate the linked data to the household level.

The administrative records contain every individual on a program case, so we can link most households in which at least one member is assigned a PIK (see Section IV.C for further details). A PIK is successfully obtained for at least one member of 93 percent of ACS households in Illinois and 95 percent of ACS households in Maryland. The rates are considerably lower for CPS households. Prior to 2005, respondents were asked to supply their SSN in the CPS to allow linking, and a PIK was not determined for those who did not supply an SSN, reducing the share of households that can be linked. We have a PIK for at least one member of 68 percent of Illinois CPS households and 81 percent of Maryland CPS households. The PIK rate is similar in the SIPP, in which 71 percent of all households have a PIK. The rates are slightly lower for those who are likely food stamp recipients in all three surveys. For example, in the ACS the rates are 89 percent in Illinois and 92 percent in Maryland for households with income below twice the federal poverty line.

The main sample for our analyses consists of households with at least one household member who has been assigned a PIK. We examine what household characteristics are associated with a household being unable to be linked to a PIK. The results of probit models for whether a household has a PIK are reported in Online Appendix Table 1. In each survey, several observable characteristics predict whether a household has a PIK, so we can reject that PIKs are missing at random. Yet there are few variables that systematically predict having a PIK in all surveys. We multiply survey weights by the inverse of the predicted probability of a household having a PIK (Wooldridge 2007). The covariates used in that prediction can be seen in Online Appendix Table 1. We discuss how the linkage process can affect estimated error rates further at the end of Section IV.

In all three surveys, the sample for our analyses is households with a householder at least 16 years of age. The food stamp assistance unit is notoriously difficult to capture in survey data, but this complication does not impinge on our analyses. We simply examine whether a household in the ACS, CPS, or SIPP that reports (or does not report) receipt of food stamps has any member that is a recipient in the administrative data. This reliance on the survey household definition greatly simplifies the analysis. Note that a survey household may contain more than one FSP assistance unit or only part of a unit.[10] For the analyses of the extent of errors and hence data accuracy in the next section, we use the entire (PIKed) population of households. To examine the accuracy of estimates typically obtained from the data, rather than its overall accuracy, we focus on a sample that would typically be used for such analyses (households below twice the poverty line).

10. To be clear, we are able to determine accurately what share of true recipient survey households report receipt, but we cannot determine what share of true recipient assistance units report receipt.

The administrative data record food stamp receipt on a monthly basis, which allows us to match the reference periods of the survey questions. The ACS asks about receipt in the past 12 months. To match this definition, we create a binary variable using the administrative data that indicates whether food stamps were received in the survey month or the previous 12 months by anyone in the household.[11] Food stamp receipt in the CPS refers to receipt in the previous calendar year, which we mimic in the administrative data. Seam bias is known to be an issue in the SIPP (Moore 2008), so we combine the four monthly reports of food stamp receipt from each interview to create an indicator for receipt during the four-month period, which we also do in the administrative data.

Downloaded from by guest on October 16, 2026. Copyright 2020

IV. Agreement between Survey and Administrative Reports

We first use our linked data to examine the differences in food stamp receipt according to the linked administrative variable and the survey reports. We take the administrative receipt measure to be accurate.[12] To obtain more precise estimates, we pool our two states in all surveys. We find substantial underreporting by true recipients and low rates, but sizable numbers of false positives in all surveys. The rates differ considerably between the three surveys, which leads the ACS and CPS to understate net food stamp receipt and the SIPP to overstate it slightly. We show that imputations are an important source of survey error, particularly of false positives. Finally, we examine to what extent the data linkage process is likely to affect our results.

A. Misclassification of Food Stamp Receipt

Table 2 presents sample sizes and statistics comparing food stamp receipt according to the administrative records and survey reports of receipt by the same household for the three surveys we examine. The first four columns contain unweighted observation counts. The population estimates and percentages in the remaining columns are weighted by household weights adjusted for missing PIKs. The first row of Table 2 contains the ACS results for Illinois and Maryland for the 2000–2001 period to which the survey refers. According to the linked administrative variable, 7.49 percent of households in Illinois and Maryland receive food stamps in a year. However, reporting errors are common; the false negative rate is 33 percent. Thus, one-third of households that receive food stamps are not recorded as recipients in the survey. The share of true nonrecipients who report receipt is 0.73 percent. Note that some recipients fail to report receipt, while some nonrecipients overreport. These two errors bias the reported receipt rate in different directions, leading to a net understatement lower than the false negative rate.[13] Overall, the high rate of false negatives

11. It is not entirely clear whether the reference period should include the month of the survey or not. We include it throughout, so that we define receipt based on a 13-month period. Error rates are only negligibly different when defining administrative receipt based on the 12 months preceding the current month.
12. As discussed above, we consider it accurate in the sense that the potential sources of error we discuss likely at most have a negligible impact on our estimates.
13. The difference is larger than the low false positive rate seems to suggest due to the much larger pool of nonrecipients.

Table 2
Reported and Administrative Food Stamp Receipt

	Unweighted Observation Counts				Receipt Rate		False Negative Rate	False Positive Rate
	Recipient		Nonrecipients					
	Report (1)	Admin. (2)	Report (3)	Admin. (4)	Report (5)	Admin. (6)	(7)	(8)
Panel A: Full Sample								
ACS	1,145	1,508	29,156	28,793	5.69%	7.49%	33.08%	0.73%
CPS	640	1,104	10,275	9,811	5.20%	8.69%	48.98%	0.84%
SIPP	817	793	10,138	10,162	6.14%	5.95%	22.82%	1.64%
Panel B: Imputed Observations								
ACS	256	219	215	252	64.93%	53.39%	3.21%	28.44%
CPS	62	108	331	285	12.07%	22.54%	79.63%	9.65%
SIPP	130	108	712	734	14.20%	11.00%	28.83%	7.16%

Notes: The false negative rate is the fraction of true recipient households with receipt not recorded in the survey. The false positive rate is the fraction of true nonrecipient households recorded as recipients in the survey. In the SIPP, we collapse receipt to the wave level. Columns 5–8 use household weights adjusted for incomplete linkage. The (weighted) imputation rates are 1.9, 3.6, and 7.3 percent in the ACS, CPS, and SIPP, respectively.

Downloaded from by guest on October 16, 2026. Copyright 2020

leads the survey report of food stamp receipt to be 5.69 percent for a net understatement of receipt of 24 percent in the ACS survey data.

The second line of Table 2 reports the same statistics for the CPS data; 8.69 percent of the households in the CPS receive food stamps in a calendar year according to the linked administrative variable. The share of food stamp recipient households that do not report receipt in the CPS is even higher than in the ACS; 49 percent of recipients do not report receipt. This share of false negatives has increased over the three (MD) or four years (IL) for which the administrative data are available. The increase is pronounced in Maryland, where by 2004 more than 60 percent of recipient households are not recorded as recipients. As in the ACS, the share of nonrecipients that report receipt is low, 0.84 percent. The net effect of false positives and false negatives is a substantial 40 percent understatement of the share of households receiving food stamps. This result accords quite closely with the net understatement by 39 percent for the Illinois time period and 38 percent for the Maryland time period that Meyer, Mok and Sullivan (2015a) find based on national aggregate data for months of participation.

The third line of Table 2 presents the same statistics for the 2001–2005 SIPP data; 5.95 percent of households in the SIPP receive food stamps according to the administrative data, and 23 percent of them fail to report receipt. Thus, the false negative rate in the SIPP is lower than in the ACS and substantially lower than in the CPS. On the other hand, the false positive rate is roughly twice as high as in the ACS and CPS; 1.64 percent of nonrecipient households report food stamp receipt. At least part of these differences is

Downloaded from by guest on October 16, 2026. Copyright 2020

likely due to the fact that we consider a household to report food stamp receipt if any household member reported receipt in any of the four reference months in the SIPP. This choice could drive down the rate of false negatives and increase the rate of false positives because anyone mistakenly reporting receipt in any of the four months results in a false positive.[14] The combination of the lower false negative and the higher false positive rate results in slight overreporting (by 3 percent) of food stamp receipt in the SIPP. Our findings support that the SIPP is the most accurate of the three data sets in measuring program receipt. It has the lowest false negative rate and the most accurate net reporting rate. Slight overreporting may well be preferable to the substantial underreporting in the ACS and CPS, particularly if one is mainly concerned with receipt rates. However, roughly half of this improvement stems from the higher false positives rate, that is, from introducing additional error, which may well aggravate the consequences of survey error in multivariate analyses, such as the models we analyze in Section VI.

In summary, we find low rates of false positives in two of the three surveys but substantial rates of false negatives in all three. These false negative rates are higher than those found in previous studies, often substantially so. The false negative rates exceed 50 percent in some cases, so analyses of government programs and the recipient population are likely to be severely biased in many situations. The low false positive rates in the ACS and the CPS imply that the aggregate underreporting rate (one minus the reporting rate) is a good approximation to the rate of false negative reports in those surveys, but not in the SIPP. This result is useful since aggregate rates are available for most years and for the entire United States, while our matched results are geographically and temporally limited. The large differences between false positive and false negative rates in all three surveys shows that misclassification is not completely random; that is, it depends on the true value. In the next section, we examine whether or not it is random conditional on truth, which is assumed in corrections such as in Hausman, Abrevaya, and Scott-Morton (1998).

