The Anatomy and Evolution of Survey Error*

Bruce D. Meyer
University of Chicago, AEI and NBER

Nikolas Mittag
CERGE-EI

Derek Wu
University of Virginia

Anthony Tatarka
University of Wisconsin

Patrick Langetieg
Internal Revenue Service

August 23, 2026

Abstract

Surveys provide much of what we know in social science and are often the only source of information on private behaviors, attitudes and expectations. Strands of the literature document four broad error sources: survey coverage, unit non-response, item non-response, and measurement error. Yet the absence of comparable error measures across sources and over time has left the relative severity of error sources difficult to assess, limiting efforts by survey producers to improve accuracy and users to interpret survey statistics. We unify the fragmented literature by analyzing and decomposing survey error in an empirical Total Survey Error (TSE) framework, quantifying for the first time every error component on a common scale of bias and absolute error. We apply our approach to comprehensively measure survey error in eighteen income and program receipt variables from the CPS ASEC – one of the oldest and most rigorously tested surveys, and the official source of income and poverty statistics in the U.S. – using linked administrative records over more than two decades. For nearly half of the variables, bias exceeds 40 percent and eliminating all individual-level error would require changing more than 80 percent of the true total. Contrary to the conventional emphasis on non-response, we find that the dominant source of bias is measurement error among respondents who fail to report receipt (false negatives). The largest potential gains therefore come from improving the accuracy of respondent reporting rather than raising response rates. This underreporting also propagates into imputed values for non-respondents, which are drawn from (misreported) responses and are particularly noisy, with positive and negative errors that tend to offset in net terms. Moreover, net bias can be an incomplete guide to data quality, as we find that absolute error has tended to rise over time even when net bias is stable. Overall, our approach provides a new method for measuring and decomposing survey error, clarifying which sources of bias require user corrections and where survey design improvements could be most consequential.

---

*Any views expressed are those of the authors and not those of the U.S. Census Bureau or the Internal Revenue Service. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2597 (CBDRB-FY26-009). We thank Gabby Rocco for superb research assistance, as well as Katharine Abraham, Thomas Crossley, and audiences at the 2026 AEA/ASSA Annual Meetings, 2026 NBER Summer Institute (CRIW), 2026 EEA/ESEM Congress, 2026 Joint Statistical Meetings, 2025 NTA Annual Conference, University College London, University of Chicago, University of Padova, and Upjohn Institute for helpful comments. We are grateful to the Alfred P. Sloan Foundation for supporting this specific research and to the Russell Sage Foundation, Charles Koch Foundation, Menard Family Foundation, National Institute on Aging, and the American Enterprise Institute for their support of the larger Comprehensive Income Dataset project. Mittag is also thankful for financial support from the Czech Science Foundation (grant number 26-21565S).

Emails: Meyer: bdmeyer@uchicago.edu; Mittag: nikolasmittag@posteo.de; Wu: derek.wu@virginia.edu; Tatarka: atatarka@wisc.edu; Langetieg: Patrick.T.Langetieg@irs.gov.

1. Introduction

Surveys are indispensable for policy and research, often serving as the main source of information on measures ranging from concrete quantities like income and expenditures to subjective expectations and opinions. Governments allocate funding, produce official statistics (including on poverty and unemployment), and officially assess legislation using survey data (NAS 2017). Pollsters run surveys to gauge public opinion (Gelman 2021) and predict election outcomes (Atkeson and Alvarez 2018). Researchers rely on surveys for representative samples and are increasingly fielding their own (e.g., Stantcheva 2023; Coibion and Gorodnichenko 2026). And while administrative data are increasingly being used in research and to inform policy, many types of information – such as private behaviors (e.g., drug use, sexual behavior, voting), time use, attitudes, opinions, and expectations – can still only be gleaned from surveys. However, surveys are known to suffer from several main sources of error: households are missed entirely by the survey (coverage/unit non-response error), questions are left unanswered by respondents (item non-response error), and answers from respondents may be inaccurate (measurement error).[1]

Each of these errors has been studied in separate strands of the literature, with their own measures of survey accuracy such as response rates and misreporting rates. This fragmentation creates three fundamental challenges. First, existing measures capture the *prevalence* of a problem rather than its *consequences* for bias or inaccuracy. For example, a decline in response rates may not increase bias if non-response is random. More generally, the distortion from any error source depends on both the prevalence of the problem (e.g., how many survey units are non-respondents) and its per-unit severity (e.g., how inaccurate are their imputations), yet conventional measures capture only the former. Second, it is difficult to compare across error sources, as each source has its own metric. For example, is a 10-point increase in the unit non-response rate more concerning than a 10-point increase in the item non-response rate? Third, it is difficult to track survey error over time, because the relationship between conventional indicators and actual bias and inaccuracy can itself change. Together, these gaps mean that survey quality has been monitored through indicators whose connection to actual bias and inaccuracy has been rarely quantified.

[1] For reviews of the literature, see e.g. Groves et al. (2002) on coverage error, Little and Rubin (2002) on item non-response error, and Bound, Brown, and Mathiowetz (2001) on measurement error. See Bee et al. (2023) for more recent references.

