Race, Ethnicity, and Measurement Error

Bruce D. Meyer[*]
University of Chicago,
NBER, and AEI

Nikolas Mittag[†]
CERGE-EI

Derek Wu[‡]
University of Virginia

March 12, 2024

Abstract

Large literatures have analyzed racial and ethnic disparities in economic outcomes and access to the safety net. For such analyses that rely on survey data, it is crucial that survey accuracy does not vary by race and ethnicity. Otherwise, the observed disparities may be confounded by differences in survey error. In this paper, we review existing studies that use linked data to assess the reporting of key programs (including SNAP, Social Security, Unemployment Insurance, TANF, Medicaid, Medicare, and private pensions) in major Census Bureau surveys, aiming to extract the evidence on differences in survey accuracy by race and ethnicity. Our key finding is a strong and robust, but previously largely unnoticed, pattern of greater measurement error for Black and Hispanic individuals and households relative to whites. As the dominant error is under-reporting for a wide variety of programs, samples, and surveys, the implication is that the safety net better supports minority groups than the survey data suggest, through higher program receipt and greater poverty reduction. We conclude that racial and ethnic minorities are inadequately served by our large household surveys and that researchers should cautiously interpret survey-based estimates of racial and ethnic differences. We briefly discuss paths forward.

[*] University of Chicago, Harris School of Public Policy Studies, Email: bdmeyer@uchicago.edu
[†] CERGE-EI, Email: nikolasmittag@posteo.de.
[‡] University of Virginia, Batten School of Leadership and Public Policy, Email: derek.wu@virginia.edu

We thank Mandana Vakil for excellent research assistance. For helpful comments, we thank Randy Akee, Larry Katz, Mark Loewenstein, Maxim Massenkoff, and audiences at the 2024 AEA Meetings and the 2023 NBER CRIW pre-conference. We are grateful to the Alfred P. Sloan Foundation, Russell Sage Foundation, Charles Koch Foundation, Menard Family Foundation, National Institute on Aging, and American Enterprise Institute for their support of the Comprehensive Income Dataset Project.

1. Introduction

For decades, researchers have examined racial disparities in economic outcomes (see, e.g., Duncan 1968, Massey and Denton 1993, Akee, Jones, and Porter 2019, Chetty et al. 2020). Survey estimates from the Census Bureau, for example, showed that median household incomes were 35% lower and the poverty rate was nearly twice as high for Black Americans in 2022 compared to white non-Hispanic Americans (Shrider and Creamer 2023, U.S. Census Bureau 2023). A growing literature using survey data has also uncovered important differences by race and ethnicity in the receipt of government programs (see, e.g., Gould-Werth and Shaefer 2012, Kuka and Stuart 2021, Forsythe and Yang 2021, U.S. GAO 2022). One group of these papers emphasizes lower receipt of transfers, particularly Unemployment Insurance, by Black and Hispanic individuals, suggesting that state laws and administrative burdens have led to disparate access to these programs. A second group of papers, focusing mostly on means-tested transfers, examines whether or not Black and Hispanic individuals are overrepresented among program recipients and the extent to which the differences can be explained by income and employment patterns (see, e.g., Moffitt and Gottschalk 2001 or the review in Currie 2006).

For such analyses that rely on survey data, it is important not only that they capture resources and program receipt well overall but also that survey accuracy does not vary by race and ethnicity. Otherwise, the racial disparities in these analyses may be confounded by differences in survey error. Survey error has long been known to be pervasive for income and program receipt. For a wide range of income sources and years, weighted totals of recipients and dollars reported in the major U.S. household surveys have been shown to severely understate administrative aggregates (Meyer, Mok, and Sullivan 2015; Rothbaum 2015). A number of papers go beyond comparisons of aggregates and link survey and administrative microdata to show substantial survey error in the receipt of transfer programs, most commonly in the form of survey underreporting (e.g., Marquis and Moore 1990, Huynh et al. 2002, Meyer and Mittag 2019a, 2021a). Despite these high error rates, differences in distributional statistics by characteristics such as race and ethnicity could still be unbiased if these characteristics do not predict survey errors. While some studies have found that survey error varies systematically with demographics (e.g., Bollinger and David 1997, Meyer, Mittag, and Goerge 2022), the literature has paid surprisingly little attention to the question of differences in reporting and how they affect analyses of disparities between racial and ethnic subgroups. Even when race and ethnicity variables are included in these studies, the empirical evidence is generally dispersed and underemphasized.

In this paper, we fill these gaps by synthesizing the evidence from eighteen empirical studies on differences in survey accuracy by race and ethnicity, focusing on differences between white, Black, and Hispanic individuals (and Asians when available) in existing studies that link administrative records to survey data. We then discuss the consequences of the differences we document and summarize the evidence on the biases found for commonly estimated statistics. We close by discussing paths forward for improving the accuracy of survey estimates for minority groups. Our main focus is on means-tested programs including the Supplemental Nutrition Program (SNAP, formerly Food Stamps), Temporary Assistance for Needy Families (TANF), and General Assistance (GA), as well as Unemployment Insurance (UI), government-provided health insurance (Medicaid, Medicare and the Indian Health Service), Social Security (Old-Age, Survivors, and Disability Income or OASDI), and private pensions in three major Census Bureau surveys: the American Community Survey (ACS), the Current Population Survey Annual Social and Economic Supplement (CPS ASEC, hereafter just CPS), and the Survey of Income and Program Participation (SIPP).

We begin by discussing how misclassification of program receipt – both reports of non-receipt by true recipients (false negatives) and reports of receipt by true non-recipients (false positives) – differs by race and ethnicity. These error rates are the most frequently reported measures of error in reports of binary variables like program receipt and are of key relevance for assessing and correcting the biases from survey error. We examine differences in raw false positive and false negative rates, as well as how they differ after controlling for other demographics. Black and Hispanic individuals have consistently higher false negative rates and usually higher false positive rates. We find striking evidence that race and ethnicity strongly and reliably predict survey error, both independently of other covariates and after adjusting for demographics. The differences are large enough to skew important conclusions. For example, OASDI receipt is roughly twice as likely to be not recorded in the survey data for minorities. These patterns are robust across multiple surveys, samples, and income sources, suggesting that the differences in survey accuracy by race and ethnicity are not limited to a particular setting but are a general phenomenon. Thus, these large surveys appear to inadequately serve racial and ethnic minorities, for whom researchers and other users of the survey data have less accurate information than for white individuals.

We then discuss implications of these pronounced and systematic differences in survey error for key statistics, specifically measures of program receipt and poverty. For each type of statistic, we briefly summarize theoretical results on the consequences of the non-classical measurement error we document, before reviewing the empirical evidence. We examine the biases in rates of program receipt and models of program take-up, before examining implications for poverty rates on the basis of the available empirical evidence. We also discuss how these results change for continuous variables (like program dollars) and examine likely biases in estimates of racial and ethnic differences when mismeasured variables are used as the dependent variable or as a covariate. However, less empirical evidence is available for these cases. Our review shows that while the magnitudes of the differences in survey accuracy suggest substantial bias, the biases also depend on the statistic of interest and which minority households are misrecorded. For example, the higher false positive rates for one group could theoretically offset the downward bias associated with higher false negative rates when estimating differences in program receipt. This possibility does not appear to be important in practice, as the surveys on net understate receipt rates more severely for racial and ethnic minorities than for whites. This pattern holds both for raw measures of receipt and for the effects of minority indicators in models of take-up. For government programs where the survey already indicated higher receipt rates among minorities, the differences are amplified. For programs where the survey indicated lower receipt rates among minorities, the differences typically disappear or change sign. Given that survey data acutely understate the resources available to minorities and the degree to which they are supported by the safety net, we also document that poverty rates are systematically overstated for minorities.

We conclude by discussing some paths forward. Improving the accuracy of surveys is difficult, as the fundamental causes of the differences in survey error by race and ethnicity are unclear. They could arise from differences in stigma or in trust in the government, with the latter consistent with well-documented patterns of institutional distrust among Black Americans borne out of historical traumas (see, e.g., Brandon, Isaac, and LaVeist 2005, Bajaj and Stanford 2021). These differences could also arise from neighborhood characteristics that make interviewers less comfortable. They may further result from the way surveys are designed and implemented, with higher error rates potentially due to minorities being assigned to interviewers who have less expertise or are more likely to be of a different race/ethnicity. Linked data can help us understand these factors better, but improving surveys in this way can be a slow and difficult process. We argue that in lieu of identifying causes, linked data can point to circumstances when errors are severe and help us devise better strategies to reduce survey error. Linked data can also help data users understand likely biases or even improve estimates. However, a key conclusion of our review is that until more accurate data (or more reliable information on the nature of errors) become available, researchers should be cautious in interpreting estimates of racial or ethnic disparities in program receipt and income from survey data alone.

2. Overview of Studies and Methodology

In this section, we provide an overview of the linked data studies that we examine. We review key methodological aspects and discuss potential problems for studies of race and ethnicity. The main obstacle to evaluating survey accuracy lies in obtaining measures of truth that are comparable to the survey concepts of interest. For transfer programs, comparisons to administrative aggregates clearly indicate substantial underreporting (Meyer, Mok, and Sullivan 2015). However, evaluating survey accuracy for subpopulations, including those by race or ethnicity, would at minimum require accurate aggregate statistics for each subpopulation. Such measures of truth are rarely available. Linking surveys to administrative microdata has emerged as a key tool for remedying this problem (see Meyer and Mittag 2021b for an overview of data linkage methods used to validate income information). In this paper, we focus on studies that link survey data from the Census Bureau to administrative records. This choice focuses on heavily used surveys, and keeps the studies we review relatively homogeneous in terms of the survey data used, while standardizing the quality of the administrative data and the methods used for data linkage. Table 1 provides an overview of the studies we rely on and discusses the datasets, income sources, and samples.