We also find large differences in the error rates and hence net reporting across the three surveys. This degree of variation is in line with the wide range of misreporting rates found in previous studies. Contrary to these studies, we were able to link the same administrative data to three surveys using the same matching procedure. Hence, the differences we find between surveys can only be due to survey-specific characteristics, such as survey design, the focus of the survey, or its target population. For example, one factor that could contribute to the lower false negative rate in the SIPP is the shorter reference period, which should mitigate recall error.[15] More generally, the differences between the surveys provide further justification for the skepticism of both Bound, Brown, and Mathiowetz (2001) and Moore, Stinson, and Welniak (2000) that a general theory of misreporting can be developed. They also emphasize that survey error heavily depends on the implementation of a survey. This observation is borne out by two surveys as similar as the ACS and CPS yielding substantially different error rates for a relatively straightforward question. Consequently, conclusions regarding important issues, such

14. Contrary to the case of false negatives, pooling four months could also reduce the false positive rate if false positives mainly stem from reporting receipt in the wrong months, but we consider this unlikely.

15. A shorter reference period likely provides less or less-relevant information, so the potential error reduction would come at a cost.

as net reporting rates or whether and how the errors are related to observable characteristics, may have survey- and program-specific answers. This finding underlines the importance of further research into the determinants of survey errors, but also makes it unlikely that they can be explained by a general theory.

B. Accuracy of Imputed Observations

An important source of error in the overall data is item nonresponse. Our linked data provide the true recipiency status of nonrespondents, providing us with a unique opportunity to examine the accuracy of the imputed values the surveys include to address the problem of item nonresponse and whether imputation improves the quality of the data.[16] The bottom panel of Table 2 reports the same statistics as the top panel, but now only for item nonrespondents. Several patterns are evident from these estimates.

First, item nonresponse is an important issue for analyses of transfer programs. Even though overall imputation rates are low at 1.9, 3.6, and 7.3 percent of the population in the ACS, CPS, and SIPP,[17] respectively, a large share of recipient households are imputed: 13.6 percent of those receiving food stamps in the ACS, 9.3 percent in the CPS, and 13.6 percent in the SIPP. These statistics imply that item nonresponse predicts true receipt. The share of true food stamp recipients is higher among those who are imputed than among respondents, so excluding nonrespondents biases estimated receipt rates downwards. This potential bias is particularly pronounced in the ACS, where 53.4 percent of the imputed households are actual recipients, compared to 6.6 percent among nonimputed observations. The shares of true recipients among imputed and nonimputed observations also differ substantively in the CPS (22.5 percent compared to 8.2 percent) and the SIPP (11 percent compared to 5.6 percent). Thus, item nonresponse is not (unconditionally) random in all three surveys because the probability of obtaining a response is lower among true recipients in all three surveys. Most imputation methods yield consistent estimates under the weaker assumption that reporting status is independent of the true value conditional on covariates. That the likelihood of item nonresponse depends so strongly on the true value casts doubt on this key assumption. The result also underlines that the nature of item nonresponse is survey-specific because the households that choose not to respond differ in their probability of receiving food stamps across the three surveys. This result suggests that item nonresponse is significantly influenced by survey design.

Downloaded from by guest on October 16, 2026. Copyright 2020

16. Food stamp receipt in the ACS, CPS, and SIPP is imputed using hot deck methods. In the ACS, households (not in group quarters) are classified into cells defined by full interactions of family type, presence of children, poverty status, and the race of the reference person in each state. The data go through what is called a "geosort" before the imputation process. The most recent nonmissing response from a given cell at the smallest level of geography available is substituted for a missing response. In the CPS hot deck, households are classified into a much larger number of cells, but at the national level. The cells are defined by full interactions of number of people in the household (six categories), household income (nine categories), household type (three categories), age of the householder (two categories), and receipt of public assistance (two categories), for a total of 648 cells. Finally, the SIPP also imputes at the national level and only uses donors from the current wave. It applies a geosort to the data, but with much less geographic detail than the ACS. Food stamp receipt is then imputed within cells formed by age (six categories), race (two categories), sex (two categories), marital status (four categories), number of children (three categories), and work experience (three categories), for a total of 864 cells.

17. Note that in the SIPP we consider an observation to be imputed if any of the four reports was imputed.

Second, the imputations also fail to capture the marginal distribution of food stamp receipt: 22.5 percent of nonrespondents in the CPS are true food stamp recipients, but the CPS imputations only assign receipt to 12 percent of nonrespondents, thereby understating the rate of receipt by 46 percent. On the other hand, the imputations overstate true food stamp receipt among nonrespondents in the ACS by 21.6 percent and in the SIPP by 29.1 percent. Another criterion to evaluate imputations is whether they make the distribution in the entire sample align better with the true distribution. The overimputation in the ACS improves the net underreporting in the ACS. However, this "improvement" comes from introducing additional error, which may have negative effects on the joint distribution of food stamp receipt and other variables in the survey. The imputations in the other two surveys make net survey error worse, by leading to more overreporting in the SIPP and adding to the underreporting in the CPS.

Third, comparing imputed receipt to administrative receipt reveals that imputations induce substantial error at the household level. False negative rates among imputations are much lower than the overall rate in the ACS (3 percent), much higher in the CPS (80 percent), and slightly higher in the SIPP (29 percent). The low rate of false negatives in the ACS comes at the expense of a staggering false positive rate of 28 percent. False positive rates are substantially higher in the two other surveys as well, at 10 percent in the CPS and 7 percent in the SIPP. Consequently, a substantial share of false positives is due to imputation. Imputed observations account for 38 percent of false positives in the ACS, despite being no more than 1.6 percent of the total sample. Similarly, 3.6 percent of the sample are imputed, but account for 37 percent of the false positives in the CPS. Despite the much higher imputation share in the SIPP (7.7 percent), the imputed observations account for a lower, but still substantial, 27 percent of the false positives. When excluding imputed observations, the slight overreporting in the SIPP changes to slight underreporting. Because of these imputed false positives, the overall false positive rate in all of the surveys is much higher than the rate of overreporting by respondents. Thus, it is not a good indicator of households' tendency to report receipt when they are not recipients.

Taken together, our findings suggest that neither including nor excluding imputed observations is likely to solve the problem of item nonresponse. Receipt rates differ between respondents and nonrespondents, so excluding them will cause sample selection bias. However, including the imputed observation leads to bias from the substantial error rates we document. Therefore, data users are faced with the dilemma that both including and excluding imputed observations causes bias, and which strategy yields less bias is application-specific and unknown.

C. Potential Biases due to the Linkage Process

The data linkage process may lead to errors in the linked data for reasons such as missing or mismatched PIKs and households moving into one of the two states during the reference period. In this section, we discuss the extent of these problems and the likely biases they may cause in our estimated error rates.

First, some individuals may receive food stamps, but have no PIK in the survey data. We include households in our samples if anyone in the household has a PIK. So as long as at least one true recipient in the household has a PIK, we are able to classify it correctly as a recipient household. However, if none of the true recipient household members has

Downloaded from by guest on October 16, 2026. Copyright 2020

a PIK (but another member does), we would falsely classify the household as a nonrecipient household. This misclassification would understate true food stamp receipt. Affected households are true recipients, so we might reasonably assume that they have reporting rates higher than nonrecipients. It may also be reasonable to assume that their reporting rates are lower than those of the average recipient households, most of which have only recipient members. Then, as shown in the Online Appendix, the false positive rate is biased upward, and the false negative rate is biased downward. About 14 percent of ACS households with at least one PIK have members without a PIK, while 24 percent of CPS households in Illinois (15 percent in Maryland) have this situation. Thus, this bias could be substantial.

Second, a small fraction of the administrative records do not have a PIK. As in the previous case, this type of error will lead some true recipient households to not appear as recipients according to our administrative measure. If such households have reporting rates higher than true nonrecipients, but lower than other true recipients, the false positive rate would be overstated, and the false negative rate understated. The first condition seems likely given that these households are true recipients, while the second inequality is less clear. The share of administrative data without PIKs is very small, however.

Third, a PIK may be incorrectly assigned to a survey individual. If the household of this individual is a true recipient household, then the situation is analogous to the second case above and likely to increase false positives. The situation is slightly better because there is still a small chance that the erroneously assigned PIK belongs to another recipient, so that the household is still correctly classified. However, if the household is a true nonrecipient household, false negatives may be overstated if the incorrectly assigned PIK is from a true recipient. This situation should be uncommon. Most households do not receive food stamps, so the incorrectly assigned PIK is more likely to belong to a true nonrecipient household, which would lead us to classify the household correctly. Thus, the incorrect false negatives require the joint occurrence of two low probability events: an incorrectly assigned PIK and administrative food stamp receipt for that PIK.

Finally, a household that moved into the current state during the reference period of the survey may have received food stamps in their previous state, but not in their current state of residence. The administrative data from their current state of residence would not report that receipt. Thus, mobility across state lines will lead to an understatement of true food stamp receipt. As above, it seems reasonable to assume that these households have higher reporting rates than nonrecipients because they are true recipients. However, they are not currently receiving the program, so it also seems likely that they report at a lower rate than the average household. Under these assumptions, the false positive rate will be biased upward and the false negative rate biased downward (see the Online Appendix for a proof). Since only about 2 percent of individuals move across state lines in a year, the likely bias is small.

Overall, three of the four sources of error likely lead the administrative variable in the linked data to understate true receipt rates, implying that the linked data understate the false negative rate and overstate the false positive rate. The third case is hard to evaluate since the frequency of incorrectly assigned PIKs is not known, but the bias seems likely to be small. In consequence, linkage error likely results in an understatement of the true

Downloaded from by guest on October 16, 2026. Copyright 2020

receipt rate and thereby an overstatement of the true reporting rate. In the presence of net underreporting, this overstatement means that the linked data make the survey look more accurate in terms of the net reporting rate of the number of recipients. In terms of error rates, linkage errors likely make the data understate false negatives and overstate false positives.