Data users, policymakers, and survey producers need measures of survey error that are comparable across time and source and directly quantify the distortion to survey estimates. Without such measures, survey users do not have the information to optimally choose samples, corrections, and robustness checks to minimize the impact of error on their estimates (e.g., Celhay, Meyer, and Mittag 2021). Policymakers likewise cannot assess whether official statistics are reliable or whether changes over time reflect true changes or changes in error, while survey producers cannot identify which design changes reduce the most error or are cost-effective. For example, non-response error has received outsized emphasis, likely because its prevalence is easy to measure while bias and inaccuracy are not. Response rates routinely serve as headline indicators of data quality, are often tied to official quality standards in government reports (e.g., OMB 2006), and feature prominently in discussions of the modern survey error crisis (Groves 2006, 2011; Brick and Williams 2013). These patterns are mirrored in the research literature, where non-response is the most studied error source and few studies examine multiple sources together (Appendix Figure A.1). But response rates alone provide little guidance for reducing survey error. Like any instrument, a survey requires calibration that pinpoints where error originates, how large it is, and whether it is getting worse.

In this paper, we develop an empirical approach to comprehensively measure survey error in terms of the bias and inaccuracy it causes, thereby quantifying the consequences of error rather than merely its prevalence. By expressing every error source on a common scale, our approach unifies the separate literatures on coverage, non-response, and measurement error, making measures of error directly comparable across sources and allowing the trade-offs between them to be assessed. Our measurement system is based on the Total Survey Error (TSE) framework (Groves 1989; Biemer 2010). However, TSE has largely remained a theoretical taxonomy and is rarely implemented empirically (Groves and Lyberg 2010). Meyer and Mittag (2021) provided a proof of concept for the idea of using data linkage to quantify the components of survey error jointly, decomposing net error for two transfer programs in a single state (see Krohn 2024 for an application). We extend their methods to additional sources of error and beyond bias to absolute error. We show the benefits of the approach by applying it at scale – nationally, across many survey variables, and over more than two decades – to draw a wide range of substantive conclusions.

Specifically, we decompose overall error into coverage error, item non-response error, measurement error, and a new component (supplement non-response error) that captures a type of unit non-response error. We further divide three of the components into false positives, false negatives, and errors in amounts. Our unified framework provides measures that are comparable over time and across income sources and error components. Moreover, our bias and absolute error-based measures jointly capture both the prevalence and per-unit severity of survey error. Our approach provides both an anatomy and an evolution of survey error, documenting which components of the survey are not functioning well and tracking how they have changed. It is not a full diagnosis, however, as the causes of these errors and their cures are left for further research. At the same time, our framework can be used to evaluate design changes to surveys, such as changes in question design or reference periods, distinguishing genuine improvements from new errors that merely offset old ones.

We use our approach to examine the accuracy of the Current Population Survey Annual Social and Economic Supplement (CPS ASEC). The CPS ASEC is the main source of income, poverty, and health insurance statistics in the United States, and it has been fielded for more than sixty years as a supplement to the Basic Monthly CPS – an even longer-running, tried-and-tested household survey that produces the nation's headline labor-force statistics and has served as a model for surveys in the U.S. and abroad. We measure survey error in the share of individuals or households receiving each of ten income sources (pensions, six cash transfers, and three of the largest in-kind transfers) and in the mean dollars received for eight sources, in 2017 and back to 1996 in four cases. Implementing the methods at scale allows us to systematically compare error patterns across income sources and thereby provide a broad assessment of survey data quality. Our measures are simple and inexpensive to produce at high frequency for many variables, making the approach well-suited for continuously and thoroughly monitoring survey accuracy.

Income and program receipt are substantively important as core inputs into measures of economic well-being (such as poverty and inequality) and into analyses of the reach and take-up of government programs. Importantly, these income sources offer close to ideal conditions for measuring survey error. We link the CPS ASEC to administrative tax records and to program records from federal and state agencies, which provide accurate information on receipt and amounts for the entire population at the individual level. This allows us to construct both aggregate benchmarks and individual-level measures of "true" receipt and amounts that match the survey's definition and timing exactly, so we can study survey error in a setting where the truth is essentially known.[2] Such a strategy can extend well beyond income and program receipt, as record linkage can supply measures of truth for a wide range of survey variables encompassing health (hospitalizations, diagnoses, prescriptions), justice (convictions, fines), demographics (age, marital status, migration), and certain forms of consumption and assets (energy use, home and vehicle ownership). Where records align less closely with the survey concept or population, one could still partly implement our validation approach using aggregate data (e.g., referenda for public opinion), partial information (e.g., credit card data for consumption), or noisy data (e.g., cell-phone data for time use). For the many measures that are hardest to validate, such as subjective expectations and opinions, inferences on their errors may need to be drawn from studies like ours.