2.1 Data Sources

All studies in this paper are based on three major U.S. household surveys conducted by the Census Bureau: the ACS, the CPS ASEC, and the SIPP.[1] The ACS is the largest household survey

[1] One study included in our review (Brummet et al. 2018) uses the Consumer Expenditure (CE) Survey, which is conducted by the Census Bureau under contract for the Bureau of Labor Statistics (BLS). However, the CE Survey is large-scale and general interest, and it can be linked to administrative records using the same probabilistic linkage system as the other Census surveys we examine. We therefore include it given the high degree of comparability.

in the U.S., with 2.8 to 3.5 million households (0.7-0.8 million prior to 2005) selected to participate during each year we examine. The CPS is one of the most important economic surveys in the U.S. and the linkage studies in our review all use the Annual Social and Economic Supplement to the CPS (the official source of income and poverty statistics in the U.S). Since 2001, the CPS ASEC has interviewed a sample of about 94,000-99,000 households, although sample sizes in earlier years were closer to 60,000-70,000. Finally, the SIPP serves as the highest quality source of information on low-income households and the receipt of government transfers. The studies in this review primarily use the 2001, 2004, and 2008 panels of the SIPP, which sampled approximately 25,000-45,000 households intended to be surveyed for a period of approximately 4 years.

These three surveys are each large-scale and general interest, but there are also pronounced differences in their design that are related to both non-response and measurement error (see, e.g., Celhay, Meyer, and Mittag 2024). The ACS questionnaire is administered by mail/internet, telephone, or an in-person interview, and the CPS conducts interviews in person and by phone. While the ACS and CPS only interview one household member (i.e., the reference person), the SIPP strives to conduct in-person interviews every four months with every member of the household over the age of 15 (see, e.g., U.S. Census Bureau 2006, 2008, 2014). In terms of reference periods, the ACS asks about income and program receipt in the 12 months prior to the interview date, the CPS asks about the previous calendar year, and the SIPP asks for monthly information in the four months preceding the interview.

The linkage studies we survey vary widely in the years and samples they use. Key sample restrictions are summarized in Table 1 and are usually based on the availability of administrative data. For linked data to provide the necessary measure of truth, the administrative records must be accurate and sufficiently detailed to exactly match the survey concept of interest. The administrative program records used in these studies are typically based on actual payments and personal information verified as part of eligibility determination. Studies that examine federal programs and income sources often utilize administrative records covering the entire U.S., with UI amounts from Internal Revenue Service (IRS) Form 1099-G, Medicare enrollment records from the Center for Medicare and Medicaid Services (CMS), Indian Health Service (IHS) receipt from IHS patient registration files, Social Security amounts from the Social Security Administration’s Payment History Update System, and private pensions from IRS Form 1099-R. For programs like SNAP and TANF that are administered by individual states, administrative records are only available for a subset of states. While Medicaid is also administered at the state level, its administrative records are available in a centralized fashion via CMS for the vast majority of states.

2.2 Data Linkage

Survey data and administrative records typically do not contain common identifiers that can be used to exactly merge the two data sources. Hence, data linkage is typically probabilistic – i.e., based on algorithms that try to find records referring to the same unit in separate data sources (see, e.g., Winkler 2021). By virtue of restricting our review to Census Bureau surveys, the studies we review all employ the same linkage methods based on person identifiers created by the Person Identification Validation System (PVS) (Wagner and Layne 2014). In short, the PVS uses person data (such as address, name, gender, and date of birth) from the administrative records and survey data to search for a matching record in a reference file containing all transactions recorded against a Social Security Number (SSN). If a matching record is found, the SSN of the record from the reference file is transformed into a Protected Identification Key (PIK) – akin to a scrambled SSN – and attached to the corresponding original record in the survey or administrative data.

In addition to the administrative data being of high quality, the linkage itself must also be sufficiently accurate for the linked administrative value to approximate a measure of truth. Linkage error can stem from missed links and wrong links, although missed links almost exclusively arise from missing PIKs in the survey data (as PIK rates are close to 100% in our administrative records). Contrary to most other linkage studies, those using the PVS can readily quantify and adjust for missed links.[2] For recent years, PIK rates in the survey data tend to be high (over 90%), although earlier survey years had lower PIK rates. Except for one (Dushi and Trenkamp 2021), all studies we consider use the linkable subsample of PIKed observations, and a number of studies also use inverse probability weighting (IPW) to adjust survey weights and restore the representativeness of the sample to its population of interest (Wooldridge 2007). Here, we are less interested in estimates being representative of some underlying population and more in estimates being comparable between subpopulations. This may not be the case if the association between having a PIK and survey accuracy differs between whites and minorities, although including race and ethnicity

[2] This advantage arises from the fact that the PVS links both the administrative and survey records to a third data source from which common identifiers are obtained. Linking both data sources to such a population register lets the researcher distinguish cases when a record was not linked to the other source because it was not included in the data (e.g., because the household did not receive the program) from cases when the record was unlinkable.

indicators in IPW models can partially address this concern. There is some evidence that PIK rates are higher for white individuals, but the evidence is mixed (Meyer, Mittag, and Goerge 2022) and the differences are generally small enough to yield only trivial effects on the patterns we find. Differences in PIK rates are more likely to arise from the uniqueness of names and the frequency of interactions with the government than from differences in reporting behavior.

Less is known about the frequency of linkage errors – i.e., the frequency with which a survey unit is linked to the wrong record from the administrative data. Such linkage errors, if random, will tend to mute differences between subpopulations and lead our results to understate the problems we identify. Thus, the main concern is whether or not the frequency of linkage errors differs by race and ethnicity. Little is known about the prevalence of linkage errors and their causes, although one may be concerned about undocumented immigrants (many of whom are Hispanic) having more linkage errors. However, the direction of the bias is not clear. Meyer, Mittag, and Goerge (2022) conclude that linkage errors are likely to overstate false positives and understate false negatives. We find false negatives to be higher for minorities, suggesting that greater linkage error for minorities would, if anything, understate the patterns we find.

Overall, the accuracy of the administrative data, the high PIK rates, and the sophisticated linkage procedure in our reviewed studies suggest that the linked administrative variables can be considered close approximations to truth. A caveat is that the information on covariates (including race and ethnicity) is self-reported. Detailed demographic information is an invaluable aspect of survey data and is often not available in administrative records. We therefore take the accuracy of these demographic variables as given throughout, although systematic differences in their reporting could bias our estimates. That being said, survey reports of race and ethnicity may in fact be more appropriate than those appearing in administrative records, as the latter indicators are not used in program administration and are likely to be less accurate and scrutinized than the information on payments used to validate receipt. And until more reliable information on these demographics becomes available, researchers will continue to rely on such self-reported measures – implying that the differences identified in our review provide the relevant measures of bias.

3. Differences in Survey Accuracy by Race and Ethnicity

For the programs we examine, the main source of survey error appears to be error at the extensive margin – i.e., whether or not receipt is recorded correctly (see, e.g., Meyer and Mittag

2021a). We start by summarizing the evidence on raw error rates for race and ethnicity subgroups, followed by an overview of error rates adjusted for demographic differences in the population. We then briefly summarize the scant evidence on differences in survey error amounts, and close by discussing other aspects of survey accuracy on which evidence is urgently needed.

3.1 Differences in Raw False Negative and False Positive Rates

Validation studies of program receipt typically report separate error rates for false negatives (true recipients recorded as non-recipients) and false positives (true non-recipients recorded as recipients). The separate focus on false positive and false negative rates (rather than a single focus on overall error rates) stems from the fact that over- and underreporting have been found to differ substantially. For government transfers, studies have found high rates of false negatives that sometimes exceed 50% (Celhay, Meyer and Mittag 2021), while false positives are much less frequent and often lower than 1%. However, these patterns differ by program, with false positive rates found to be higher for reports of Social Security and Medicare (Bee and Mitchell 2017, Bhaskar et al. 2019). In our review, we are interested less in the overall levels of error and more in how error rates differ between minorities and whites.

Table 2 reports false positive and false negative rates from the studies that estimate these rates separately by race and ethnicity. Almost all error rates are higher for minorities, with the magnitudes of the differences being substantively important. First, we find that minorities have higher false negative rates than whites for all income sources examined outside of private pensions. Most differences are large in both absolute and relative terms. In absolute terms, the differences are largest for UI, with false negative rates in the CPS being a staggering 18.4 percentage points (55%) higher for Hispanic individuals, 15.9 percentage points (48%) higher for Black individuals, and 10.1 percentage points (30%) higher for Asian individuals relative to whites. The UI results from the SIPP show largely comparable patterns. For OASDI, the gaps in false negative rates are similar in relative terms: 11.5%, 11.2%, and 10.2% of true OASDI recipients among Black, Hispanic, and Asian individuals do not have payments recorded in the survey, which is around twice as high as the rate for whites (5.6%). These gaps in false negative rates are fractionally smaller but still pronounced for Medicare (5.3-7.9% for minorities vs. 3.8% for whites).

For SNAP, using SIPP data for 2010-2013, Colby et al. (2016) document sizeable differences in false negative rates of 3.6 percentage points (22%) for Black individuals and 4.1

percentage points (26%) for Hispanic individuals relative to whites. Using more recent SIPP data for 2014-2020, Giefer et al. (2022) find false negative rates for Black individuals to be only 0.5 percentage points higher than for white individuals. For Medicaid, Davern et al. (2009a) find similarly high false negative rates for all subgroups, with Black and Hispanic individuals having slightly higher rates (within a percentage point) and Asian individuals having meaningfully higher false negative rates by 4.3 percentage points (10%).[3] In contrast, false negative rates for pensions are 3-11 percentage points lower for minorities relative to whites, with this exception potentially due to pension receipt rates being relatively lower among minorities.