V. What Affects the Agreement between the Survey Reports and the Administrative Records?

We next examine how misreporting of food stamp receipt differs across households. The previous section shows that error rates differ by true receipt status, so we analyze how errors vary with household characteristics conditional on true receipt. If misreporting does not depend on household characteristics conditional on true receipt, then it is fairly straightforward to analyze the bias it causes and correct estimates of take-up and the distributional effects of programs. Examples of such corrections can be found in Hausman, Abrevaya, and Scott-Morton (1998) and under the assumption of no false positives in Meyer, Mok, and Sullivan (2015a) and Meyer (2010). However, if misreporting is correlated with household characteristics, such corrections do not work well, and the biases are difficult to assess (Meyer and Mittag 2017; Mittag 2019; Nguimkeu, Denteh, and Tchernis 2019). Nonetheless, in such cases, we can use models of survey error, such as the ones we estimate in this section, to adjust statistical analyses (Bollinger and David 1997). We first examine the determinants of false negatives and then examine the determinants of false positives. We examine households with income less than twice the poverty line to focus on a group for whom food stamp receipt is especially relevant.[18] Due to the smaller SIPP sample, we continue to pool the data from Illinois and Maryland. Online Appendix Table 2 provides summary statistics for these samples.

Table 3 reports probit estimates for the determinants of false negative reporting in the ACS, the CPS, and the SIPP. Here the subsample consists of those who, according to the administrative data, are recipients of food stamps. We report average marginal effects on the probability of being a false negative reporter rather than coefficients to aid the interpretation of the magnitudes. The explanatory variables differ slightly due to availability in the three surveys, but all models include family type, number of adults and children, number of members who had a PIK, age categories, gender, education, ethnicity and employment status of the householder, whether the household is in a rural area, income relative to the household poverty line, reported receipt of other programs, receipt of TANF, and length of food stamp receipt from administrative data, as well as whether food stamp receipt was imputed. In the ACS and SIPP, we also examine whether the householder is disabled or a U.S. citizen and the role of language. In the CPS and the SIPP, we control for the time period. In the SIPP, we also include time in months since last food stamp receipt, a dummy if the household is in Maryland, and variables that are related to the quality of the interview.

18. We focus on this sample to make the results informative about the bias in the estimated models of receipt in the next section. Qualitative conclusions on the determinants of errors remain unchanged when analyzing the entire sample.

Downloaded from by guest on October 16, 2026. Copyright 2020

Downloaded from by guest on October 16, 2026. Copyright 2020

Table 3
Determinants of False Negatives, Probit Average Derivatives, Households with Income Less Than Twice the Poverty Line

	ACS		CPS		SIPP
	Illinois	Maryland	Illinois	Maryland	IL & MD
Single, no children	–0.0862	0.0437	–0.1312*	0.0558	–0.1644**
	(0.0716)	(0.0877)	(0.0779)	(0.1755)	(0.0662)
Single, with children	–0.0802	0.1203	–0.0227	–0.0323	–0.1426***
	(0.0539)	(0.0753)	(0.0620)	(0.1203)	(0.0481)
Multiple adults, no children	–0.1036	–0.0135	–0.0245	0.0668	0.0323
	(0.0857)	(0.1067)	(0.0739)	(0.1416)	(0.0843)
Number of members 18 or older	–0.0248	0.0405	0.0391	0.0370	0.0159
	(0.0341)	(0.0363)	(0.0371)	(0.0794)	(0.0210)
Number of members under 18	–0.0306	–0.0185	–0.0230	–0.0968	–0.0361*
	(0.0264)	(0.0329)	(0.0224)	(0.0616)	(0.0219)
Number of members PIKed	0.0308	0.0358	–0.0171	0.0484	–0.0060
	(0.0268)	(0.0333)	(0.0194)	(0.0433)	(0.0205)
Age ≥50	0.1514***	0.1319**	0.0881*	–0.1418*	–0.0044
	(0.0513)	(0.0663)	(0.0525)	(0.0832)	(0.0347)
Male	0.0877*	–0.0335	–0.0603	0.0195	–0.0798**
	(0.0356)	(0.0483)	(0.0446)	(0.0858)	(0.0359)

(*continued*)

Downloaded from by guest on October 16, 2026. Copyright 2020

Table 3 (*continued*)

	ACS		CPS		SIPP
	Illinois	Maryland	Illinois	Maryland	IL & MD
Less than high school	0.0688 (0.0431)	0.0659 (0.0589)	–0.0695 (0.0479)	–0.0620 (0.1111)	–0.1572*** (0.0517)
High school graduate	–0.0001 (0.0425)	0.1147** (0.0576)	–0.0293 (0.0463)	–0.0002 (0.0926)	0.0425 (0.0327)
College graduate and beyond	0.2197*** (0.0745)	–0.0586 (0.1201)	0.0373 (0.1103)	–0.0295 (0.1223)	–0.0460 (0.0463)
White	–0.0897** (0.0368)	–0.1110*** (0.0422)	–0.0503 (0.0415)	–0.0509 (0.0810)	–0.0672* (0.0355)
Employed					0.0428 (0.0273)
Unemployed	–0.0206 (0.0554)	–0.2504*** (0.0668)	0.0396 (0.0664)	0.0235 (0.1532)	
Not in labor force	–0.0077 (0.0404)	–0.0627 (0.0513)	0.0199 (0.0447)	–0.0074 (0.0832)	
Income divided by poverty line	0.0010*** (0.0003)	0.0008** (0.0004)	0.0010** (0.0004)	–0.0003 (0.0008)	–0.0082 (0.0304)
Disabled	–0.0637* (0.0386)	–0.0333 (0.0584)	n.d.	n.d.	0.0249 (0.0280)

(*continued*)

Downloaded from by guest on October 16, 2026. Copyright 2020

Table 3 (*continued*)

	ACS		CPS		SIPP
	Illinois	Maryland	Illinois	Maryland	IL & MD
Disabled, not working	–0.0382	0.1179**			
	(0.0465)	(0.0505)			
Speaks English only	0.0455	–0.1448*			
	(0.0507)	(0.0838)			
Speaks no or poor English					0.1690***
					(0.0454)
Non-U.S. citizen	–0.1545***	0.0697			–0.2039*
	(0.0327)	(0.1011)			(0.1162)
Rural	–0.1000**	–0.1079**	–0.0276	–0.0684	0.0145
	(0.0472)	(0.0476)	(0.0548)	(0.1346)	(0.0542)
Reported public assistance receipt	–0.2693***	–0.2453***	–0.3293***	n.d.	–0.2283***
	(0.0549)	(0.0632)	(0.0722)		(0.0749)
Reported housing assistance receipt	–0.0336	–0.0248	–0.1753***	–0.2732***	–0.1180***
	(0.0397)	(0.0481)	(0.0409)	(0.0871)	(0.0305)
Food Stamp receipt imputed	–0.3115***	–0.3833***	0.3580***	0.1932*	–0.1029**
	(0.0647)	(0.0899)	(0.0552)	(0.1103)	(0.0525)
Length of Food Stamp receipt spell	–0.0275***	–0.0384***	–0.0281***	–0.0196**	–0.0452***
	(0.0034)	(0.0036)	(0.0051)	(0.0086)	(0.0158)

(*continued*)

Downloaded from by guest on October 16, 2026. Copyright 2020

Table 3 (*continued*)

	ACS		CPS		SIPP
	Illinois	Maryland	Illinois	Maryland	IL & MD
Administrative TANF receipt	0.0658 (0.0446)	0.0273 (0.0514)	0.0986* (0.0580)	0.2466*** (0.0766)	0.0812 (0.0676)
Linear time trend			0.0222 (0.0157)	0.0980*** (0.0373)	
Household had bad data record					0.0936** (0.0450)
Months since last admin. Food Stamp receipt					0.0414*** (0.0132)
Household in MD					0.0893** (0.0442)
Observations	789	344	689	136	640

Notes: The unreported omitted family type is multiple adults with children, the education category is some college, the employment category is employed, the race group is nonwhite, and the geographic area is within-MSA. n.d. indicates that the variable was controlled for, but the estimate was not disclosed by the Census Bureau. In the ACS, we also controls for mode of interview (mail-back, CATI, CAPI). For the CPS, samples are pooled across all years (2002–2005). In the SIPP, we also control for year dummies, number of members interviewed, whether there was no interview with the reference person or an interview with a household member with no PIK. "Household had a bad data record" indicates that the reference person refused to answer many related survey questions. For the SIPP, samples include both states, are collapsed to the wave level, and are pooled across all years (IL: 10/2000–10/2004, MD: 10/2000–12/2003). All analyses conducted using household weights adjusted for PIK probability. Standard errors in parentheses. *$p<0.1$, **$p<0.05$, ***$p<0.01$.

Despite fairly small samples, there are many statistically significant determinants of false negative reporting. In all surveys, we easily reject the hypothesis that errors are unrelated to household characteristics. Consequently, misreporting is not conditionally random, because reporting rates vary with household characteristics even among true recipients. This finding violates the assumption of most corrections for misreporting and implies that (linear and nonlinear) instrumental variable methods are unlikely to give consistent estimates. Nguimkeu, Denteh, and Tchernis (2019) show that instrumental variable estimates of treatment effects can be severely biased. It also implies that the bias caused by survey error depends on the covariates that predict reporting errors. Intuitively, misreporting will cause larger downward biases for survey estimates of the receipt rates of subpopulations that are less likely to report true receipt. Attempting to address the errors by scaling up receipt rates by the net underreporting rate leads to overestimation for good reporters and underestimation for groups that report poorly. Consequently, understanding which variables predict survey error is important to assess what kind of analyses are likely to be biased and to examine the likely bias in practice. The predictors of errors are also informative about some of the theories of misreporting discussed above.