We start by constructing measures of what the survey intends to yield (i.e., survey targets) for our eighteen measures of income or program receipt and amounts to determine the overall bias in surveys. Our results confirm that the survey misses substantial shares of recipients and dollars received for nearly all income sources and programs, suggesting that survey-based estimates of take-up rates and (low) incomes should be treated with caution (e.g., Meyer, Mok, and Sullivan 2015; Meyer and Mittag 2019). Bias is 40 percent or more for nearly half of the variables in a recent year (2017). We then decompose TSE into the four broad components defined above to understand the sources of overall bias. Here, our results cut against the conventional focus on non-response, as we find that measurement error makes up the majority of the overall bias for seven income sources (for recipients) and eight income sources (for dollars). This finding suggests that replacing survey responses with linked administrative variables or imputing additional receipt, as discussed in Mittag (2019), may be the most direct routes to reducing bias. Our separate estimates of supplement non-response error – where the entire ASEC record is imputed for households that respond to the monthly CPS but not the ASEC – and, to a lesser degree, item non-response error reveal that imputations also cause meaningful negative bias, though less than measurement error.

Next, we further decompose supplement non-response, item non-response, and measurement error into extensive margin errors in receipt (false negatives and false positives) and intensive margin errors in amounts. Failure to report receipt constitutes the dominant source of

[2] We do not argue that the linked measure is entirely error-free, as errors may arise from incomplete administrative records, linkage, an inexact match to survey concepts, and other sources. However, we consider these errors to be sufficiently rare to be of little or no practical relevance in our case, with some exceptions. As we discuss below, this accuracy is not a general feature of linked data, but arises from having high-quality administrative records, reliable linkage, and carefully aligned concepts. We also focus on income sources where these problems are least relevant (as opposed to, for example, earnings, asset income, or business income).

bias. This points to reducing extensive margin underreporting (e.g., by probing more regarding receipt) as a promising strategy to reduce bias. This error source matters doubly, as errors among respondents propagate into the imputations built from their responses, so improving respondent reporting would reduce both measurement and imputation error. One program that deviates from the strong dominant patterns is Supplemental Security Income (SSI), which has a substantial false positive rate due to confusion with Social Security (OASDI).

A central extension of our earlier work is our decomposition of absolute error. Low net error may reflect either accurate responses or offsetting errors. We therefore complement our measures of bias with measures of absolute error, which capture overall inaccuracy and matter especially for multivariate analyses (Meyer and Mittag 2017; Greene 2017). Absolute errors are large, with nearly half of the variables requiring changes exceeding 80 percent of their true totals to eliminate all individual-level error. Offsetting errors are particularly concerning for non-respondents, who add substantial absolute error despite their smaller contribution to net error. As a result, non-response error tends to generate noise while measurement error tends to drive bias. Combined with our finding of generally low coverage error (indicating that weighting adjustments largely restore sample representativeness), this result suggests that excluding non-respondents and reweighting the sample may substantially reduce error.

Finally, we analyze how net and absolute error have changed over time, focusing on the four income sources with administrative data in both 1996 and 2017 (pensions, OASDI, SSI, and housing assistance). These income sources are among the least problematic in terms of net error trends, but the ambiguous patterns in net error mask unambiguous increases in absolute error for every source. This deterioration is driven primarily by rising non-response rates that shift the sample toward observations that are less accurate, compounded by worsening error within some respondent types. This finding shows that our distinction between net and absolute error is crucial, as it separates changes that genuinely improve survey accuracy from changes that merely reduce bias by adding offsetting error. In this sense more is not always better, since design changes that reduce net bias can degrade data quality by making individual observations less accurate. Our results over time also clarify when conventional indicators such as non-response rates track actual bias and when they do not. For supplement non-response, per-unit error is fairly stable over time and across programs, so the rate is roughly proportional to bias. But for item non-response, per-unit error varies substantially over time and across programs, making the rate a less reliable guide.

The breadth of our analysis, spanning eighteen variables over two decades, lets us identify patterns that are likely to hold beyond the variables we examine. Some of our findings, particularly on coverage, unit non-response, and imputation accuracy, are likely informative about other survey variables, since these errors stem from the survey's design rather than the specific question asked. Reporting error likely depends more on the specific measure of interest, but the high rates of underreporting we document are plausible wherever recall difficulty or stigma are at work. And because we find substantial error even in an established survey asking clear, well-tested questions about regular events, our results tell a cautionary tale for the more sensitive, irregular, and subjective topics for which surveys are often our only source of information.

The remainder of the paper proceeds as follows. Section 2 explains how we implement the empirical Total Survey Error decomposition. Section 3 describes our data sources, linkage methods, and construction of survey targets. Section 4 discusses measures of overall error and Section 5 decomposes net and absolute TSE at a point in time (for 2017). Section 6 examines how survey error and its components have changed over time (between 1996 and 2017). Section 7 discusses implications for improving survey statistics, and Section 8 concludes.

2. An Empirical Framework for Decomposing TSE

Our approach rests on the ability to create measures of truth for key survey concepts by linking accurate records for the entire population to survey data. As Meyer and Mittag (2021) showed, such linked data make it possible to estimate TSE and its components. While they provided a proof of concept, here we reintroduce and extend the methods that we apply in later sections. We build on the earlier decomposition which focused on average non-sampling error and decomposed TSE into three error components: generalized coverage error, item non-response error, and measurement error.[3] We additionally estimate an error component arising from unit non-response to the ASEC part of the CPS.