False positive rates are also higher for racial and ethnic minorities for nearly all income sources examined. Some differences are large in absolute terms. For example, false positive rates for SNAP in the 2010-2013 SIPP are 6.5 percentage points higher for Black than for white individuals. The difference in OASDI false positive rates between Asian and white individuals is of a similar magnitude at 6.7 percentage points. As false positive rates tend to be much lower than false negative rates (with the exception of Medicare), the differences also tend to be smaller in absolute terms. Yet, they tend to be much larger in relative terms. False positive rates among Black individuals are more than twice as high than among white individuals in 4 out of 9 cases in Table 2. The patterns persist but are less pronounced for Hispanic and Asian individuals, for whom false positive rates exceed those of white individuals in all but one case (Medicare). Thus, Table 2 shows a clear and alarming pattern that both false positive and false negative rates tend to be more frequent for minorities, particularly Black and Hispanic individuals.

3.2 Differences in Covariate-Adjusted False Negative and False Positive Rates

While the error rates above differ substantially between groups, we do not have the results of tests for the statistical significance of such differences. Therefore, we also examine covariate-adjusted error rates in the form of marginal effects of race/ethnicity indicators from regressions where the dependent variable is a binary indicator for the misreporting of program receipt. A key advantage of looking at covariate-adjusted error rates is that one can use simple *t*-tests of the marginal effects to assess whether or not the differences are statistically significant for minorities

[3] Note that Davern et al. (2007, 2009b) also examine false negative and false positives for Medicaid by race/ethnicity, but their analyses condition on being a recipient or non-recipient based on survey records (rather than administrative records). Given this lack of comparability to the other estimates reviewed, we omit these two papers from our review.

relative to whites (the reference group). Another advantage is that they are based on models that control for demographic characteristics. Both receipt rates and the demographics associated with survey error differ substantially between racial and ethnic groups and between programs. Consequently, it is particularly useful to understand whether survey data from units with similar demographic characteristics differ systematically by race and ethnicity.

Studies typically show separate models for whether receipt is missing for true recipients (false negatives) and whether true non-recipients are recorded as recipients in the survey (false positives). Table 3 provides an overview of the estimated covariate-adjusted differences in error rates in the literature, showing a clear pattern of significantly higher false positive and false negative rates for racial and ethnic minorities – even for early linkage studies with small samples. We focus our discussion on the average partial effects in Panel A of Table 3, as they are easier to compare and interpret (although the logit coefficients in Panel B yield similar conclusions). For Black individuals, the false negative differences are all positive and for the most part statistically significant. The differences are large – tending to hover around 7-11 percentage points higher than the false negative rates for whites when analyzing SNAP, TANF, and UI. Black individuals also have higher covariate-adjusted false positive rates than whites, with the sign of all statistically significant differences being positive across all studies. In line with the raw error rates discussed above, the differences in false positive rates are smaller in absolute terms (usually up to 2 percentage points higher for minorities) but still large in relative terms.

The patterns for Hispanics are similar, as they also have more frequent instances of both types of survey error than whites. Yet, there are some slight differences. The magnitudes of the false negative gaps (relative to whites) appear to be smaller and less statistically significant for Hispanic than for Black individuals. However, the range of estimates is large for Hispanics, with the difference in covariate-adjusted false negative rates being as large as 12 percentage points relative to whites for UI (in the CPS). The covariate-adjusted false positive rates for Hispanics also exceed those of white individuals for most programs, but contrary to the case of Black individuals, there are some exceptions. Hispanic individuals have lower false positive rates for cash assistance (TANF and General Assistance combined) and Medicare. The results are more sparse and noisier for the remaining minority groups, with estimates for Asians available only for UI, Medicare, and Medicaid and estimates for American Indians available only for Medicaid and the IHS.

Nevertheless, false negative and false positive rates appear to be higher for these groups, consistent with the more general patterns of covariate-adjusted errors being higher for minorities.

Overall, the covariate-adjusted error rates confirm the pattern of systematically higher survey error for minorities. After conditioning on covariates, the effect sizes tend to be slightly smaller and vary less across programs. This finding suggests that some of the differences between groups and programs are explained by demographics. However, large differences remain for all programs even after conditioning on many covariates.

It is worth noting that the marginal effects we document for the race/ethnicity indicators are large compared to those for other predictors of misreporting. For example, Celhay, Meyer, and Mittag (2021) show that the difference in the false negative rates between Black and white individuals is consistently larger than (and up to almost 4 times as large as) the increase in false negatives when moving from income at the poverty line to twice the poverty line. The racial and ethnic differences in survey accuracy they find vastly exceed differences by education or household structure and are of similar magnitude or larger than differences by gender or elderly status.[4] Consequently, the marginal effects of the minority dummies tend to be among the strongest predictors of survey error in the models that Celhay, Meyer and Mittag (2021) estimate, only falling short of other program receipt indicators and a few demographic characteristics.[5]

In conclusion, our findings in this section paint a striking picture of lower survey accuracy among racial and ethnic minorities. Race and ethnicity reliably predict survey error, with differences large enough to skew important conclusions. These patterns hold for multiple programs, surveys, and years and also persist across different minority groups (including Black individuals, Hispanics, and Asians), suggesting that our main household surveys poorly serve minorities by providing less accurate information about their incomes and program receipt.

---

[4] Researchers have been concerned about measurement error differences in program receipt by age (Haider, Jacknowitz, and Schoeni 2003). Here, we see that differences by race and ethnicity are much larger than differences by age.

[5] These variables include indicators for household income being above 10x the poverty line (which systematically exceeds race and ethnicity effects) and whether or not the household head is disabled or employed (which exceeds race and ethnicity effects in some cases).

4. Consequences of Survey Error

The patterns of survey error documented in Section 3 raise concerns about the extent to which survey estimates accurately reflect key statistics of interest for racial and ethnic minorities. The nature of measurement error in surveys is non-classical, given that survey error is correlated with key covariates. Only a few general results exist for the consequences of such non-classical measurement error, and they usually depend on the correlation structure of the data and errors. Here, we are particularly interested in differences between groups, so most statistics of interest can be recovered as coefficients on binary explanatory variables. This simplification makes it possible to gain some insights for common types of statistics from theory, but most of what we know about bias stems from empirical assessments using linked data. In this section, we discuss the empirical evidence on bias from survey error in common statistics such as rates of program receipt, determinants of program take-up, and poverty rates. Our discussion shows that survey errors can introduce substantial biases in estimates of racial or ethnic differences, and we highlight conditions under which researchers using survey data should be particularly cautious.

4.1 Bias in Estimates of Program Receipt and Average Amounts Among True Recipients

For differences such as raw gaps in average dollars received or receipt rates, the nature of the biases is simple and determined by the measures of error we summarize above. For average differences in continuous variables (e.g., average benefit amounts received), the bias in the survey estimate for a group is simply the difference between average survey reports and administrative values for that group. If this difference in survey error between groups is known, then the relative bias is straightforward to assess. Unfortunately, only a few papers report such estimates (see, e.g., Shantz and Fox 2018, Meyer et al. 2023). Similarly, for binary variables like program receipt, the bias in statistics is simply the difference between average misclassification rates – equivalent to the false positive rate weighted by the share of non-recipients net of the false negative rate weighted by the share of true recipients.

Table 4 reports measures of program receipt by race and ethnicity, both in terms of the number of recipients and average dollars received (conditional on receipt). For each subgroup, we report the statistic of interest according to the survey data and the linked administrative variable, as well the statistical significance of the difference (if reported in the study). Due to the severe net underreporting of program receipt documented in the previous literature, receipt rates are markedly higher using the administrative values than the survey values. This finding holds for all programs, samples, and minority groups studied. In line with the higher rates of survey error we document above, the increases in receipt rates when correcting for survey error tend to be largest for Black and Hispanic individuals. SNAP and pension receipt rates approximately double for all minority groups, and TANF receipt rates increase by approximately 50% for Black and Hispanic individuals and about a quarter for Asians. In addition, UI receipt rates are higher by around 75% for Black and Hispanic individuals (compared to approximately 50% for Asians). For white individuals, the biases in receipt rates are smaller in fractional terms, with SNAP and pension receipt less than doubling and TANF and UI receipt increasing by about a third. While Tables 2 and 3 showed that minority groups have larger errors in terms of both false positives and false negatives, the results here suggest that they translate to greater net underreporting of program receipt. Thereby, survey error leads us to understate both the degree to which the safety net supports minorities and its contribution to reducing inequality.

Several studies also calculate average benefit amounts among true reporting recipients (i.e., those reporting receipt in both the survey and administrative data) and find more muted evidence of these patterns. True reporting recipients have higher average amounts of SNAP and UI and lower average amounts of TANF when replacing survey values with administrative data. For all recipients, the fractional gaps in dollar biases tend to be larger for minority groups than whites (although these biases are positive for SNAP and UI and negative for TANF). While these results could be indicative of real intensive margin differences in dollars reported, they could also reflect differential selection of survey individuals into reporting by race/ethnicity. Fewer recipients report receipt among minorities, so those who are correctly recorded as recipients in both sources may be differentially selected in terms of reporting accuracy among minorities. However, we can conclude overall that the survey data greatly understate the access of racial and ethnic minorities to the safety net, with clear differences attributable to extensive margin variation in reporting rates.

4.2 Bias in Racial and Ethnic Predictors of Program Receipt

The raw differences in program means by race and ethnicity are what are used in many contexts to characterize disparities across groups. However, several studies go beyond raw means and account for differences in demographic or other characteristics between individuals when analyzing differences in receipt rates by race or ethnicity. In binary choice models of program receipt differences by race or ethnicity, there is both an attenuation of estimated effects due to mismeasurement of the dependent variable and an additional bias due to correlation of the explanatory variable with the error rates. We first discuss the main patterns in past results, and then at the end of this subsection discuss this econometric relationships that both explain the patterns and provide predictions for future analyses of racial and ethnic differences.

To examine whether the patterns for raw differences hold for estimated differences in receipt after controlling for other factors, Table 5 reports empirical estimates of the effects of race and ethnicity indicators on program receipt from probit models. For each minority group, the table reports the average marginal effect of the respective minority indicator in two receipt models – one using survey-reported receipt and the other using receipt from administrative data – as well as the level at which the difference in estimates between the two models is statistically significant. A useful aspect of these models is that they include covariates for other demographic characteristics.