Even though the marginal effects of many of the variables are imprecisely estimated, some common themes emerge. Households with a householder 50 years or older are more likely to be false negatives (by 9–15 percentage points), except in the Maryland CPS sample, where the effect is large and negative. Several papers (discussed below) argue that the elderly are less likely to report program receipt for reasons such as stigma. Except for the positive effect in the Maryland CPS, our results support this hypothesis. As a consequence, part of the decline in estimated participation rates with age comes from decreasing reporting rates rather than decreasing rates of program receipt. The fact that higher income increases the likelihood that a recipient will not report receipt is also consistent with stigma being among the causes of misreporting. We also find that households where a language other than English is spoken (ACS) or where the householder speaks poor or no English (SIPP) are much more likely to fail to report food stamp receipt. This result can be taken as evidence that comprehension of the question is among the causes of misreporting. However, we also find that non-U.S. citizens are surprisingly less likely to fail to report, and the difference is significant in the SIPP and the ACS Illinois sample.

In terms of other demographic characteristics, households with a white householder are less likely to fail to report in all samples. The difference is sizeable (5–11 percentage points) and significant in the ACS and the SIPP. For the remaining demographic variables, our results are mixed or inconclusive. The marginal effects for households in rural areas are negative in four out of five models. They are less likely to fail to report in the ACS and CPS. The difference in the probability of reporting is large (ten percentage points) and significant in the ACS, but insignificant and small in the CPS and SIPP. Misreporting seems to be related to the gender of the householder, but the signs of the marginal effects differ across surveys. There is some evidence that households with a more educated householder are more likely to underreport, but the estimates are imprecise. Similarly, the marginal effect of being a single parent household with children is negative in four out of five models, but only significant in the SIPP. The effect of the number of adults is positive in four out of five cases, but always imprecisely estimated.

Downloaded from by guest on October 16, 2026. Copyright 2020

Downloaded from by guest on October 16, 2026. Copyright 2020

Quite uniformly, true recipients who report receipt of other programs (public assistance, housing assistance) are more likely to report food stamp receipt. The difference is large—for example, in the ACS, food stamp recipient households reporting public assistance receipt are nearly 20 percentage points less likely to fail to report food stamp receipt. We also find sizable effects of the duration of receipt in the reference period. An additional month of food stamp receipt is estimated to decrease the probability of failing to report by two to five percentage points. This result agrees with the idea that regularity of receipt is important and is also consistent with recall error being one of the reasons for false negatives. The SIPP provides further evidence of recall error, where we show that the number of months since last food stamp receipt in the reference period increases the probability of false negatives, by four percentage points per month. As the earlier analysis of imputed observations in Section IV.B presages, an imputation indicator is significant in all samples. We find little evidence of any effect of other variables related to the quality of the data, the interview, and the matching process.

In addition to our analysis of underreporting, we also examine the frequency of reporting receipt by those who are truly nonrecipients in Table 4. The sample for this false positive analysis, those who are truly nonrecipients, is much larger than that used for the false negative analysis. However, the false positive rate is so low that the number of false positives is much smaller than the number of false negatives. We can still easily reject the hypothesis that overreporting is unrelated to household characteristics, which confirms that survey error is not random, even after conditioning on truth. Given the small number of "ones" in this probit analysis, there are fewer significant determinants of reporting in these equations. Households with a householder 50 or older are less likely to misreport if they do not receive food stamps. The effect is negative throughout and significant for Illinois in the ACS and CPS. The fact that both recipient and nonrecipient elderly households are less likely to report receipt may indicate that stigma plays a larger role for the elderly. Similarly, income relative to the poverty line decreases the probability of false positives. The effect is significant except for Maryland in the ACS. This result may be additional evidence of stigma, but could also be explained by the fact that these households are less likely to receive food stamps and thus are less likely to make mistakes about their recipiency status. Households with a disabled householder are more likely to overreport. The number of household members under 18 matters, but goes in different directions in the surveys. It is significant and positive in the ACS (IL only) and the SIPP, but negative for Maryland in the CPS. There is some evidence that true recipient households with a white householder report more accurately. Reporting receipt of other programs, particularly a report of public assistance receipt, increases the probability of a false positive. This finding supports the hypothesis that misreporting is partly due to respondents confusing government programs (Nicholas and Wiseman 2009). The marginal effects of the imputation indicators confirm the finding of the last section that many false positives are due to imputation and that imputations are worse than reports by true nonrecipients in all surveys.

In conclusion, we show that both false positives and false negatives are systematically related to household characteristics in all surveys. Nonetheless, we find few consistent patterns. This finding may be due to the small sample sizes or because misreporting is mainly survey-specific. The variables that consistently predict survey error support common explanations for misreporting, such as comprehension, salience, recall, confusing government programs, and stigma. Even though specific effects are imprecisely

Downloaded from by guest on October 16, 2026. Copyright 2020

Table 4
Determinants of False Positives, Probit Average Derivatives, Households with Income Less than Twice the Poverty Line

	ACS		CPS		SIPP
	Illinois	Maryland	Illinois	Maryland	IL & MD
Single, no children					0.0026
					(0.0177)
Single, with children					0.0214*
					(0.0124)
Multiple adults, no children					–0.0337*
					(0.0174)
Number of members 18 or older	–0.0024	0.0053	0.0092	–0.0170	0.0130**
	(0.0034)	(0.0050)	(0.0067)	(0.0130)	(0.0061)
Number of members under 18	0.0069**	–0.0020	0.0044	–0.0251**	0.0142**
	(0.0033)	(0.0036)	(0.0049)	(0.0120)	(0.0065)
Number of members PIKed	–0.0085**	0.0060	–0.0047	0.0222*	–0.0117**
	(0.0038)	(0.0040)	(0.0044)	(0.0118)	(0.0048)
Age ≥50	–0.0225***	–0.0063	–0.0382***	–0.0010	–0.0100
	(0.0075)	(0.0086)	(0.0147)	(0.0109)	(0.0103)
Male	–0.0106*	0.0032	–0.0130	0.0106	–0.0032
	(0.0061)	(0.0080)	(0.0104)	(0.0094)	(0.0084)

(*continued*)

Downloaded from by guest on October 16, 2026. Copyright 2020

Table 4 (*continued*)

	ACS		CPS		SIPP
	Illinois	Maryland	Illinois	Maryland	IL & MD
Less than high school	0.0140**	0.0063	0.0193	n.d.	0.0091
	(0.0068)	(0.0099)	(0.0134)		(0.0089)
High school graduate	–0.0032	0.0111	–0.0001	0.0008	0.0086
	(0.0085)	(0.0126)	(0.0117)	(0.0079)	(0.0081)
College graduate and beyond					–0.0051
					(0.0127)
White	–0.0239***	–0.0082	0.0046	0.0094	–0.0207**
	(0.0071)	(0.0083)	(0.0098)	(0.0096)	(0.0083)
Employed	–0.0054	–0.0261*	–0.0016	0.0012	–0.0003
	(0.0066)	(0.0151)	(0.0117)	(0.0089)	(0.0070)
Income divided by poverty line	–0.0001***	–0.0001	–0.0003***	–0.0002**	–0.0257***
	(0.0000)	(0.0001)	(0.0001)	(0.0001)	(0.0069)
Disabled	0.0076	–0.0069	n.d.	n.d.	0.0441***
	(0.0084)	(0.0084)			(0.0073)
Disabled, not working	0.0159*	0.0226**			
	(0.0082)	(0.0097)			
Speaks no or poor English					0.0027
					(0.0125)

(*continued*)

Downloaded from by guest on October 16, 2026. Copyright 2020

Table 4 (*continued*)

	ACS		CPS		SIPP
	Illinois	Maryland	Illinois	Maryland	IL & MD
Non-U.S. citizen					–0.0129
					(0.0132)
Rural	–0.0051	n.d.			0.0120
	(0.0088)				(0.0095)
Reported public assistance receipt	0.0442***	0.0622***	0.0957***	0.0872***	0.1020***
	(0.0091)	(0.0186)	(0.0197)	(0.0332)	(0.0210)
Reported housing assistance receipt	0.0108	0.0007	0.0571***	–0.0032	0.0523***
	(0.0070)	(0.0081)	(0.0146)	(0.0116)	(0.0089)
Food Stamp receipt imputed	0.0700***	0.0447***	0.0544***	0.0443***	0.0470***
	(0.0110)	(0.0139)	(0.0113)	(0.0156)	(0.0130)
Administrative TANF receipt					–0.1093***
					(0.0422)
Linear time trend			0.0018	0.0000	
			(0.0047)	(0.0056)	
Household in MD					–0.0309*
					(0.0163)
Observations	3,357	1,455	1,462	504	2,333

Notes: The unreported omitted education category for the false negative probits is some college or more, the race group is nonwhite, the employment category is not employed, and the geographic area is within-MSA. n.d. indicates that the variable was controlled for, but the estimate was not disclosed by the Census Bureau. In the ACS, we also controls for mode of interview (mail-back, CATI, CAPI). For the CPS, samples are pooled across all years (2002–2005). In the SIPP, we also control for year dummies, number of members interviewed, whether there was no interview with the reference person or an interview with a household member with no PIK, and whether the reference person had a bad data record (refused to answer many related survey questions). For the SIPP, samples include both states, are collapsed to the wave level, and are pooled across all years (IL: 10/2000–10/2004, MD: 10/2000–12/2003). All analyses conducted using household weights adjusted for PIK probability. Standard errors in parentheses. $*p<0.1$, $**p<0.05$, $***p<0.01$.

estimated, they are jointly significant, so that we reject the hypothesis that errors are random conditional on truth in all surveys. It is unclear how this systematic survey error affects estimates, such as program effects or those from the binary choice models, that are commonly used to examine program take-up. Coefficient estimates, such as the ones produced here, could be used to correct such models as in Bollinger and David (2001) or Meyer and Mittag (2017). These estimates also enter the formulas for the bias in estimated program effects in Nguimkeu, Denteh, and Tchernis (2019).