Each component is substantively important and isolates a distinct source of error in the survey. Generalized coverage error captures error from an incorrect survey frame, unit non-response, and the possibly offsetting effects of adjustments such as sampling weight

[3] We abstract from sampling error throughout, since it is small relative to the bias in our large samples. See e.g., Alwin (1991, 2007) for more extensive discussions of sampling error as a TSE component and Meyer and Mittag (2021) for ways to add sampling error to our empirical TSE framework.

modifications.[4] A key part of coverage error arises from unit non-response, which is normally addressed by adjusting the sampling weights such that survey estimates match known population demographics.[5] This is usually the only feasible correction since the survey collects little or no information on the missing units, and it also makes unit non-response notoriously difficult to study (Bee, Gathright and Meyer, 2015). Yet, the two-stage design of the CPS allows us to separately estimate error for a type of unit non-respondent (whom we call supplement non-respondents): those who respond to the Basic CPS but not the ASEC, whose entire supplement the CPS then imputes.[6] Because we know some information about these observations from the Basic CPS, we can link them to administrative data at the individual level to study the nature and consequences of unit non-response and the suitability of imputations meant to address it. Item non-response error is arguably the most studied of the error components, with recent validation studies showing that the common mitigation approach of imputation leads to high error rates at the household level (Chenevert, Klee, and Wilkins 2016; Hokayem, Bollinger, and Ziliak 2015; Bollinger et al. 2019; Celhay, Meyer, and Mittag 2021; Meyer, Mittag, and Goerge 2022). These validation studies also provide evidence that measurement error – i.e., respondents providing inaccurate answers – is substantial and pervasive (see Bound, Brown, and Mathiowetz 2001 for an extensive review).

To formally define these error components, let μ be the parameter of interest, which we want to estimate from a random variable X. We refer to μ as the survey target. Let x_i denote the realization of X for unit i. In our application, x_i is either an indicator for program receipt or the amount of dollars received by unit i and μ is the receipt rate or average dollars received. We have two measures of x_i: x_i^A from the administrative records and x_i^S from the survey data. The survey uses weights w_i^S to account for sampling probabilities and unit non-response. All summations are taken over the survey observations that can be linked to the administrative records (the "linked data"). We assume below that the linked data can be reweighted to be representative of the population the survey actually covers. In our case, this amounts to assuming that multiplying the final survey weights by an inverse probability weight adjustment, $\widehat{w}_i$, addresses the problem of

[4] Survey coverage is key for estimating population statistics and analyzing subpopulations, as discussions of coverage of the Decennial Census show (e.g., Mulry 2007, O'Hare et al. 2016, Seeskin and Spencer 2018). See Krohn (2024) for a recent application and de Leeuw and de Heer (2002) and Mulry and Spencer (2001) for reviews of the extent of error and methods to address it.
[5] Unit non-response has been a focus in the CPS and more generally due to its rapidly increasing prevalence (de Leeuw and de Heer 2002, National Research Council 2013, Rothbaum and Bee 2021, Bee and Rothbaum 2025).
[6] Meyer and Mittag (2021) include this error component as part of measurement error.

incomplete linkage (see Meyer and Mittag 2021 for discussion of this assumption). Finally, let r_i^U indicate whether unit i responded to the ASEC and r_i^I indicate whether unit i responded to the relevant survey question (i.e., is not an item non-respondent).

TSE is the difference between the survey estimate of the parameter of interest and the survey target, μ. We estimate the TSE as:

$$\hat{\varepsilon}_{TSE} = \frac{1}{\sum \widehat{w}_i w_i^S} \left[\sum \widehat{w}_i w_i^S x_i^S\right] - \hat{\mu}, \tag{1}$$

where the first term is the survey estimate of the population mean using the reweighted linked data. The second term, $\hat{\mu}$, is the estimated survey target. If the administrative records cover the same population as the survey and linkage is error-free, then it is straightforward to calculate the survey target from the administrative microdata. However, the administrative data may contain records the survey does not intend to cover (e.g., non-residents or the institutionalized) and records it intends to cover but cannot, such as those without a linking key. Carefully constructing an accurate measure of the survey target is one of the key challenges of measuring TSE. Section 3 discusses in detail how we estimate and adjust our survey targets.

We then additively decompose our estimate of TSE into four components:

$$\begin{aligned}\hat{\varepsilon}_{TSE} = & \underbrace{\frac{1}{\sum \widehat{w}_i w_i^S} \left[\sum \widehat{w}_i w_i^S x_i^A\right] - \hat{\mu}}_{\hat{\varepsilon}_{GCE}} + \underbrace{\frac{1}{\sum \widehat{w}_i w_i^S} \sum \left(1 - r_i^U\right) \widehat{w}_i w_i^S (x_i^S - x_i^A)}_{\hat{\varepsilon}_{SNR}} \\ & + \underbrace{\frac{1}{\sum \widehat{w}_i w_i^S} \sum r_i^U (1 - r_i^I) \widehat{w}_i w_i^S (x_i^S - x_i^A)}_{\hat{\varepsilon}_{INR}} + \underbrace{\frac{1}{\sum \widehat{w}_i w_i^S} \sum r_i^U r_i^I \widehat{w}_i w_i^S (x_i^S - x_i^A)}_{\hat{\varepsilon}_{ME}}. \\ & = \hat{\varepsilon}_{GCE} + \hat{\varepsilon}_{SNR} + \hat{\varepsilon}_{INR} + \hat{\varepsilon}_{ME}. \end{aligned} \tag{2}$$