The results for Black and Hispanic individuals are remarkably uniform across datasets and samples. As a starting point, the SNAP and TANF receipt rates according to the survey data tend to be higher for Black and Hispanic individuals than for whites, with the magnitudes of these coefficients varying across samples. The results are more mixed for UI, where we tend to observe survey-only receipt rates being not significantly different for Black individuals, higher and lower for Hispanic individuals, and lower for Asian individuals relative to whites. In comparing differences between the survey and administrative data, we find that the survey data understate receipt rates among all subgroups, even holding constant other covariates. The magnitudes of these biases are large and important: the covariate-adjusted Black-White and Hispanic-White gaps in the probability of program receipt typically more than double after correcting for survey error, and these differences are statistically significant for the most part. For SNAP, the differences in receipt rates (relative to whites) increase by 4 to 8 percentage points after replacing survey values with administrative values. The differences are of the same magnitude for cash assistance (TANF and GA combined) at 4 to 6 percentage points. They tend to be a bit smaller for UI, but still reach up to 5 percentage points.

However, survey data can sometimes get the sign and/or statistical significance of the difference wrong. This is particularly for the analyses of UI (Meyer et al. 2023), where the survey data tend to suggest that minorities are no more likely to receive UI payments than whites, conditional on being eligible. When using administrative values instead, Black individuals have covariate-adjusted receipt rates that are higher by up to 7.5 percentage points in the CPS and 5.7 percentage points in the SIPP, relative to white individuals. Hispanic individuals also tend to have significantly higher UI receipt rates (by up to 3.4 percentage points in the CPS and 8.9 percentage points in the SIPP) after substituting administrative values, with the few cases of lower receipt rates not statistically significant at conventional levels. Asians tend to have lower rates of UI take-up than white individuals using either survey or administrative data, but the overall tendency is still for survey data to understate receipt rates relative to the administrative data.

At a simple level, when analyzing means like receipt rates for a particular group, the presence of false positives and false negatives can potentially cancel out biases from both overreporting and underreporting. When analyzing *differences* in receipt rates, however, the joint presence of false positives and negatives serves to attenuate these differences across subgroups. This result comes from Meyer and Mittag (2017), who discuss the general bias from misclassification of a binary dependent variable in modeling the determinants of program receipt or take-up. They find that the bias in the coefficient of a predictor (such as a dummy variable for non-white) is proportional to the sum of the false positive and false negative rates for the full sample. Consequently, even if there are no differences in error rates across subgroups, the differences in receipt rates would still be biased towards zero in the presence of overall error. When the correlation between the predictor and the error is positive (e.g., non-whites have higher error rates), then that leads to a further negative bias. This theoretical framework helps to rationalize the empirical results we document: the differences in receipt rates between whites and minority groups are understated simply in the presence of overall error, but they are further understated given the higher error rates associated with racial and ethnic minorities.

Beyond binary choice models, survey error can also lead to measurement error in the dependent variable of multivariate linear regressions (such as continuous amounts of income). If the dependent variable is affected by measurement error that differs across subgroups, then we may obtain biased estimates of regressors corresponding to such groups (e.g., defined by race and ethnicity). To our knowledge, there is little empirical evidence on conditional differences in the reporting of continuous variables by minorities or estimates of the corresponding bias. More generally, it can be difficult to predict the sign and size of biases, even when analytic results and empirical evidence on the errors are available. This is because biases depend on the distribution of other variables and their relation to the error and variable of interest. This problem is greatly amplified for statistics that depend on the mismeasured variable in nonlinear ways, including simple descriptive statistics such as poverty rates.

4.3 Bias in Poverty Rates and Other Statistics

The impact of survey error on poverty rates depends not only on whether income misreporting is more severe closer to the poverty line but also on the distribution of income itself (i.e., what share of households have incomes close to the poverty line). This complex relationship may also differ by race and ethnicity. Consequently, it is impossible to predict how patterns of survey error translate to biases in poverty rates, even for the specific samples for which we have extensive misreporting information from linkage studies.

That being said, some empirical evidence exists on the implications of survey misreporting for poverty rates. Table 6 reports poverty rates by race and ethnicity that prior studies have calculated separately using survey data and administrative measures, as well as the percentage change when correcting for survey error. We do this for the poverty rate and when available for the share with incomes below various multiples of the poverty line, including 50%, 125%, 150%, and 200%. Poverty measures are lower for all racial and ethnic subgroups after correcting for survey error, which is not surprising given that the surveys are known to miss so much program receipt (Meyer and Mittag 2019a). In most cases, the fractional reductions are larger for Black and Hispanic individuals than for white individuals, although the patterns vary considerably. When using the administrative data to correct for misreporting of SNAP or SNAP and TANF together, we find reductions in poverty for Black and Hispanic individuals of 6-9% relative to 4% for whites (Fox et al. 2017), 4% relative to 2% for whites (Shantz and Fox 2018), and 5-6% relative to 2% for Stevens et al. (2018). The one exception here is Rothbaum et al. (2021), who find reductions in poverty of 1-3% for Black and Hispanic individuals relative to 4% for whites.

The effects are more pronounced in the studies correcting simultaneously for survey error in multiple income sources, with some papers finding poverty measures to be two or three times higher according to the survey data. Meyer and Wu (2023), for example, find that correcting for measurement error in multiple income sources leads to reductions in poverty for Black and Hispanic individuals of 70-96% relative to 66% for whites. The larger fractional declines for Black and Hispanic subgroups persist when looking at multiples of the poverty line. Dushi and Trenkamp (2021), in bringing in multiple administrative data sources for the elderly, find that poverty rates decline by 43% for Black individuals relative to 25% for whites. The fractional decline for Hispanics, however, is only 6%. Finally, Bee and Mitchell (2017) also bring in multiple sources of administrative data to correct measurement error for the elderly. They find a slightly different pattern, with poverty rates falling by 38% for white individuals, compared to 31% for Black individuals and 9% for Hispanics. This result is consistent with the higher false negative rates they find for pensions among whites.

In summary, the percentage reductions in poverty due to the administrative records are larger for Black individuals (and to a lesser extent for Hispanic individuals) than for white individuals in most studies examined. These results are a direct implication of program receipt – and thus the poverty reduction effects of programs – being more understated for Black and Hispanic individuals. In contrast, the fractional reductions in poverty tend to be smaller for Asians, for whom differences in error and receipt rates tend to be smaller, relative to white individuals in the reviewed studies. Overall, these results provide further evidence for the extent to which survey data understate how well the safety net serves racial and ethnic minorities – particularly Black and Hispanic individuals.

5. Discussion of Paths Forward

In this paper, we consolidate evidence from a series of data linkage studies documenting high survey error rates for income sources that significantly and substantially differ between racial and ethnic groups. While few studies explicitly discuss the differences they find by race and ethnicity, their results paint a clear and alarming picture of lower survey accuracy for minorities. These patterns hold across programs and surveys and persist even after controlling for differences in demographics. The errors in program receipt lead to large biases in common estimates. For simple statistics such as receipt rates, the error rates we document allow researchers to gauge the size of the bias. However, for more complicated statistics such as regression coefficients or poverty rates, the biases quickly become intractable even when information from linked data is available. These biases are sizeable and systematically distort our understanding of the well-being of racial and ethnic minorities, as well as differences between racial and ethnic groups.

While we unearth a surprising number of estimates by race and ethnicity that paint a remarkably clear picture of lower survey accuracy, our review leaves many questions open. It would be important to understand to what extent the patterns we find also affect other programs and other survey variables. Our review focuses on errors in program receipt, since these variables are easier to validate than other variables such as earnings or education for which there is no readily available measure of truth.[6] Yet, to fully understand how differences in survey accuracy skew our estimates of differences between racial and ethnic groups, we also need to know whether the patterns of severe survey error extend to these variables. There is some evidence that inaccuracies in one dimension of survey error predict inaccuracies in other dimensions, suggesting that the pronounced errors we document extend to other variables in some form.[7]

The studies we review only examine the accuracy of a given sample, but to understand survey accuracy more comprehensively requires understanding how other aspects of survey accuracy like coverage and weighting adjustments differ across groups. To fully comprehend how survey accuracy differs for minorities, we also need more research into whether other the effect of other predictors of survey error such as income or education differ between racial and ethnic groups. To our knowledge, no paper includes interactions of other predictors of error with race and ethnicity indicators. More evidence is also needed on joint misreporting (i.e., whether survey error is correlated across different questions), which can be crucial for further understanding the biases from survey error. While our focus in this review is on household surveys administered by the Census Bureau, these patterns likely affect survey data more generally. Several studies have documented concerning rates of misreporting of transfer receipt for other surveys, including the National Health and Nutrition Examination Survey (Kirlin and Wiseman 2014), the ACCESS study (Rosen, McMahon, and Rosenheck 2007), the FoodAPS Survey (Courtemanche, Denteh, and Tchernis 2019, Meyer and Mittag 2019b), and the Health and Retirement Survey (Dushi, Iams, and Trenkamp 2017, Dushi and Trenkamp 2021). These and similar questions are beyond the scope of this review but could be examined in future studies that make use of linked data.

[6] We identified several studies that compare survey reports of earnings to administrative records for subgroups defined by race and ethnicity (see, e.g., Pedace and Bates 2000, Cristia and Schwabish 2007, Bricker and Engelhardt 2008, Kim and Tamborini 2012, Abowd and Stinson 2013, Kim and Tamborini 2014, Chenevert, Klee, and Wilkin 2016, Brummet et al. 2018). While these analyses reveal important discrepancies across subgroups that strongly suggest the presence of differential measurement error, no measure of truth is available for earnings. Therefore, we focus on the clearer case of programs for which measures of truth can be easily gleaned from the administrative records. Black, Sanders, and Taylor (2003) also look at measurement error in the reporting of education in surveys, but they focus on higher education for which they bring in the National Survey of College Graduates.