VI. The Effect of Survey Error on Estimates of Program Receipt

The previous sections show substantial misreporting of food stamp receipt and that it is systematically related to household characteristics. It is well known that such nonclassical measurement error causes bias, but little is known about the direction and magnitude of the bias in general. We use our measure of truth in the linked administrative data to analyze an important case, binary choice models of program receipt. Such models are often used to analyze program targeting (see Currie 2006 or Haider, Jacknowitz, and Schoeni 2003).[19] Meyer and Mittag (2017) derive the bias for probit models with reported receipt as the dependent variable that these analyses usually employ. Their results imply a tendency of marginal effects to retain the correct sign. However, this prediction can be overturned, for example, when the covariates strongly predict survey errors. Thus, it is important to assess whether such results provide a useful characterization of the consequences of survey error in practice.

Having true food stamp receipt matched to survey data gives us the opportunity to estimate this bias directly and examine whether the use of administrative data provides a different understanding of the determinants of food stamp receipt than the survey data alone. We first estimate the determinants of receipt using only survey data. We then reestimate the determinants of receipt using the survey covariates, but with the administrative measure of receipt as the dependent variable. We then compare the two equations for the determinants of food stamp use. Throughout this section, we report average marginal effects[20] and restrict our sample to households with income below twice the poverty line to have a sample for which food stamp receipt is a likely possibility.

The determinants of food stamp receipt in the ACS are in the first four columns of Table 5. The results using only survey data are in Column 1 for Illinois and in Column 3 for Maryland. The survey estimates suggest that, controlling for household income, a single parent household is about ten percentage points more likely to be a recipient than a married couple household in both states. Those 50 or older are much less likely to be reported participants than those ages 40–49 in Illinois, while in Maryland the effect is only evident for those 60 or older. The differences in receipt for these older groups are large—ten percentage points in Illinois and nine percentage points in Maryland—

19. A related strand of literature examines estimates of program effects in the presence of misclassification of program receipt, see, for example, Gundersen and Kreider (2009); Kreider (2010); Millimet (2011); Gundersen, Kreider, and Pepper (2012); Kreider et al. (2012); Alamada, McCarthy, and Tchernis (2016); Jensen, Kreider, and Zhylyevskyy (2019); and Nguimkeu, Denteh, and Tchernis (2019).

20. The overall results are very similar for the coefficients, though the differences are smaller in some cases, but not uniformly so.

Downloaded from by guest on October 16, 2026. Copyright 2020

Downloaded from by guest on October 16, 2026. Copyright 2020

Table 5
Determinants of Reported and Administrative Food Stamp Receipt, Probit Marginal Effects, Linked Households with Income Less than Twice the Poverty Line

	ACS				CPS				SIPP	
	Illinois		Maryland		Illinois		Maryland		IL & MD	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	Survey Report	Admin. Receipt	Survey Report	Admin. Receipt	Survey Report	Admin. Receipt	Survey Report	Admin. Receipt	Survey Report	Admin. Receipt
Single, no children	0.0670	0.1164[a]	0.0861	0.1485	–0.0119	0.0001	–0.0687	–0.0229	0.0505	0.0329
	(0.0320)	(0.0361)	(0.0461)	(0.0515)	(0.0256)	(0.0386)	(0.0511)	(0.0623)	(0.0223)	(0.0278)
Single, with children	0.1076	0.1429[a]	0.1083	0.1965[b]	0.0547	0.1333[b]	0.0133	0.0775	0.1262	0.1280
	(0.0247)	(0.0272)	(0.0351)	(0.0389)	(0.0214)	(0.0308)	(0.0437)	(0.0491)	(0.0149)	(0.0185)
Multiple adults, no children	0.0696	0.0959	0.0547	0.0975	0.0192	0.0664	–0.0509	0.0235	0.0186	–0.0064
	(0.0344)	(0.0392)	(0.0500)	(0.0547)	(0.0226)	(0.0346)	(0.0413)	(0.0560)	(0.0245)	(0.0307)
Number of members under 18	0.0188	–0.0066[b]	0.0202	0.0027	0.0227	0.0309	0.0235	0.0541[a]	0.0199	0.0052
	(0.0099)	(0.0145)	(0.0144)	(0.0191)	(0.0058)	(0.0087)	(0.0117)	(0.0181)	(0.0065)	(0.0113)
Number of members 18 or older	0.0027	–0.0201[a]	0.0039	0.0153	–0.0069	0.0128	–0.0213	0.0055	0.0155	–0.0121[c]
	(0.0111)	(0.0138)	(0.0174)	(0.0208)	(0.0104)	(0.0143)	(0.0258)	(0.0246)	(0.0074)	(0.0096)
Number of members PIKed	0.0145	0.0692[c]	0.0165	0.0612[c]					0.0048	0.0281[c]
	(0.0076)	(0.0131)	(0.0118)	(0.0183)					(0.0036)	(0.0069)

(*continued*)

Downloaded from by guest on October 16, 2026. Copyright 2020

Table 5 (*continued*)

	ACS				CPS				SIPP	
	Illinois		Maryland		Illinois		Maryland		IL & MD	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Age 16–29	–0.0208	–0.0055	0.0274	0.0141	–0.0111	–0.0378	–0.0086	–0.0428	0.0298	0.0210
	(0.0231)	(0.0264)	(0.0300)	(0.0332)	(0.0198)	(0.0291)	(0.0287)	(0.0431)	(0.0159)	(0.0205)
Age 30–39	0.0061	0.0061	–0.0386	–0.0454	–0.0118	0.0040	–0.0285	–0.0043	0.0338	–0.0151[c]
	(0.0221)	(0.0262)	(0.0288)	(0.0323)	(0.0194)	(0.0280)	(0.0257)	(0.0419)	(0.0150)	(0.0188)
Age 50–59	–0.0981	–0.0405[b]	–0.0315	–0.0375	0.0016	0.0287	0.0249	0.0382	0.0112	–0.0269[b]
	(0.0261)	(0.0294)	(0.0366)	(0.0369)	(0.0228)	(0.0369)	(0.0291)	(0.0461)	(0.0171)	(0.0190)
Age 60–69	–0.1144	–0.0806	–0.0856	–0.0702	–0.0110	–0.0625	0.0372	–0.0052	–0.0039	–0.0075
	(0.0278)	(0.0320)	(0.0358)	(0.0384)	(0.0240)	(0.0353)	(0.0344)	(0.0519)	(0.0175)	(0.0219)
Age ≥70	–0.1641	–0.1619	–0.1346	–0.1354	–0.1313	–0.1579	–0.0714	–0.1675[a]	0.0080	0.0316
	(0.0278)	(0.0321)	(0.0359)	(0.0386)	(0.0254)	(0.0352)	(0.0353)	(0.0599)	(0.0174)	(0.0206)
Less than high school	0.0648	0.0687	0.0739	0.1089[a]	0.0503	0.0455	–0.0056	0.0073	0.0100	0.0077
	(0.0184)	(0.0218)	(0.0237)	(0.0271)	(0.0165)	(0.0248)	(0.0262)	(0.0405)	(0.0122)	(0.0155)
High school graduate	0.0239	0.0318	0.0130	0.0510	0.0266	0.0409	0.0031	–0.0085	0.0137	0.0387[a]
	(0.0186)	(0.0212)	(0.0232)	(0.0255)	(0.0158)	(0.0236)	(0.0241)	(0.0360)	(0.0114)	(0.0139)

(*continued*)

Downloaded from by guest on October 16, 2026. Copyright 2020

Table 5 (*continued*)

	ACS				CPS				SIPP	
	Illinois		Maryland		Illinois		Maryland		IL & MD	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
College graduate and beyond	−0.0584	−0.0569	0.0114	−0.0147	−0.0892	−0.1557	0.0191	−0.0420	−0.0223	0.0004
	(0.0313)	(0.0329)	(0.0361)	(0.0407)	(0.0267)	(0.0442)	(0.0300)	(0.0510)	(0.0152)	(0.0187)
White	−0.0380	−0.0801[c]	0.0055	−0.0355[b]	−0.0211	−0.0762[c]	0.0048	−0.0118	−0.0767	−0.1069[c]
	(0.0178)	(0.0191)	(0.0187)	(0.0211)	(0.0133)	(0.0196)	(0.0182)	(0.0261)	(0.0100)	(0.0112)
Employed	−0.0380	−0.0217	−0.0488	−0.0078[a]	−0.0399	−0.0665	−0.0391	−0.0633	−0.0146	−0.0275
	(0.0164)	(0.0188)	(0.0227)	(0.0247)	(0.0141)	(0.0207)	(0.0191)	(0.0280)	(0.0092)	(0.0112)
Income divided by poverty line	−0.0007	−0.0007	−0.0010	−0.0013[b]	−0.0009	−0.0015[c]	−0.0003	−0.0003	−0.0841	−0.1028[a]
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0002)	(0.0083)	(0.0100)
Disabled	0.0906	0.0774	0.0773	0.0933	0.0466	0.0377	0.1046	0.0022[a]	0.1220	0.1550[b]
	(0.0182)	(0.0209)	(0.0235)	(0.0249)	(0.0451)	(0.0719)	(0.0629)	(0.0867)	(0.0116)	(0.0143)
Disabled, not working	0.0271	0.0086	0.0093	0.0465						
	(0.0193)	(0.0224)	(0.0242)	(0.0266)						
Speaks English only	0.0343	0.0850[c]	0.0716	0.0772						
	(0.0207)	(0.0245)	(0.0306)	(0.0393)						