The first term is generalized coverage error, $\hat{\varepsilon}_{GCE}$, which is the difference between the error-free estimate of the parameter of interest for the population that the survey actually covers and the survey target. The second term is supplement non-response error, $\hat{\varepsilon}_{SNR}$, which is the difference between the statistic of interest calculated using imputed values (x_i^S) and actual values (x_i^A) for supplement non-respondents. Item non-response error, $\hat{\varepsilon}_{INR}$, is the difference between the statistic of interest calculated using imputed values and actual values for item non-respondents. Finally, measurement error is the difference between the statistic of interest calculated using reported and actual values of x_i for respondents. We report these error measures as a percentage of the survey target to put them on a comparable scale. In addition, we examine how supplement and item non-response rates, which measure error prevalence, relate to our measure of bias. To do so, we also report measures of error per unit, simply defined as the bias divided by the respective prevalence (i.e., the supplement or item non-response rate). This measure of bias per unit is useful for comparisons across error components that abstract from differences in the response rates (i.e., when the sample sizes differ between components).

We further decompose each of the three error components estimated from unit-level observations – supplement non-response, item non-response, and measurement error – into two parts due to misclassification, false positives (non-recipients who receive benefits according to the survey: $\mathcal{FP} = \{ x_i^S > 0 \,\&\, x_i^A = 0\}$) and false negatives (recipients not receiving benefits according to the survey: $\mathcal{FN} = \{x_i^S = 0 \,\&\, x_i^A > 0\}$), and, when analyzing dollars received, a part due to errors in amounts among those correctly classified ($\mathcal{CC} = \{x_i^S > 0 \,\&\, x_i^A > 0\}$). Applying this decomposition to measurement error, we can decompose $\hat{\varepsilon}_{ME}$ into three subcomponents:

$$\hat{\varepsilon}_{ME} = \underbrace{\frac{1}{\sum \widehat{w}_i w_i^S} \sum_{i \in \mathcal{FP}} r_i^U r_i^I \cdot \widehat{w}_i w_i^S x_i^S}_{\hat{\varepsilon}_{ME}^{FP}} - \underbrace{\frac{1}{\sum \widehat{w}_i w_i^S} \sum_{i \in \mathcal{FN}} r_i^U r_i^I \cdot \widehat{w}_i w_i^S x_i^A}_{\hat{\varepsilon}_{ME}^{FN}}$$
$$+ \underbrace{\frac{1}{\sum \widehat{w}_i w_i^S} \sum_{i \in \mathcal{CC}} r_i^U r_i^I \cdot \widehat{w}_i w_i^S \left(x_i^S - x_i^A\right)}_{\hat{\varepsilon}_{ME}^{Amount}}$$

We can decompose item non-response error similarly, with $r_i^U r_i^I$ changed to $1 - r_i^U$ for supplement non-response error and to $r_i^U(1 - r_i^I)$ for item non-response error.

For these three components, we also measure overall data inaccuracy by summing the magnitudes of errors (so that offsetting errors no longer cancel). We first estimate total absolute error in amounts, denoted by $\hat{\varepsilon}_{ME}^{|Amount|}$, by replacing $x_i^S - x_i^A$ in the definition of $\hat{\varepsilon}_{ME}^{Amount}$ by its absolute value $|x_i^S - x_i^A|$. We then estimate total absolute error from the respective component by adding this new term to false positive error and subtracting false negative error (which is negative by definition). That is, for supplement non-response error, we estimate our absolute error components as $|\hat{\varepsilon}|_{ME} = \hat{\varepsilon}_{ME}^{FP} - \hat{\varepsilon}_{ME}^{FN} + \hat{\varepsilon}_{ME}^{|Amount|} = |\hat{\varepsilon}_{ME}^{FP}| + |\hat{\varepsilon}_{ME}^{FN}| + \hat{\varepsilon}_{ME}^{|Amount|}$. We define absolute error components for supplement and item non-response error analogously. Total absolute survey error is then the sum of the absolute values of the four error components:[7]

[7] Our measure of absolute error is thereby a measure of average absolute prediction error – i.e., how well the survey values "predict" the administrative variable at the individual level. Alternative measures include the mean squared

$$\hat{\varepsilon}_{TASE} = |\hat{\varepsilon}_{GCE}| + |\hat{\varepsilon}|_{SNR} + |\hat{\varepsilon}|_{INR} + |\hat{\varepsilon}|_{ME}.$$

Our measure of absolute error can be interpreted as the number of program recipients/dollars that would need to be changed in order to eliminate all error in the survey (expressed as a percentage of true total recipients/dollars).[8] We again define per-unit inaccuracy as absolute error divided by the respective prevalence (e.g., the non-response rate).