[7] For example, item non-response has been shown to predict measurement error (Bollinger and David 2001; Celhay, Meyer, and Mittag 2021) and measurement error in one variable appears to predict future measures of the same variable and errors in other variables (Bollinger and David 2005; Celhay, Meyer, and Mittag 2021; Bollinger and Tasseva 2023).

Perhaps the most important question raised by our paper pertains to how we can still make progress on understanding economic disparities by race and ethnicity in the presence of alarming differences in survey accuracy and their corresponding biases for key estimates. The ideal solution would be to improve survey accuracy, but this can be costly and difficult – especially when little is known about the causes of survey error (Bound, Brown, and Mathiowetz 2001, Celhay, Meyer, and Mittag 2024). Differences in survey accuracy could arise from cultural factors, given that trust in the government may be lower among minority groups or that program receipt carries more stigma (see, e.g., Brandon, Isaac, and Laveist 2005, Bajaj and Stanford 2021). While strategies exist to reduce the effects of mistrust and stigma, they may only go so far if the root causes of deep-seated inequities remain present. Differences could also arise from neighborhood characteristics or from the way the surveys are designed and implemented. Such factors may be easier to address. For example, interviewers have been found to vary substantially in their error rates (e.g., Meyer and Mittag 2019b), and the differences in survey error between minorities and whites could either stem from interviewers with higher error rates being more likely to interview minority groups or from match effects between interviewers and respondents. Differences arising from interviewer assignment or match effects could be reduced through changes in assignment and interviewer training.

Linked data can help us better understand these causes and the extent to which they contribute to differences in survey error. For example, the evidence from linkage studies casts doubt on a large role for stigma. While stigma should lead to higher false negative rates and lower false positive rates, we find both rates to be higher for racial and ethnic minorities. Our review also shows that rates of underreporting do not seem to be systematically lower for programs where stigma is likely to be lower, such as UI and private pensions. In addition, Celhay, Meyer, and Mittag (2022) find that reporting accuracy is better in neighborhoods with higher receipt rates, suggesting that stigma may be lower and reporting better in neighborhoods with a higher fraction of minorities with high receipt rates. Yet, linked data can help survey producers to devise pragmatic strategies to reduce survey error. For example, linked data can be used to examine which neighborhood or demographic characteristics predict misreporting and compel interviewers to probe more when survey error appears likely. Linked data can also be used to study the role of interviewers to examine whether there are systematic differences in interviewer assignment or interviewer performance can provide a practical way to substantially increase survey accuracy.

Ultimately, survey data are unlikely to become completely error-free and any improvements are likely to take time. In the short run, researchers will have to take as given the data with their errors, raising the question of how survey users can harness the available information to improve estimates of racial and ethnic disparities. Surveys are increasingly linked to administrative data, which is one way of creating more accurate data on a timely basis. However, linked data are often not accessible to external researchers or may be unavailable for an exact population of interest. Without direct access to the linked data, researchers could still apply corrections to the survey estimates based on information derived from the linked data. For example, linked survey and administrative data allow for the estimation of models of the determinants of survey errors. Making the parameters of these models available to the public would allow survey data users to correct not only unconditional statistics, but also estimates of racial and ethnic differences. For example, several corrections for misclassified dependent variables exist (Hausman, Abrevaya, and Scott-Morten 1998) and have been found to work well when information from administrative data is incorporated (Meyer and Mittag 2017). More generally, estimates can still be improved by imputing or integrating out mismeasured variables (Blackwell, Honaker, and King 2017, Mittag 2019). Such corrections may work reasonably well when parameters from prior years or some states are available (Mittag 2019, Meyer and Mittag 2019b). In order for the corrections to work well for minority subgroups, the models also need to include indicators for these groups of interest. However, this strategy crucially depends on data producers providing the required estimates from the linked data on a timely and systematic basis.

In the absence of any information from linked data (i.e., for estimates based on survey data alone), our overview of the literature sends a clear message of caution for interpreting racial and ethnic differences. Corrections based solely on survey data and aggregate receipt rates have been found to work reasonably well for analyses of unconditional statistics such as SNAP receipt rates (Mittag 2019) but are unlikely to improve estimates of racial and ethnic differences as these corrections tend to replicate differences recorded in surveys. Our discussion above clearly shows that the bias not only depends on the frequency and severity of survey error, but also on the statistic of interest. Thus, even if researchers do not have the information required to correct estimates, linkage studies can still provide them with useful guidance about when increased caution is warranted. A better understanding of the extent and nature of survey error can help researchers judge whether or not the bias is likely to be large enough to affect their substantive conclusions.

But in the presence of error rates as high as the ones we document above, biases can often be large, and our review provides a strong call for caution.

In conclusion, we find high rates of survey error for racial and ethnic minorities. Survey error differs substantially between whites and minorities, leading to large biases in important statistics. From a methodological perspective, these results indicate that more research is needed into the sources of survey error and how survey error and its differences by group can be reduced. Such research would help to improve both the quality of survey data and the ability of researchers to work with error-ridden data. Until these problems are ameliorated or at least better understood, researchers should be cautious when analyzing survey estimates of differences between racial and ethnic groups. Substantively, our findings show that these large surveys do not serve racial and ethnic minorities adequately, as they provide less accurate information about minorities than whites. On the positive side, correcting for differences in survey error shows that the safety net serves minorities much more favorably than the survey estimates we currently rely on suggest.

References

Abowd, John M. and Martha H. Stinson. 2013. “Estimating Measurement Error in Annual Job Earnings: A Comparison of Survey and Administrative Data.” *Review of Economics and Statistics*, 95(5): 1451-1467.

Akee, Randall, Maggie R. Jones, and Sonya R. Porter. 2019. “Race Matters: Income Shares, Income Inequality, and Income Mobility for All U.S. Races.” *Demography*, 56(3): 999-1021.

Aigner, Dennis J. 1973. “Regression with a Binary Independent Variable Subject to Errors of Observation.” *Journal of Econometrics*, 1: 49-60.

Bajaj, Simar Singh and Fatima Cody Stanford. 2021. “Beyond Tuskegee – Vaccine Distrust and Everyday Racism.” *New England Journal of Medicine*, 384:e12.

Bee, Adam, and Joshua Mitchell. 2017. “Do Older Americans Have More Income Than We Think?” SESHD Working Paper 2017-39. Washington, D.C.: U.S. Census Bureau.

Bhaskar, Renuka, James Noon, & Brett J. O’Hara. 2019. “The Errors in Reporting Medicare Coverage: A Comparison of Survey Data and Administrative Records.” *Journal of Aging and Health*, 31(10), 1806-1829.

Bhaskar, Renuka, Rachel Shattuck, and James Noon. 2018. “Reporting of Indian Health Service Coverage in the American Community Survey.” CARRA Working Paper 2018-04. Washington, D.C.: U.S. Census Bureau.

Black, Dan, Seth Sanders, and Lowell Taylor. 2003. “Measurement of Higher Education in the Census and Current Population Survey.” *Journal of the American Statistical Association*, 98(463): 545-554.

Blackwell, Matthew, James Honaker, and Gary King. 2017. “A Unified Approach to Measurement Error and Missing Data: Overview and Applications.” *Sociological Methods & Research*, 46(3): 303-341.

Bollinger, Christopher R. and Martin H. David. 1997. “Modeling Discrete Choice with Response Error: Food Stamp Participation.” *Journal of the American Statistical Association*, 92(439): 827-835.

_____. 2001. “Estimation with Response Error and Nonresponse: Food-Stamp Participation in the SIPP.” *Journal of Business & Economic Statistics*, 19(2): 129-141.

_____. 2005. "I Didn't Tell, and I Won't Tell: Dynamic Response Error in the SIPP." *Journal of Applied Econometrics*, 20(4): 563-569.

Bollinger, Christopher R. and Iva Valentinova Tasseva. 2023. "Income Source Confusion Using the SILC." *Public Opinion Quarterly*, 87(S1): 542-574.

Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. "Measurement Error in Survey Data." In J. J. Heckman and E. Leamer (eds.), *Handbook of Econometrics, Volume 5*, 3705–3843. Amsterdam: Elsevier Science.

Brandon, Dwayne T., Lydia A. Isaac, and Thomas A. LaVeist. 2005. "The Legacy of Tuskegee and Trust in Medical Care: Is Tuskegee Responsible for Race Differences in Mistrust of Medical Care?" *Journal of the National Medical Association*, 97(7): 951-956.

Bricker, Jesse and Gary V. Engelhardt. 2008. "Measurement Error in Earnings Data in the Health and Retirement Study." *Journal of Economic and Social Measurement*, 33(1): 39-61.

Brummet, Quentin, Denise Flanagan-Doyle, Joshua Mitchell, John Voorheis, Laura Erhard, and Brett McBride. 2018. "Investigating the Use of Administrative Records in the Consumer Expenditure Survey." CARRA Working Paper #2018-01. Washington, D.C.: U.S. Census Bureau.

Celhay, Pablo, Bruce D. Meyer, and Nikolas Mittag. 2021. "Errors in Reporting and Imputation of Government Benefits and Their Implications." NBER Working Paper 29184.

_____. 2022. "Stigma in Welfare Programs." NBER Working Paper 30307.

_____. 2024. "What Leads to Measurement Error? Evidence from Reports of Program Participation in Three Surveys." *Journal of Econometrics*, 238(2): 105581.

Chenevert, Rebecca L., Mark A. Klee, and Kelly R. Wilkin. 2016. "Do Imputed Earnings Earn Their Keep? Evaluating SIPP Earnings and Nonresponse with Administrative Records." SESHD Working Paper 2016-18. Washington, D.C.: U.S. Census Bureau.

Chetty, Raj, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. 2020. "Race and Economic Opportunity in the United States: An Intergenerational Perspective." *Quarterly Journal of Economics*, 135(2): 711-783.