(*continued*)

Downloaded from by guest on October 16, 2026. Copyright 2020

Table 5 (*continued*)

	ACS				CPS				SIPP	
	Illinois		Maryland		Illinois		Maryland		IL & MD	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Speaks no or poor English									0.0393	0.0814[c]
									(0.0138)	(0.0167)
Rural	0.0293	0.0458	0.0499	0.0491	0.0275	0.0383	0.0495	0.0682	0.0386	0.0384
	(0.0191)	(0.0189)	(0.0183)	(0.0225)	(0.0167)	(0.0262)	(0.0278)	(0.0388)	(0.0122)	(0.0160)
Reported public assistance receipt	0.3189	0.2386[b]	0.3020	0.3728	0.2179	0.2077	0.1934	0.2246	0.1676	0.0883[c]
	(0.0240)	(0.0315)	(0.0324)	(0.0408)	(0.0268)	(0.0432)	(0.0327)	(0.0590)	(0.0244)	(0.0262)
Reported housing assistance receipt	0.1461	0.1811[b]	0.1021	0.1337	0.1517	0.1999	0.1378	0.1593	0.1522	0.1355
	(0.0184)	(0.0217)	(0.0198)	(0.0241)	(0.0147)	(0.0243)	(0.0221)	(0.0364)	(0.0116)	(0.0132)
Linear time trend					0.0039	0.0180[a]	–0.0002	0.0329[b]		
					(0.0053)	(0.0079)	(0.0096)	(0.0164)		
2001									–0.0373	–0.0687[b]
									(0.0124)	(0.0152)
2002									–0.0360	–0.0517
									(0.0133)	(0.0164)

(*continued*)

Downloaded from by guest on October 16, 2026. Copyright 2020

Table 5 (*continued*)

	ACS				CPS				SIPP	
	Illinois		Maryland		Illinois		Maryland		IL & MD	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
2003									–0.0228	–0.0581[b]
									(0.0137)	(0.0180)
2005									0.0188	–0.0207[c]
									(0.0167)	(0.0195)
Household in MD									–0.0493	–0.0008[c]
									(0.0146)	(0.0161)
Observations	4,591	4,146	1,945	1,799	2,981	2,151	808	640	4,177	2,973
p-value admin equal survey (jointly)		<0.0001		0.0004		<0.0001		0.0085		<0.0001

Notes: For each survey, the first column contains estimates from a probit model using reported receipt as the dependent variable. The second column estimates the same model using the administrative receipt measure as the dependent variable. Superscript a, b, and c in these columns indicate that a χ^2-test whether the survey and administrative estimates are equal rejects at the 10, 5, and 1 percent level, respectively. Standard errors of the estimates are in parentheses. All analyses conducted using household weights adjusted for PIK probability. All demographic characteristics refer to the reference person. The unreported omitted family type is multiple adults with children, the age group is 40–49, the education group is some college, the race group is nonwhite, the employment group is not employed, and the geographic area is within MSA. For the CPS, samples are pooled across all years (2002–2005). For the SIPP, samples include both states, are collapsed to the wave level, and are pooled across all years (IL: 10/2000–10/2004, MD: 10/2000–12/2003).

compared to those 40–49. The marginal effects of education and income have the expected signs, with high school dropouts six percentage points more likely to report participation in Illinois and seven percentage points more likely in Maryland than those with some college. Income is a strong predictor of reported food stamp receipt. In Illinois, households with income equal to half the poverty line are seven percentage points more likely to report food stamp receipt than households with income 1.5 times the poverty line. In Maryland, the difference is ten percentage points. The survey estimates also suggest that households with a nonemployed or disabled householder are much more likely to receive food stamps. In Illinois, nonwhites are more likely to report participation, while there is little difference by race in Maryland. According to survey reports, those reporting housing assistance receipt are more than 1.5 times as likely to be recipients than an average household. Those reporting public assistance receipt are more than twice as likely to be recipients.

Replacing the mismeasured ACS survey receipt variable with the administrative measure of receipt paints a different picture of the determinants of food stamp participation. Columns 2 and 4 of Table 5 repeat the analysis, substituting an administrative dependent variable for the poorly reported survey measure of receipt. Superscript letters in Columns 2 and 4 indicate the level of significance from tests of equality of the marginal effects based on the survey data alone and those based on the survey and administrative combined data. The joint χ^2-tests in the last row clearly reject that the combined data yield the same estimated marginal effects as the ACS survey data alone for both states. Single households, both with and without children, are much more likely to be recipient households in the combined data. In Illinois, the difference is four to five percentage points, while in Maryland, it is six to nine percentage points, and most differences are at least marginally statistically significant. The average marginal effects for race also differ significantly, with the administrative specifications indicating that participation is four percentage points greater for nonwhites than the survey data only specifications indicate in each state. Most marginal effects for reported receipt of public assistance or housing benefits are significantly different. In Illinois, the marginal effect of age, particularly for age 50–59, is quite different in the combined data, and the difference is statistically significant. The association with speaking English only is also significantly different. For Maryland, the association with income is quite different in the combined data, indicating a substantially quicker decline in participation with income. Overall, 16 marginal effects differ significantly, but only four out of 46 marginal effects change their direction due to survey errors. This pattern is in line with the theoretical results from Meyer and Mittag (2017).

We report the determinants of food stamp participation using the CPS data in Columns 5–8 of Table 5. Again, Columns 5–7 of this table provide the average marginal effects for the models that use only survey data. As in the ACS survey data results, all else equal, single parent households are more likely to be recipients, though the relationship is not significant in Maryland. Households with many children are more likely to report food stamp receipt, and this difference is significant in both states. Households with householders 70 years or older are less likely to receive food stamps, while those that have very low income, a nonemployed householder, or who report receipt of either public assistance or housing benefits are significantly more likely to receive food stamps in both states according to the CPS reports. In Illinois, those without a high school

Downloaded from by guest on October 16, 2026. Copyright 2020

degree are more likely, and those with a college degree less likely, to receive than those with some college. The survey data alone do not suggest that food stamp receipt has been rising with time in either of the states.

When we substitute the administrative measure of receipt for the poorly reported survey measure in Columns 6 and 8 of Table 5, the determinants of reporting change in important ways. As the χ^2-test p-values at the bottom of Columns 6 and 8 indicate, in both states we reject that the marginal effects are jointly the same using the administrative and the survey dependent variable. Eleven out of 42 marginal effects in the CPS change their sign due to survey error. On one hand, this pattern is surprising given the prediction that sign changes require strong conditions. On the other hand, it is not surprising that we find sign changes to be most frequent in the survey with the highest misclassification rate. Substantively, the difference in participation between single parent families and those with married parents changes from five percentage points to 13 in Illinois and from one percentage point to eight in Maryland with the administrative data measure. In Illinois, the change is statistically significant, while it is not in Maryland. In Maryland, there is some evidence of an increased marginal effect of the number of children in a household. Food stamp participation is also much higher among nonwhites and drops off more quickly with income than in the survey data alone in Illinois. Contrary to the survey data, which showed no time trend, the combined data provide evidence of increasing receipt in both states, which is relevant to recent research by Mulligan (2012) and Ganong and Liebman (2018).

The results using the SIPP survey reports are in Column 9 of Table 5 and are similar to the other two surveys. Single parents are again more likely to receive food stamps, all else being the same. In the SIPP this also applies to single individuals. As in the other two surveys, income relative to the poverty line has a negative impact on reported program take-up. Households in rural areas and with a householder reporting a disability or poor English skills are more likely to receive benefits according to SIPP reports. Households with a nonwhite householder are more likely to be (reported) participants. There is a strong positive association between reporting food stamps and receipt of other programs (housing assistance and TANF). Reported participation seems to decline with age, but the evidence is weak. Contrary to the CPS, there is a time trend in the survey reports, but it is flat until 2003 and then increases sharply.

Column 10 of Table 5 reports the SIPP results that use the administrative dependent variable. The joint test rejects that the results from the two dependent variables are the same. The SIPP estimates align well with the prediction from Meyer and Mittag (2017), as only six out of 27 marginal effects change sign. Several marginal effects are significantly different. The number of adults has a pronounced negative effect in the administrative data. As in the other two surveys, the effects of race and income are more pronounced when using administrative food stamp receipt, while the association with reporting other programs is weaker. The marginal effects of two age categories (30–39 and 50–59) change significantly. While the survey data suggest that participation declines over the life-cycle, the relation is U-shaped in the administrative data, increasing sharply after age 50, though the marginal effects are imprecisely estimated. Despite being equally likely to receive food stamps as those in Illinois, households in Maryland are almost five percentage points less likely to report receiving food stamps. The time trend is clearly different when using the administrative dependent variable. Growth in program

Downloaded from by guest on October 16, 2026. Copyright 2020

participation is more rapid in the first half of the time period and slower in the second half using the accurate data.