3. Data

Implementing this decomposition requires a number of ingredients: (1) a measure of truth with definitions and timing that match the survey concepts being validated, (2) aggregate benchmarks (to construct survey targets) and individually linkable administrative microdata (to estimate unit-level error components), (3) knowledge of the survey design, including which populations are intentionally excluded from coverage, and (4) a way to assess the completeness of the administrative records. Income sources and government transfers are well-suited to this approach because administrative records for these sources tend to be accurate, obtainable, and well-documented.

3.1. Survey Data

We rely on survey data from the Current Population Survey Annual Social and Economic Supplement (CPS ASEC), which surveys approximately 100,000 households in person and by phone between February and April of each calendar year. The CPS ASEC asks detailed questions on income received during the previous year along with extensive demographic information (U.S. Census Bureau 2025). In this paper, we focus on the CPS ASEC from survey years 1997 and 2018, which correspond to reference years 1996 and 2017, respectively.[9] The CPS ASEC is intended to

error and variance of survey estimates, but we prefer our measure because it is intuitive and directly summarizes overall inaccuracies. Note that in contrast to variance, on which biases in multivariate analyses typically depend, our measure is not centered. Thus, it combines noise and bias. In the presence of substantial net error, the bias in the variance may thus be a preferable measure if one wants to predict bias rather than characterize survey accuracy.

[8] For supplement non-response, item non-response, and measurement error, this measure is built up from individual-level errors (false positives, false negatives, and absolute errors in amounts), so it captures offsetting errors that cancel in net terms. In contrast, coverage error enters as the absolute value of its net aggregate, since individual-level decomposition is not possible for units outside the survey frame. As a result, coverage error contributes identically to net and absolute TSE in magnitude, such that the expansion from net to absolute error is driven entirely by the three unit-level components.

[9] We focus on 1996 as it is one of the first years for which administrative microdata are available and on 2017 as it is among the last years before CPS data collection and response patterns were disrupted by the COVID-19 pandemic.

be representative of the U.S. civilian non-institutional population. It is also the official source of income, poverty, and health insurance statistics in the U.S., making it one of the most important surveys in the country. Because it is fielded as a supplement to the Basic Monthly CPS, its observations are a subset of those who answer the Basic Monthly CPS (U.S. Census Bureau 2019).

There are two types of non-respondents to the CPS ASEC.[10] Supplement non-respondents (also called "full supplement" imputations) answer the Basic Monthly CPS but not the ASEC, so their entire ASEC record is imputed from donor records via a hot deck procedure that sorts observations into cells based on observed characteristics from the Basic Monthly CPS (U.S. Census Bureau 2002; Hokayem, Raghunathan, and Rothbaum 2015). Item non-respondents complete the ASEC supplement but fail to answer specific questions (e.g., recipiency or dollar amount for a given income source).[11] Imputation again follows a hot deck approach, but the details of implementation vary across variables. For example, the hot deck for SNAP sorts households into 648 cells that are specific to the SNAP income source (Celhay, Meyer, and Mittag 2021; U.S. Census Bureau 2021). A key aspect of the imputation process is that measurement error among respondents will naturally propagate to item non-respondents whose imputed values are based on reports. These errors will further propagate to supplement non-respondents imputed from donor records that may have some variables item imputed and other variables reported. Conceptually, this suggests that imputations should be at least as noisy as reports.

Table 1 lists the ten income sources we examine. Our unit of analysis corresponds to the level at which each source is reported in the CPS: individuals aged 15+ for seven income sources (pensions, OASDI, OASI, DI, SSI, UI, and VA disability compensation), individuals of all ages for Medicaid, and households for SNAP and housing assistance. The CPS asks separate questions about the receipt of each income source and collects amounts for all of them except Medicaid and housing assistance.[12] All income sources are available in both 2017 and 1996, with the exception of OASI and DI (which are not available in 1996, since the survey question that separated OASDI

[10] Note that the Basic Monthly CPS also suffers from non-response, but that is not a bias that we explicitly examine and it will ultimately be embedded in our coverage error term.

[11] We classify an income source as item imputed only if its receipt is imputed (which in turn implies that amounts are imputed). In some cases, respondents report receipt but have their amounts imputed – in this case we do not classify that income source as imputed. When an income source is constructed as the sum of components (e.g., for pensions), we call the overall income source imputed if any component is imputed.

[12] We define "pensions" as the sum of several components in the survey (pensions, other private retirement distributions, private survivors' income, and private disability income). In addition, we separate OASDI into OASI and DI based on an additional survey question that asks about the reason for receiving Social Security.

into its components only started being asked in 2000). Appendix Section A.1 provides a detailed summary of the survey questions and variables for each income source.

3.2. Administrative Data

We draw on a variety of administrative sources to provide measures of "truth" against which to evaluate survey accuracy. A key challenge is finding relevant administrative data sources that align with survey income concepts. As a result, we focus on income sources for which the administrative microdata are the gold standard, typically government benefits recorded by the administering agencies (which are therefore accurate and nearly complete). In contrast, we do not analyze income sources like earnings, business income, and asset income, where survey and administrative income concepts can diverge more substantially. For several sources we do analyze (namely SSI and housing assistance), the administrative data cover a narrower concept of benefits than the survey. For these sources, any negative (positive) survey errors are likely to be understated (overstated). Table 1 lists the administrative sources we use. They cover the entire U.S. for every source except SNAP (for which we obtain records from five states). Administrative records for all ten sources are available in 2017, and our analyses over time focus on four sources with both survey and administrative data in 1996 (pensions, OASDI, SSI, and housing assistance).