Colby, Sandy, Jose Debora, and Misty L. Heggeness. 2016. "How Well Do Individuals Report Supplemental Nutrition Assistance Program (SNAP) Take Up in Household Surveys?." SEHSD Working Paper 2017-03. Washington, D.C.: U.S. Census Bureau.

Cristia, Julian and Jonathan A. Schwabish. 2007. "Measurement Error in the SIPP: Evidence from Matched Administrative Records." Working Paper 2007-03. Washington, D.C.: Congressional Budget Office.

Currie, Janet. 2006. "The Take-Up of Social Benefits." In A. Auerbach, D. Card, and J. Quigley (eds.), *Public Policy and the Income Distribution*, 80-148. New York: Russell Sage Foundation.

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): 202-228.

Davern, Michael, Jacob Alex Klerman, and Jeanette Ziegenfuss. 2007. "Medicaid Under-reporting in the Current Population Survey and One Approach for a Partial Correction." RAND Corporation, WR-532.

Davern, Michael, Jacob Alex Klerman, David K. Baugh, Kathleen Thiede Call, and George D. Greenberg. 2009a. "An Examination of the Medicaid Undercount in the Current Population Survey: Preliminary Results from Record Linking." *Health Services Research*, 44(3): 965-987.

Davern, Michael, Jacob Klerman, Jeanette Ziegenfuss, Victoria Lynch, and George Greenberg. 2009b. "A Partially Corrected Estimate of Medicaid Enrollment and Uninsurance: Results from an Imputational Model Developed off Linked Survey and Administrative Data." *Journal of Economic and Social Measurement*, 34(4): 219-240.

Duncan, Otis Dudley. 1968. "Inheritance of Power or Inheritance of Race?" In Moynihan, D. (ed.), *On Understanding Poverty: Perspectives from the Social Sciences*, 85-110. New York: Basic Books.

Dushi, Irena, Howard Iams, and Brad Trenkamp. 2017. "The Importance of Social Security Benefits to the Income of the Aged Population." *Social Security Bulletin*, 77(2): 1-12.

Dushi, Irena, and Brad Trenkamp. 2021. "Improving the Measurement of Retirement Income of the Aged Population." ORES Working Paper No. 116. Washington, D.C.: Social Security Administration.

Forsythe, Eliza, and Hesong Yang. 2021. "Understanding Disparities in Unemployment Insurance Recipiency." Report to the Department of Labor.

Fox, Liana E., Misty L. Heggeness, Jose Pacas, and Kathryn Stevens. 2017. "Precision in Measurement: Using SNAP Administrative Records to Evaluate Poverty Measurement." SEHSD Working Paper 2017-49. Washington, D.C.: U.S. Census Bureau.

Giefer, Katherine G., Michael D. King, and Veronica L. Roth. 2022. "SNAP Receipt in SIPP: Using Administrative Records to Evaluate Data Quality." SEHSD Working Paper 2022-22. Washington, D.C.: U.S. Census Bureau.

Gould-Werth, Alix, and H. Luke Shaefer. 2012. "Unemployment Insurance Participation by Education and by Race and Ethnicity." *Monthly Labor Review* 135: 28-41.

Haider, Steven J., Alison Jacknowitz, and Robert F. Schoeni. 2003. "Food Stamps and the Elderly: Why is Participation so Low?" *Journal of Human Resources*, 38(S): 1180-1220.

Hausman, Jerry A., Jason Abrevaya, and Fiona M. Scott-Morton. 1998. "Misclassification of the Dependent Variable in a Discrete-Response Setting." *Journal of Econometrics*, 87(2): 239–269.

Huynh, Minh, Kalman Rupp, and James Sears. 2002. "The Assessment of Survey of Income and Program Participation (SIPP) Data using Longitudinal Administrative Records." SIPP Working Paper No. 238. Washington, D.C.: Census Bureau.

Kim, ChangHwan and Christopher R. Tamborini. 2012. "Do Survey Data Estimate Earnings Inequality Correctly? Measurement Errors Among Black and White Male Workers." *Social Forces*, 90: 1157-1181.

_____. 2014. "Response Error in Earnings: An Analysis of the Survey of Income and Program Participation Matched with Administrative Data." *Sociological Methods and Research*, 43(1): 39-72.

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." Working Paper.

Kuka, Elira, and Bryan A. Stuart. 2021. "Racial Inequality in Unemployment Insurance Receipt and Take-Up." NBER Working Paper 29595.

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

Massey, Douglas S. and Nancy A. Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, MA: Harvard University 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): 230-253.

_____. 2021a. "An Empirical Total Survey Error Decomposition using Data Combination." *Journal of Econometrics*, 224(2): 286-305.

_____. 2021b. "Combining Administrative and Survey Data to Improve Income Measurement." In A.Y. Chun, M.D. Larsen, G. Durrant, and J.P. Reiter (eds.), *Administrative Records for Survey Methodology*, 297-322. New York: John Wiley and Sons.

Meyer, Bruce D., Nikolas Mittag, and Robert M. Goerge. 2022. "Errors in Survey Reporting and Imputation and Their Effects on Estimates of Food Stamp Program Participation." *Journal of Human Resources*, 57(5): 1605-1644.

Meyer, Bruce D. and Derek Wu. 2023. "Poverty in the United States." Working Paper.

Meyer, Bruce D., Derek Wu, Matthew Stadnicki, and Patrick Langetieg. 2023. "The Reporting of Unemployment Insurance and Unemployment in Survey and Administrative Sources." Working Paper.

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

Mittag, Nikolas. 2019. "Correcting for Misreporting of Government Benefits." *American Economic Journal: Economic Policy*, 11(2): 142-164.

Moffitt, Robert A. and Peter T. Gottschalk. 2001. "Ethnic and Racial Differences in Welfare Receipt in the United States." In *America Becoming: Racial Trends and Their Consequences*, Vol. 2, Chapter 7, 152-173. Washington, D.C.: The National Academies Press.

Noon, James M., Leticia E. Fernandez, and Sonya R. Porter. 2019. "Response Error and the Medicaid Undercount in the Current Population Survey." *Health Services Research*, 54(1): 34-43.

Pedace, Roberto and Nancy Bates. 2000. "Using Administrative Data to Assess Earnings Reporting Error in the Survey of Income and Program Participation." *Journal of Economic and Social Measurement*, 26: 173-192.

Rosen, Marc I., Thomas J. McMahon, and Robert A. Rosenheck. 2007. "Homeless People Whose Self-Reported SSI/DI Status is Inconsistent with Social Security Administration Records." *Social Security Bulletin*, 67(1): 53-62.

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 1. Washington, D.C.: U.S. Census Bureau.

Rothbaum, Jonathan, Liana Fox, and Kathryn Shantz. 2021. "Fixing Errors in a SNAP: Addressing SNAP Under-reporting to Evaluate Poverty. Working Paper. Washington, D.C.: U.S. Census Bureau.

Shantz, Kathryn and Liana E. Fox. 2018. "Precision in Measurement: Using State Level Supplemental Nutrition Assistance Program and Temporary Assistance for Needy Families Administrative Records and the Transfer Income Model (TRIM3) to Evaluate Poverty Measurement." SEHSD Working Paper 2018-30. Washington, D.C.: U.S. Census Bureau.

Shrider, Emily A., and John Creamer. 2023. "Poverty in the United States: 2022." Current Population Report P60-280. Washington, D.C.: U.S. Census Bureau.

Stevens, Kathryn, Liana E. Fox, and Misty L. Heggeness. 2018. "Precision in Measurement: Using SNAP Administrative Records and the Transfer Income Model (TRIM3) to Evaluate Poverty Measurement." SESHD Working Paper 15. Washington, D.C.: U.S. Census Bureau.

U.S. Census Bureau. 2006. "Design and Methodology: Current Population Survey." Technical Paper 66, U.S. Census Bureau.

_____. 2008. "Survey of Income and Program Participation: User's Guide." Washington, D.C.: U.S. Census Bureau.

_____. 2014. "American Community Survey: Design and Methodology." Washington, D.C.: U.S. Census Bureau.

_____. 2023. "Historical Income Tables: Households; Table H-5. Race and Hispanic Origin of Householder – Households by Median and Mean Income." https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-households.html.

U.S. General Accounting Office (GAO). 2022. "Pandemic Unemployment Assistance: Federal Program Supported Contingent Workers Amid Historic Demand, but DOL Should Examine Racial Disparities in Benefit Receipt" Report to Congressional Committee GAO-22-104438, Washington, D.C.

Wagner, Deborah and Mary Layne. 2014. "The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) Record Linkage Software." CARRA Working Paper 2014-01. Washington, D.C.: U.S. Census Bureau.

Winkler, William E. 2021. "Cleaning and Using Administrative Lists: Enhanced Practices and Computational Algorithms for Record Linkage and Modeling/Editing/Imputation." In A.Y. Chun, M.D. Larsen, G. Durrant, and J.P. Reiter (eds.), *Administrative Records for Survey Methodology*, 297-322. New York: John Wiley and Sons.