In summary, survey error clearly changes what we learn about program receipt. One of the key differences between the combined administrative and survey data and the survey data alone is in participation by age. Haider, Jacknowitz, and Schoeni (2003) and Wu (2010) emphasize lower food stamp take-up by older households in survey data. Gundersen and Ziliak (2008) find a more complicated pattern by age. In some cases, the differences in misreporting by age we document in Section V make the combined data show much less of a difference between the aged and the nonaged, thus explaining a significant part of the puzzle in past work. We see this pattern in our largest sample, that for Illinois using ACS data, although it is not evident in the CPS data. Another noteworthy difference is the impact of income relative to the poverty line. Food stamp receipt declines more rapidly with income in the administrative data, so analyses using survey data only are likely to understate the distributional consequences of the Food Stamp Program. Finally, survey error has a pronounced impact on the time trend in food stamp receipt. In the CPS, the survey reports conceal the time trend, while in the SIPP they suggest a flat profile followed by a steep increase instead of a more steady increase. The time pattern of receipt has been a key issue in recent work on food stamps, such as Mulligan (2012) and Ganong and Liebman (2018).

However, while the survey data alone would lead one to make incorrect inferences in some cases, the overall picture obtained from the survey data is fairly accurate in qualitative terms. Most of the significant marginal effects remain significant, and changes in the sign of marginal effects are rare when one goes from the survey data alone to the combined data. Overall, only 21 out of 115 marginal effects change sign. This pattern holds even in the CPS, where half of true food stamp recipients fail to report. That few marginal effects change sign conforms to theoretical predictions in Meyer and Mittag (2017), suggesting that the asymptotic biases can be useful in assessing the bias in practice. If the tendency for misclassification to not affect the sign of estimates in such models holds more generally, we may still be able to draw important qualitative conclusions from contaminated survey data. Future research should explore the generality of this result.

Downloaded from by guest on October 16, 2026. Copyright 2020

VII. Conclusions

Benefit receipt in major household surveys is often misreported, hindering our understanding of government programs and the economic circumstances of disadvantaged populations. We use administrative data on Food Stamp Program participation from Illinois and Maryland matched to ACS, CPS, and SIPP household survey data to examine the extent and consequences of such survey errors. We show that more than 30 percent of true recipient households do not report receipt in the ACS, approximately 50 percent do not report receipt in the CPS, and 23 percent do not report in the SIPP. False positive rates are much lower, at less than 1 percent in the ACS and CPS and 1.6 percent in the SIPP. Imputation matters for analysis of program receipt because item nonresponse is frequent among recipients. Receipt rates differ between respondents and the overall population, so only using respondents results in biased

population estimates. Imputed observations introduce substantial error, with a large share of false positives being due to imputation in all three surveys. Imputations do not correctly reproduce the probabilities of receipt among item nonrespondents. We discuss the potential bias from linkage errors on these error rates, finding that such errors likely lead to an understatement of false negative errors and an overstatement of false positive errors.

Misreporting, both false negatives and false positives, varies with household characteristics such as income, race, and age. The relation of these errors to frequently used covariates will lead to biases that are difficult to assess and complicated to correct for because it renders most corrections for misreporting invalid and makes it difficult to distinguish the effect of such characteristics on reporting from their effect on true receipt. See Mittag (2019) and Davern, Meyer, and Mittag (2019) for discussion and methods of correction based on estimates such as the ones provided here. The characteristics that predict misreporting suggest that comprehension, salience, recall, stigma, and complex patterns of program receipt are among the determinants of survey errors, as theories of misreporting predict. However, with our small sample, it is difficult to provide definitive tests of theories of misreporting.

Finally, we examine bias in the determinants of program receipt using our combined administrative and survey data, which include accurate participation from the administrative data and household explanatory characteristics from the survey that are missing in the administrative data. Our food stamp participation results differ from conventional estimates using only survey data in several important ways. Participation is higher among single parents and nonwhites and declines more quickly with income than the survey data alone suggest. Participation by age and the patterns of multiple program participation are also different using the administrative variable. The results indicate that underreporting is part of the explanation for the low receipt rate among the elderly. Lastly, using only the survey data, one would miss much of the rise in food stamp participation. It is also possible to think of the glass as half full, rather than half empty. It is striking that the signs of most determinants of food stamp receipt in the survey data alone match those in the combined administrative and survey data, even in the CPS, where half of true food stamp recipients are not recorded as recipients. Further evidence on this pattern might clarify the conditions under which this finding holds more generally.

Our results also suggest biases in other studies where program receipt is used as an explanatory variable in a regression. We show that the errors of measurement are correlated with the true values as well as with a range of explanatory variables. This nonclassical form of the errors means that the bias will usually take a complicated form. Substantively, erroneous program receipt will affect studies of who receives benefits and why they do and of program effects on labor supply, health, consumption, and other outcomes. Studies that examine the extent to which food stamps increase the resources of poor families will tend to understate their impact. A better understanding of underreporting and how it may bias program receipt estimates is important for both policymakers and researchers. Accurate estimates of program receipt are needed to know who benefits from programs, why some choose not to participate in certain programs, and how individual characteristics affect participation. Since we find that survey error leads to biased estimates of the determinants of program receipt, policies based on survey data alone may be misguided.

Downloaded from by guest on October 16, 2026. Copyright 2020

References

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

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.

Almond, Douglas, Hilary W. Hoynes, and Diane Whitmore Schanzenbach. 2011. "Inside the War on Poverty: The Impact of Food Stamps on Birth Outcomes." *Review of Economics and Statistics* 93(2):387–403.

Bitler, Marianne P., Janet Currie, and John-Karl Scholz. 2003. "WIC Eligibility and Participation." *Journal of Human Resources* 38(S):1139–79.

Blank, Rebecca M., and Patricia Ruggles. 1996. "When Do Women Use AFDC & Food Stamps? The Dynamics of Eligibility vs. Participation." *Journal of Human Resources* 31(1):57–89.

Blundell, Richard, and Luigi Pistaferri. 2003. "Income Volatility and Household Consumption: The Impact of Food Assistance Programs." *Journal of Human Resources* 38(S):1032–50.

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

———. 2001. "Estimation with Response Error and Nonresponse: Food-Stamp Participation in the SIPP." *Journal of Business and Economic Statistics* 19(2):129–41.

———. 2005. "I Didn't Tell, and I Won't Tell: Dynamic Response Error in the SIPP." *Journal of Applied Econometrics* 20(4):563–69.

Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. "Measurement Error in Survey Data." In *Handbook of Econometrics*, Volume 5, ed. J.J Heckman and E. Leamer, 3705–843. Amsterdam: Elsevier.

Celhay, Pablo, Bruce D. Meyer, and Nikolas Mittag. 2022. "What Leads to Measurement Error? Evidence from Reports of Program Participation in Three Surveys." NBER Working Paper 29652.

———. 2021. "Errors in Reporting and Imputation of Government Benefits and Their Implications." NBER Working Paper 29184. Cambridge, MA: NBER.

Cerf Harris, Ben. 2014. "Within and across County Variation in SNAP Misreporting: Evidence from Linked ACS and Administrative Records." U.S. Census Bureau CARRA Working Paper 2014-05. Washington, DC: U.S. Census Bureau.

Coder, John, and Lydia Scoon-Rogers. 1996. "Evaluating the Quality of Income Data Collected in the Annual Supplement to the March Current Population Survey and the Survey of Income and Program Participation." Working Paper. Washington, DC: U.S. Census Bureau Housing and Household Economic Statistics Division.

Courtemanche, Charles, Augustine Denteh, and Rusty Tchernis. 2019. "Estimating the Associations between SNAP and Food Insecurity, Obesity, and Food Purchases with Imperfect Administrative Measures of Participation." *Southern Economic Journal* 86(1):202–28.

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, 80–148. New York: Russell Sage Foundation.

Currie, Janet, and Jeffrey Grogger. 2001. "Explaining Recent Declines in Food Stamp Program Participation." *Brookings-Wharton Papers on Urban Affairs* 2001:203–29.

Davern, Michael, Bruce D. Meyer, and Nikolas Mittag. 2019. "Creating Improved Survey Data Products Using Linked Administrative-Survey Data." *Journal of Survey Statistics and Methodology* 7(3):440–63.

Downloaded from by guest on October 16, 2026. Copyright 2020

Figlio, David, Craig Gundersen, and James Ziliak. 2000. "The Effects of the Macro-Economy and Welfare Reform on Food Stamp Caseloads." *American Journal of Agricultural Economics* 82(3):635–41.

Ganong, Peter, and Jeffrey B. Liebman. 2018. "The Decline, Rebound, and Further Rise in SNAP Enrollment: Disentangling Business Cycle Fluctuations and Policy Changes." *American Economic Journal: Economic Policy* 10(4):153–76.

Groves, Robert M. 1989. *Survey Errors and Survey Costs*. New York: John Wiley & Sons.

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):352–82.

———. 2009. "Bounding the Effects of Food Insecurity on Children's Health Outcomes." *Journal of Health Economics* 28(5):971–83.

Gundersen, Craig, Brent Kreider, and John V. Pepper. 2012. "The Impact of the National School Lunch Program on Child Health: A Nonparametric Bounds Analysis." *Journal of Econometrics* 166(1):79–91.

Gundersen, Craig, and Victor Oliveira. 2001. "The Food Stamp Program and Food Insufficiency." *American Journal of Agricultural Economics* 83(4):875–87.

Gundersen, Craig, and James P. Ziliak. 2003. "The Role of Food Stamps in Consumption Stabilization." *Journal of Human Resources* 38(S):1051–79.

———. 2008. "The Age Gradient in Food Stamp Program Participation: Does Income Volatility Matter?" In *Income Volatility and Food Assistance in the United States*, ed. Dean Jolliffe and J.P. Ziliak, 171–214. Kalamazoo, MI: Upjohn.