Data on pensions come from IRS Form 1099-R, which covers gross payments from employer-sponsored plans and Individual Retirement Account (IRA) withdrawals and excludes rollovers. Bee and Mitchell (2017) discuss the conceptual alignment between these records and the CPS variables. OASDI benefits come from the SSA's Payment History Update System (PHUS) file, and we use the SSA's Master Beneficiary Record (MBR) to separate payments from Old-Age and Survivors Insurance (OASI) and Disability Insurance (DI). Our preferred measures of OASI and DI are gross amounts that include medical insurance premiums withheld from the net disbursement. SSI benefits come from the SSA's Supplemental Security Record (SSR) file, which includes federal and federally-administered state payments but excludes state-administered payments (generally less than 5% of all SSI payments, which the CPS definition covers). OASDI and SSI present a unique challenge because benefits are sometimes paid on behalf of children under 15, who appear as recipients in the administrative data but cannot report receipt in the survey.

We attempt to reallocate these children's administrative values to the adults in their household, matching the survey's convention that an adult reports benefits received on a child's behalf.[13]

Unemployment Insurance (UI) records come from Box 1 of IRS Form 1099-G, which are filed with the IRS by state agencies and report the total amount of UI benefits and other government benefits received by an individual during the tax year.[14] Records on service-connected disability payments to veterans are available from the Department of Veterans Affairs (VA). Supplemental Nutrition Assistance Program (SNAP) records come from state agencies. Medicaid enrollment records come from the Medicaid Statistical Information System (MSIS), which is a national database managed by the Centers for Medicare and Medicaid Services (CMS). Finally, housing assistance records come from the Public and Indian Housing Information Center (PIC) and Tenant Rental Assistance Certification System (TRACS) files, which cover almost all public and subsidized housing assistance programs under the jurisdiction of the Department of Housing and Urban Development (HUD). These HUD data cover only federally administered programs and thus miss programs administered by the U.S. Department of Agriculture (USDA) and by states and localities, which may be included by respondents in the CPS.

3.3. Linking Data Sources

We link the administrative microdata to the CPS at the individual level using person identifiers created by the Person Identification Validation System (PVS) of the U.S. Census Bureau (Wagner and Layne 2014). 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 that contains all transactions recorded against a Social Security Number (SSN). If a matching record is found, the SSN from the reference file is transformed into a Protected Identification Key (PIK) and attached to the corresponding records in our data. For the administrative data, PIK rates

[13] The reallocation depends on what the household's adults report, and we handle three cases: (1) If no adults (age 15+, PIKed) in the household report receipt in the survey, then we do not identify any false negatives but include dollars and people in coverage error, which understates false negatives and overstates coverage error. (2) If one or more adults (age 15+, PIKed) report receipt in the survey but are not linked to the admin data, then we assign each child's admin dollars to one of the adults, allowing each child to cancel out one false positive; if there are more children than adults, the extra children's dollars are counted in coverage error. (3) If there are only correct reporting recipients among adults (age 15+, PIKed), then we add the linked children's dollars to the adults, equally splitting total children's dollars among the correct reporting adults; missing recipients are counted in coverage error.

[14] In some states, Box 1 can also include payments from Paid Family Leave or Short-Term Disability Assistance programs, but these are only present for four states in 2017 and thus likely to be trivial (see Meyer et al. 2023 for more details).

tend to be nearly complete at around 99 percent for most income sources (see Table 2). In the survey sample, the shares of individuals who have a PIK are 81% and 85% in 1996 and 2017, respectively, while the shares of households where at least one member is PIKed are 86% and 91% in 1996 and 2017.

We cannot assess survey error for survey observations without a PIK, so all analyses are based on the survey sample with a PIK (the "linked data"). Despite the low rate of missed links, PIKs are not missing completely at random in the survey data. To restore representativeness of the linked data for the survey population, we use inverse probability weighting (Wooldridge 2007). Specifically, we estimate probit models to predict the probability that an observation has a PIK and multiply the survey weights by the inverse of this predicted probability (see Section A.2 of the Data Appendix for more details).[15] We adjust for non-PIKing at the level corresponding to our unit of analysis – i.e., at the individual level for all income sources except SNAP and housing assistance, for which we adjust at the household level. To validate our reweighting methodology, Appendix Table A.1 compares weighted totals of recipients and dollars for the income sources we examine between the full sample with original weights and the reweighted linked subsample. The estimates are very similar (always within 1% of each other), suggesting that conditional on the covariates in the probit model, PIK status does not predict receipt or reporting.