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

Table 1: Overview of Studies

Study	Survey Data	Programs	Geography	Main Sample
Panel A: Means-Tested Transfer Programs				
Colby, Debora, Heggeness (2016)	2010-3 SIPP	SNAP	IL, MD, VA	All individuals
Fox, Heggeness, Pacas, Stevens (2017)	2010-6 CPS	SNAP	IL, MD, OR, VA	All individuals, no imputed benefit amounts
Shantz and Fox (2018)	2010-6 CPS	SNAP, TANF	AZ, MD, ND, TN, ID, MI, VA	All individuals, not whole imputed, no imputed benefit amounts
Stevens, Fox, Heggeness (2018)	2010-5 CPS	SNAP	AZ, IL, MD, OR, TN, ID, VA	All individuals, no imputed benefit amounts
Celhay, Meyer, Mittag (2021)	2008-12 ACS, 2008-13 CPS, 2007-12 SIPP	SNAP, TANF+GA	NY	All households
Rothbaum, Fox, Shantz (2021)	2014 CPS	SNAP	AZ, ID, MD, MI, NY, ND, TN, VA	All individuals
Giefer, King, Roth (2022)	2014-20 SIPP	SNAP	12 States	All individuals (person-months), no imputed observations
Meyer, Mittag, Goerge (2022)	2001 ACS, 2002-5 CPS, 2001-5 SIPP	SNAP	IL, MD	All households below 2xFPL, at least one household member age ≥ 16
Panel B: Health Insurance Programs				
Klerman, Ringel, Roth (2005)	1990-2000 CPS	Medicaid, Welfare	CA	All CA residents aged 15-65, no imputed benefit amounts, no imputed demographic variables
Davern, Klerman, Baugh, Call, Greenberg (2009a)	2001-2 CPS	Medicaid	USA	All individuals, no partial Medicaid benefit enrollees, no SCHIP enrollees
Bhaskar, Shattuck, Noon (2018)	2014 ACS	IHS	USA	All individuals
Bhaskar, Noon, O'Hara (2019)	2014 CPS	Medicare	USA	All individuals age ≥ 65
Noon, Fernandez, Porter (2019)	2011 CPS	Medicaid	USA	All individuals age ≥ 18, no partial Medicaid benefit enrollees, no SCHIP enrollees, no imputed or edited responses
Panel C: Old Age Income and Other Income Sources				
Bee and Mitchell (2017)	2013 CPS	OASDI, Pensions	USA	All individuals age ≥ 65
Brummet, Flanagan-Doyle, Mitchell, Voorheis, Erhard, McBride (2018)	2013-14 CE	Retirement Income	USA	All individuals
Dushi and Trenkamp (2021)	2016 CPS	Income (in retirement)	USA	All individuals age ≥ 65
Meyer, Wu, Stadnicki, Langetieg (2023)	2011 CPS, 2010 SIPP	UI	USA	All individuals age ≥ 18, not whole imputed
Meyer and Wu (2023)	2017 CPS	Retirement Income, Social Security, SSI, SNAP, Housing Assistance, TANF	USA, 23 states for SNAP, 18 states for TANF	All individuals, no whole imputed SPM units

Notes: This table provides an overview of the studies surveyed in this paper, including the data used, programs examined and key sample restrictions. The original papers provide additional detail on the samples used and key definitions, which vary between studies.

Table 2: False Negative and False Positive Rates by Race and Ethnicity

Study	Program	Sample	White		Black		Hispanic		Asian	
			F. Neg (1)	*F. Pos* (2)	*F. Neg* (3)	*F. Pos* (4)	*F. Neg* (5)	*F. Pos* (6)	*F. Neg* (7)	*F. Pos* (8)
Colby et al. (2016)	SNAP	2010-13 SIPP	15.9%	1.5%	19.5%	8.0%	20.0%	4.9%	-	-
Giefer et al. (2022)	SNAP	2014-20 SIPP	24.5%	0.4%	25.1%	1.5%	-	-	-	-
Meyer et al. (2023)	UI	2011 CPS	33.2%	0.3%	49.0%	0.8%	51.5%	0.7%	43.3%	0.5%
	UI	2010 SIPP	33.3%	0.9%	50.0%	1.6%	42.2%	1.3%	45.7%	1.1%
Klerman et al. (2005)	Medicaid	1990-2000 CPS	-	-	29.0%	5.0%	33.0%	5.0%	-	-
	Welfare	1990-2000 CPS	-	-	42.0%	2.0%	59.0%	2.0%	-	-
Davern et al. (2009a)	Medicaid	2002 CPS	42.8%	2.3%	42.9%	5.6%	43.7%	4.5%	47.1%	2.5%
Bhaskar et al. (2018)	IHS	2014 ACS	-	-	-	-	42.9%	-	-	-
Bhaskar et al. (2019)	Medicare	2014 CPS	3.8%	54.5%	5.4%	55.9%	5.3%	41.4%	7.9%	50.8%
Bee and Mitchell (2017)	OASDI	2013 CPS	5.6%	5.1%	11.5%	7.2%	11.2%	8.4%	10.2%	11.8%
	Pensions	2013 CPS	29.8%	3.5%	26.0%	5.0%	20.9%	3.5%	23.3%	5.0%
	Pensions	2013 ACS	28.3%	3.2%	21.5%	5.4%	17.4%	4.8%	18.2%	4.0%

Notes: This table summarizes estimates of the rate at which true recipients are recorded as non-recipients in the data (columns labeled *F. Neg*) and the rate at which true non-recipients are recorded as recipients in the survey data (columns labeled *F. Pos*) for studies that report these statistics separately for subgroups defined by race and ethnicity. See Table 1 for details on the studies and their methodology. Error rates in Colby et al. (2016), Giefer et al. (2022), and Bhaskar et al. (2018) are reported in a different format and were converted to conform to our definition. Hispanic is mutually exclusive with other race categories in Colby et al. (2016), Giefer et al. (2022), Meyer et al. (2023), Bhaskar et al. (2019), and Bee and Mitchell (2017), while it overlaps with other race categories in Klerman et al. (2005), Davern et al. (2009a), and Bhaskar et al. (2018).

Table 3: Differences in Covariate-Adjusted Error Rates Between Minorities and Whites

Study	Program	Sample	Black		Hispanic		Asian		AI/AIAN	
			F. Neg	*F. Pos*	*F. Neg*	*F. Pos*	*F. Neg*	*F. Pos*	*F. Neg*	*F. Pos*
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: LPM Coefficients and Probit Average Partial Effects										
Fox et al. (2017)	SNAP	2010-16 CPS	0.0680***	-	0.0390	-	-0.0420	-	-	-
Celhay et al. (2021)	SNAP	2008-12 ACS	0.0678***	0.0098***	0.0410***	0.0085***	-	-	-	-
		2008-13 CPS	0.0958***	0.0116***	0.0554***	0.0139***	-	-	-	-
		2007-12 SIPP	0.0841***	0.0137***	0.0382**	0.0044	-	-	-	-
	TANF+GA	2008-12 ACS	0.0877***	0.0005	0.0467***	-0.0028***	-	-	-	-
		2008-13 CPS	0.0683	0.0015	0.0744*	-0.0017	-	-	-	-
		2007-12 SIPP	0.0800*	0.0059***	0.0508	-0.0027*	-	-	-	-
Meyer et al. (2022)	SNAP	2001 ACS, IL	0.0897**	0.0239***	-	-	-	-	-	-
		2001 ACS, MD	0.1110***	0.0082	-	-	-	-	-	-
		2002-5 CPS, IL	0.0503	-0.0046	-	-	-	-	-	-
		2002-5 CPS, MD	0.0509	-0.0094	-	-	-	-	-	-
		2001-5 SIPP, IL+MD	0.0672*	0.0207**	-	-	-	-	-	-
Meyer et al. (2023)	UI	2011 CPS	0.0945***	0.0036***	0.1190***	0.0037***	0.1120***	0.0031**	-	-
		2010 SIPP	0.1100***	0.0023**	0.0583***	0.0015	0.0866**	0.0012	-	-
Panel B: Logistic Regression Coefficients										
Colby et al. (2016)	SNAP	2010-13 SIPP (2008 panel)	0.1480	1.5300***	0.0860	0.5770**	-	-	-	-
Klerman et al. (2005)	Medicaid	1990-2000 CPS	0.3670	0.9270	0.7070	N/A	-	-	-	-
	Welfare	1990-2000 CPS	N/A	N/A	0.3470	N/A	-	-	-	-
Bhaskar et al. (2018)	IHS	2014 ACS, AIAN only	-	-	0.4300***	-	-	-	0.0900	-
		2014 ACS, AIAN only, age>=25	-	-	0.4400**	-	-	-	0.0100	-
		2014 ACS, non-AIAN	-	-	0.1900	-	-	-	-2.1200***	-
Bhaskar et al. (2019)	Medicare	2014 CPS	0.5200**	-0.1500	0.4000*	-1.2000*	0.3900	-0.4600	-	-
Noon et al. (2019)	Medicaid	2011 CPS	0.1900**	0.3000	0.2900***	0.8300***	0.3000*	0.9600**	0.2900	0.6300

Notes: This table reports estimated effects of minority indicators from models where the binary dependent variable indicates not receiving a program in the survey conditional on being a true recipient (columns labeled *F. Neg*) and models where the dependent variable indicates receiving a program in the survey conditional on being a true non-recipient (columns labeled *F. Pos*). The models vary in the sample and covariates used; see Table 1 and the original studies for additional details. Logistic coefficients for Colby et al. (2016), Bhaskar et al. (2018), Bhaskar et al. (2019), and Noon et al. (2019) are reported as odds ratios in the published papers and were converted to logistic regression coefficients to facilitate comparability. Hispanic is mutually exclusive with other race categories in Celhay et al. (2021), Meyer et al. (2023), Colby et al. (2016), and Bhaskar et al. (2019), while it overlaps with other race categories in Klerman et al. (2005), Bhaskar et al. (2018), and Noon et al. (2019).