Haider, Steven, Alison Jacknowitz, and Robert Schoeni. 2003. "Food Stamps and the Elderly: Why Is Participation so Low?" *Journal of Human Resources* 38(S):1180–220.

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):239–69.

Hoynes, Hilary W., and Diane Whitmore 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.

Jensen, Helen H., Brent Kreider, and Oleksandr Zhylyevskyy. 2019. "Investigating Treatment Effects of Participating Jointly in SNAP and WIC When the Treatment Is Validated Only for SNAP." *Southern Economic Journal* 86(1):124–55.

Kang, Kyung M., and Robert A. Moffitt. 2019. "The Effect of SNAP and School Food Programs on Food Security, Diet Quality, and Food Spending: Sensitivity to Program Reporting Error." *Southern Economic Journal* 86(1):156–201.

Kapteyn, Arie, and Jelmer Y. Ypma. 2007. "Measurement Error and Misclassification: A Comparison of Survey and Administrative Data." *Journal of Labor Economics* 25(3):513–51.

Kirlin, John A., and Michael Wiseman. 2014. "Getting it Right, or at Least Better: Improving Identification of Food Stamp Participants in the National Health and Nutrition Examination Survey." Unpublished.

Kreider, Brent. 2010. "Identification Decay of Regression Coefficients in the Presence of Infrequent Classification Errors." *Review of Economics and Statistics* 92(4):1017–23

Kreider, Brent, John V. Pepper, Craig Gundersen, and Dean Jolliffe. 2012. "Identifying the Effects of SNAP (Food Stamps) on Child Health Outcomes When Participation Is Endogenous and Misreported." *Journal of the American Statistical Association* 107(499):958–75.

Marquis, Kent H., Naihua Duan, M. Susan Marquis, J. Michael Polich, J.E. Meshkoff, Donna S. Schwarzbach, and Cathy Stasz. 1981. "Response Errors in Sensitive Topic Surveys: Estimates, Effects, and Correction Options." Santa Monica, CA: RAND Corporation.

Marquis, Kent H., and Jeffrey C. Moore. 1990. "Measurement Errors in SIPP Program Reports." SIPP Working Paper. Washington, DC: U.S. Bureau of the Census.

Downloaded from by guest on October 16, 2026. Copyright 2020

Meijer, Erik, Susann Rohwedder, and Tom Wansbeek. 2012. "Measurement Error in Earnings Data: Using a Mixture Model Approach to Combine Survey and Register Data." *Journal of Business & Economic Statistics* 30(2):191–201.
Meyer, Bruce D. 2010. "The Effects of the Earned Income Tax Credit and Recent Reforms." In *Tax Policy and the Economy*, Volume 24, ed. Jeffrey Brown, 153–80. Cambridge, MA: MIT Press.
Meyer, Bruce D., and Nikolas Mittag. 2017. "Misclassification in Binary Choice Models." *Journal of Econometrics* 200(2):295–311.
———. 2019a. "Using Linked Survey and Administrative Data to Better Measure Income: Implications for Poverty, Program Effectiveness and Holes in the Safety Net." *American Economic Journal: Applied Economics* 11(2):176–204.
———. 2019b. "Misreporting of Government Transfers: How Important Are Survey Design and Geography?" *Southern Economic Journal* 86(1):230–53.
———. 2021a. "An Empirical Total Survey Error Decomposition." *Journal of Econometrics* 224(2):286–305.
———. 2021b. "Combining Administrative and Survey Data to Improve Income Measurement." In *Administrative Records for Survey Methodology*, ed. Asap Y. Chun, Michael D. Larsen, Gabriela Durrant, and Jerome P. Reiter, 297–322. New York: John Wiley and Sons.
Meyer, Bruce D., Wallace K.C. Mok, and James X. Sullivan. 2015a. "The Underreporting of Transfers in Household Surveys: Its Nature and Consequences." Working Paper. Chicago, IL: Harris School of Public Policy Studies, University of Chicago
———. 2015b. "Household Surveys in Crisis." *Journal of Economic Perspectives* 28(4):199–226.
Millimet, Daniel L. 2011. "The Elephant in the Corner: A Cautionary Tale about Measurement Error in Treatment Effects Models." In *Missing Data Methods: Cross-Sectional Methods and Applications*, Advances in Econometrics, Volume 27, Part 1, ed. D. Drukker, 1–39. Bingley, UK: Emerald Publishing.
Mittag, Nikolas. 2019. "Correcting for Misreporting of Government Benefits" *American Economic Journal: Economic Policy* 11(2):142–64.
Moore, Jeffrey C. 2008. "Seam Bias in the 2004 SIPP Panel: Much Improved, but Much Bias Still Remains." U.S. Census Bureau Statistical Research Division Survey Methodology Research Report Series 2008-3. Washington, DC: U.S. Census Bureau.
Moore, Jeffrey C., Kent H. Marquis, and Karen Bogen. 1996. "The SIPP Cognitive Research Evaluation Experiment: Basic Results and Documentation." SIPP Working Paper 212. Washington, DC: U.S. Census Bureau.
Moore, Jeffrey C., Linda L. Stinson, and Edward J. Welniak, Jr. 2000. "Income Measurement Error in Surveys: A Review." *Journal of Official Statistics* 14(4):331–61.
Mulligan, Casey. 2012. *The Redistribution Recession: How Labor Market Distortions Contracted the Economy.* Oxford, UK: Oxford University Press.
Nguimkeu, Pierre, Augustine Denteh, and Rusty Tchernis. 2019. "On the Estimation of Treatment Effects with Endogenous Misreporting." *Journal of Econometrics* 208(2):487–506.
Nicholas, Joyce, and Michael Wiseman. 2009. "Elderly Poverty and Supplemental Security Income." *Social Security Bulletin* 69(1):45–73.
———. 2010. "Elderly Poverty and Supplemental Security Income 2002–2005." *Social Security Bulletin* 70(2):1–29.
Niehaus, Paul, and Sandip Sukhtankar. 2013. "Corruption Dynamics: The Golden Goose Effect." *American Economic Journal: Economic Policy* 5(4):230–69.
NORC. 2011. "Assessment of the US Census Bureau's Person Identification Validation System." NORC Final Report presented to the U.S. Census Bureau.
Oberski, Daniel L., Antje Kirchner, Stephanie Eckman, and Frauke Kreuter. 2017. "Evaluating the Quality of Survey and Administrative Data with Generalized Multitrait-Multimethod Models." *Journal of the American Statistical Association* 112(520):1477–89.

Downloaded from by guest on October 16, 2026. Copyright 2020

Pierret, Charles R. 2001. "Event History Data and Survey Recall: An Analysis of the National Longitudinal Survey of Youth 1979 Recall Experiment." *Journal of Human Resources* 36(3): 439–66.
Remler, Dahlia K., and Sherry A. Glied. 2003. "What Other Programs Can Teach Us: Increasing Participation in Health Insurance Programs." *American Journal of Public Health* 93(1):67–74.
Roemer, Marc I. 2000. "Assessing the Quality of the March Current Population Survey and the Survey of Income and Program Participation Income Estimates 1990–1996." Staff Paper. Washington, DC: U.S. Census Bureau Household Economic Statistics Division.
Rothbaum, Jonathan L. 2015. "Comparing Income Aggregates: How Do the CPS and ACS Match the National Income and Product Accounts 2007–2012." SEHSD Working Paper 2015-01. Washington, DC: U.S. Census Bureau.
Schmidt, Lucie, Lara Shore-Sheppard, and Tara Watson. 2016. "The Effect of Safety-Net Programs on Food Insecurity." *Journal of Human Resources* 51(3):589–614.
Scholz, John K., Robert A. Moffitt, and Benjamin Cowan. 2009. "Trends in Income Support." In *Changing Poverty, Changing Policies*, ed. M. Cancian and S. Danziger, 1–32. Washington, DC: Russell Sage Foundation.
Sudman, Seymour, and Norman M. Bradburn. 1974. *Response Effects in Surveys.* Chicago: Aldine.
Sudman, Seymour, Norman M. Bradburn, and Norbert Schwarz. 1996. *Thinking about Answers: The Application of Cognitive Processes to Survey Methodology.* San Francisco: Josey-Bass.
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." Working Paper. Washington, DC: U.S. Census Bureau.
Tiehen, Laura, Dean Jolliffe, and Craig Gundersen. 2012. "Alleviating Poverty in the United States: The Critical Role of SNAP Benefits." ERR-132. Washington, DC: U.S. Department of Agriculture, Economic Research Service.
U.S. General Accounting Office. 2004. "Food Stamp Program: Steps Have Been Taken to Increase Participation of Working Families, but Better Tracking of Efforts Is Needed." GAO-04-346. Washington, DC: GAO.
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." Washington, DC: U.S. Census Bureau.
Wheaton, Laura. 2007. "Underreporting of Means-Tested Transfer Programs in the CPS and SIPP." In *2007 Proceedings of the American Statistical Association, Social Statistics Section*. Alexandria, VA: American Statistical Association
Wooldridge, Jeffrey M. 2007. "Inverse Probability Weighted Estimation for General Missing Data Problems." *Journal of Econometrics* 141(2):1281–301.
Wu, Yanyuan. 2010. "Essays on the Economic Well-Being of the Elderly and Public Policy." Ph.D. Dissertation. Chicago: University of Chicago.
Ziliak, James, Craig Gundersen, and David Figlio. 2003. "Food Stamp Caseloads over the Business Cycle." *Southern Economic Journal* 69(4):903–19.

Downloaded from by guest on October 16, 2026. Copyright 2020