Table 4: Bias in Estimates of Program Receipt and Average Amounts for True Reporting Recipients

Study	Program	Sample	Statistic	White			Black			Hispanic			Asian		
				Survey	*Admin*	*Δ Sig?*	*Survey*	*Admin*	*Δ Sig?*	*Survey*	*Admin*	*Δ Sig?*	*Survey*	*Admin*	*Δ Sig?*
				(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Fox et al. (2017)	SNAP	2010-6 CPS	Receipt rate	6.0%	11.0%	-	17.0%	38.0%	-	16.0%	32.0%	-	5.0%	10.0%	-
			Average amount	$3,324	$3,660	-	$3,648	$4,080	-	$3,576	$4,092	-	$3,708	$3,912	-
Shantz and Fox (2018)	SNAP	2010-6 CPS	Receipt rate	8.5%	14.6%	***	21.8%	38.8%	***	22.9%	37.7%	***	5.4%	10.1%	***
			Average amount	$3,409	$3,607	***	$3,825	$4,172	***	$3,528	$3,638		$2,803	$3,286	*
	TANF	2010-6 CPS	Receipt rate	0.9%	1.2%	***	4.6%	7.1%	***	2.0%	3.0%	***	0.4%	0.5%	
			Average amount	$2,567	$2,196	**	$3,420	$2,640	***	$3,006	$1,455	**	*S*	*S*	*S*
Meyer et al. (2023)	UI	2011 CPS	Receipt rate	4.2%	5.9%	-	4.8%	8.3%	-	3.4%	6.2%	-	2.9%	4.5%	-
			Average amount	$8,065	$8,808	-	$6,917	$7,990	-	$7,556	$8,803	-	$9,421	$10,990	-
		2010 SIPP	Receipt rate	4.7%	6.1%	-	5.1%	8.2%	-	5.8%	8.8%	-	3.7%	5.3%	-
			Average amount	$7,194	$9,301	-	$5,762	$8,573	-	$5,963	$8,562	-	$8,023	$9,664	-
Brummet et al. (2018)	Pensions	2013-4 CE	Receipt rate	40.3%	74.6%	-	31.6%	65.7%	-	23.6%	48.3%	-	21.0%	45.3%	-

Notes: This table reports receipt rates and average annual amounts received for studies that provide these statistics separately for subgroups defined by race and ethnicity. For each subgroup, the first column provides the estimate according to the survey data and the second column provides the same statistic according to the linked administrative variable. The third column for each group indicates the significance level of the difference between the two estimates (* 10%, ** 5%, *** 1%, with blank indicating not significant at conventional levels) if such a test was performed and "-" otherwise. See Table 1 for additional details on the data and samples used. *S* indicates suppressed results. Brummet et al. (2018) also report results for AIANs and Pacific Islanders that we omit in this paper. Average amounts are conditional on receiving benefits in both the survey and administrative data (i.e., being a true reporting recipient). As a result, differences in average amounts reflect errors on the intensive margin, while differences in receipt rates reflect errors on the extensive margin. Average amounts in Fox et al. (2017) were originally reported at the monthly level and we have multiplied them by 12 to represent annual levels. Rates of program receipt and average amounts from Meyer et al. (2023) are reported in a different format and were converted to conform to our definition. Hispanic is mutually exclusive with other race categories in Fox et al. (2017), Shantz and Fox (2018), and Meyer et al. (2023), while it overlaps with other race categories in Brummet et al. (2018).

Table 5: Bias in Marginal Effect of Minority Variable in Models of Program Receipt

Study	Program	Sample	Black			Hispanic			Asian		
			Survey	*Admin*	*Δ Sig?*	*Survey*	*Admin*	*Δ*	*Survey*	*Admin*	*Δ Sig?*
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Celhay et al. (2021)	SNAP	2008-12 ACS	0.1098***	0.1708***	***	0.0987***	0.1405***	***	-	-	-
		2008-13 CPS	0.0291*	0.1075***	***	0.0743***	0.1445***	***	-	-	-
		2007-12 SIPP	0.0759***	0.1232***	***	0.0697***	0.1096***	*	-	-	-
	TANF+GA	2008-12 ACS	0.0177***	0.0753***	***	0.0114***	0.0579***	***	-	-	-
		2008-13 CPS	0.0204**	0.0835***	***	0.0122	0.0761***	***	-	-	-
		2007-12 SIPP	0.0258***	0.0647***	***	-0.0011	0.0286***	***	-	-	-
Meyer et al. (2022)	SNAP	2001 ACS, IL	0.0380**	0.0801***	***	-	-	-	-	-	-
		2001 ACS, MD	-0.0055	0.0355	**	-	-	-	-	-	-
		2002-5 CPS, IL	0.0211	0.0762***	***	-	-	-	-	-	-
		2002-5 CPS, MD	-0.0048	0.0118		-	-	-	-	-	-
		2001-5 SIPP, IL+MD	0.0767***	0.1069***	***	-	-	-	-	-	-
Meyer et al. (2023)	UI	2011 CPS, all	0.0060	0.0280***	***	-0.0020	0.0140***	***	-0.0090	-0.0090**	
		2011 CPS, svy eligible	0.0060	0.0310	*	-0.0390**	-0.0240		-0.0860***	-0.0830**	
		2011 CPS, adm eligible	0.0230**	0.0750***	***	-0.0220***	0.0340***	***	-0.0350***	-0.0150	*
		2011 CPS, comb eligible	-0.0080	0.0040		-0.0570***	-0.0250	**	-0.0810**	-0.0540*	
		2010 SIPP, all	-0.0030	0.0170***	***	0.0070*	0.0240***	***	-0.0200***	-0.0170***	
		2010 SIPP, svy eligible	0.0250	0.0230		0.0320	0.0790***	**	-0.0480	-0.0700*	
		2010 SIPP, adm eligible	-0.0070	0.0570***	***	0.0340**	0.0890***	***	-0.0480**	-0.0500**	
		2010 SIPP, comb eligible	0.0150	0.0390*		0.0270	0.0740***	**	-0.0790**	-0.1020**	

Notes: This table reports the estimated effects of minority indicators in models of program receipt or take-up (i.e., from models where program receipt is the dependent variable). For each minority subgroup, the first column provides the estimate using the survey measure of receipt as the dependent variable and the second column provides the same statistic using the administrative measure of receipt as the dependent variable. The third column for each group indicates the significance level of the difference between the two estimates (* 10%, ** 5%, *** 1%, with blank indicating not significant at conventional levels) if such a test was performed and "-" otherwise. Table 1 provides additional detail on the studies and the data and samples used. For Meyer et al. (2023), we report estimates in each survey for four samples: 1) "all" refers to all individuals aged 18+, 2) "svy eligible" refers to those simulated to be eligible for UI using survey information only, 3) "adm eligible" refers to those simulated to be eligible for UI using administrative information only, and 4) "comb eligible" refers to those simulated to be eligible for UI using a combination of survey and administrative information. All estimates are average partial effects from Probit models. Hispanic is mutually exclusive with other race categories in Celhay et al. (2021), and Meyer et al. (2023), whereas Meyer et al. (2022) report results for non-white individuals whom we classify under the "Black" race/ethnicity category.

Table 6: Bias in Poverty Rates (%)

Study	Program	Sample	White				Black				Hispanic				Asian			
			Svy	*Admin*	*% Diff*	*Δ Sig?*	*Svy*	*Admin*	*% Diff*	*Δ Sig?*	*Svy*	*Admin*	*% Diff*	*Δ Sig?*	*Svy*	*Admin*	*% Diff*	*Δ Sig?*
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Fox et al. (2017)	SNAP	2010-6 CPS	8.6	8.3	-4%	*	18.8	17.8	-6%	*	21.4	19.7	-9%	*	12.9	12.4	-4%	*
Shantz and Fox (2018)	SNAP & TANF	2010-6 CPS	9.4	9.2	-2%	***	20.2	19.4	-4%	***	23.3	22.3	-4%	***	13.4	12.9	-4%	*
Stevens et al. (2018)	SNAP	2010-5 CPS	9.0	8.8	-2%	*	17.9	17.0	-5%	*	22.3	21.1	-6%	*	12.9	12.8	-1%	*
Rothbaum et al. (2021)	SNAP	2014 CPS	10.0	9.6	-4%	***	23.7	23.4	-1%		24.9	24.1	-3%	*	15.8	15.8	0%	
Bee and Mitchell (2017)	Various	2013 CPS	6.6	4.8	-38%	***	18.4	14.0	-31%	***	20.9	19.1	-9%		12.4	11.7	-6%	
Dushi and Trenkamp (2021)	Various	2016 CPS	7.5	6.0	-25%	-	18.4	12.9	-43%	-	17.5	16.5	-6%	-	-	-	-	-
…<1.25x FPL	Various	2016 CPS	12.1	9.5	-27%	-	26.1	19.5	-34%	-	25.7	24.2	-6%	-	-	-	-	-
Meyer and Wu (2023)	Various	2017 CPS	6.4	3.9	-66%	-	14.7	7.5	-96%	-	13.6	8.0	-70%	-	8.3	5.7	-46%	-
…<0.5x FPL	Various	2017 CPS	2.5	1.1	-122%	-	5.2	1.6	-217%	-	4.4	1.8	-140%	-	4.2	2.3	-80%	-
…<1.5x FPL	Various	2017 CPS	14.7	10.4	-41%	-	32.7	20.5	-59%	-	34.5	22.6	-52%	-	16.4	12.4	-32%	-
…<2x FPL	Various	2017 CPS	25.7	20.0	-28%	-	49.4	38.5	-28%	-	52.8	41.0	-29%	-	26.5	22.0	-20%	-

Notes: This table reports estimated poverty rates from papers that estimate poverty rates (or variants of poverty rates) separately for subgroups defined by race and ethnicity. For each subgroup, the first column provides the estimate using survey measures only and the second column provides estimates after replacing survey measures of the relevant income sources with the linked administrative variables. The third column for each subgroup indicates the fractional difference between the survey and administrative estimates (using the survey value as the baseline), and the fourth column indicates the significance level of the difference between the two estimates (* 10%, ** 5%, *** 1%, with blank indicating not significant at conventional levels) if such a test was performed and "-" otherwise. Table 1 provides additional detail on the studies, data, and samples used. Unless specified otherwise, all poverty rates are defined as having incomes below 100% the federal poverty line. Bee and Mitchell (2017) correct earnings, Social Security, SSI, interest and dividends, and retirement income. Dushi and Trenkamp (2021) correct earnings, Social Security, SSI, and interest and dividends. Meyer and Wu (2023) correct earnings, interest and dividends, retirement income, Social Security, SSI, AGI and other cash income, tax liabilities and credits, SNAP, housing assistance, and TANF. Hispanic is mutually exclusive with other race categories in all of these studies.