3.4. Constructing Survey Targets

To estimate TSE, we need an aggregate measure of what the survey intends to yield – i.e., the survey target. A natural starting point is to calculate the total recipients and dollars for each income source in the administrative microdata, which should approximate totals available from published program statistics. However, these aggregates may include recipients not covered by the survey. To account for these coverage differences, we implement and extend several adjustments to the administrative aggregates from previous studies, including Roemer (2000), Rothbaum

[15] We model the probability of obtaining a PIK as a function of a rich set of observed survey characteristics. This includes demographic characteristics such as age, sex, race and ethnicity, nativity, citizenship, marital status, education, and disability; household composition and structure; state and rural residence; household income and earnings relative to the SPM poverty threshold; labor-force status, hours worked, and health-insurance coverage; and indicators for the receipt and, where feasible, amounts of major income sources, including earnings, self-employment income, Social Security, SSI, UI, TANF, SNAP, housing assistance, VA benefits, pensions, interest, dividends, and rental income. We also include measures of self-employment losses and reliance on business income, along with selected interactions. The household-level model uses the household head's characteristics and household-level measures, while the individual-level model additionally incorporates person-specific income-source, employment, and disability measures.

(2015), Meyer, Mok, and Sullivan (2015), Meyer and Mittag (2021), and Corinth, Meyer, and Wu (2022). Section A.3 of the Data Appendix provides a detailed overview of these adjustments.

We first implement a series of adjustments for intentional coverage differences. First, we subtract records corresponding to decedents – specifically, those in the administrative records for reference year *t* who died before they were eligible to be surveyed – by linking to individual-level death dates in the SSA Numident file to identify recipients who died during year *t* or in January-February of year $t + 1$.[16] Second, we subtract records corresponding to those living in group quarters (GQs) – both institutional GQs (not covered in the CPS frame) and non-institutional GQs (under-surveyed in the CPS).[17] To do this, we link administrative records for year *t* to GQ records in the American Community Survey (ACS) from survey year $t + 1$ and calculate the weighted total of linked dollars in GQs.[18] Third, in the case of SNAP, our only administrative income source with incomplete geographies, we subtract records linked to residents who appear in a different state when surveyed in year $t + 1$. Fourth, to account for records not covered by the linked data, we subtract administrative records that do not link to a PIK.[19] We apply this methodology to every source with universe administrative data – i.e., all sources except those from the SSA (OASDI, OASI, DI, and SSI), for which we have administrative extracts only for records that link to the CPS ASEC. For these SSA sources, we instead apply public-use analogs of the decedent and GQ adjustments to publicly available administrative totals and do not apply an adjustment for non-PIKing in the administrative data.[20]

We must also align the administrative data with the survey's units of analysis (individual or household) and reference periods (calendar year). Most income sources in the administrative data are already at the relevant unit. The exception is SNAP, whose administrative records are at

[16] This is straightforward for individual-level administrative sources. For administrative sources at the household or assistance unit level (housing assistance and SNAP), we only subtract cases if all members are decedents.
[17] While in principle the CPS includes non-institutional GQs, in practice the counts are very small. Moreover, starting in 2017 (one of the years we analyze), the CPS stopped including college dormitories (one of the largest non-institutional GQ types) in its sample.
[18] Because we use all months of the ACS in survey year $t + 1$, there might be slight misclassification relative to the CPS ASEC if individuals move in and out of GQs between February–April and other months of the year. Because the ACS GQ files are only available starting in 2006, for 1996 we apply the GQ share from 2006 and assume that the share is the same in 1996.
[19] Subtracting administrative records that do not link to a PIK means we avoid overstating survey error and, if anything, understate it. It also makes the survey target depend on the rate of non-linkage in the administrative data, which is conceptually less clean, though in practice these adjustments are small.
[20] Because the publicly available administrative totals for OASDI, OASI, DI, and SSI may include payments to those living overseas, we subtract these records given that the CPS frame covers only the residential U.S. population.

the assistance unit (case) level, which differs slightly from a survey household. We scale the SNAP target down using the average number of cases per survey household in our linked data (1.07 across our five states in 2017). For the reference period, the tax-based sources (pensions and UI) are already at the calendar year level, and many other income sources are at the monthly level (which straightforwardly convert to a calendar-year basis). For the SSA income sources (OASDI, OASI, DI, and SSI), the publicly available recipient targets are for December, which we convert to annual counts by adding total terminations at any point during the year (also published by SSA).

Table 2 shows the adjustments for recipients and dollars in 2017, and Appendix Table A.2 reports them for 1996. The largest adjustment for most sources is for GQs, which account for at least 2% of observations for every source except pensions, UI, and VA disability. The GQ adjustment is particularly large for SNAP (7.2% of recipients and 6.8% of dollars), SSI (3.0% of recipients and 4.5% for dollars), and housing assistance (3.6% of recipients). The adjustment for decedents never exceeds 3%, though as expected it is larger for income sources targeted to the elderly or disabled (for recipients: 2.8% for pensions, 2.2% for OASI, and 2.0% for OASDI). The adjustment for un-PIKed administrative records is less than half a percent for nearly all sources, because one generally needs an SSN to receive a government benefit, and there is a one-to-one relationship between SSN and PIK.[21] Finally, for SNAP we estimate that 1.1% of recipients and 1.2% of dollars are surveyed in a different state from where they received benefits per the administrative data.

4. Measures of Survey Error

To establish the need for our approach, we first review commonly used indicators of survey accuracy: unit non-response rates, item non-response rates, and measurement error rates. While these indicators raise clear concerns about data quality, they are indirect measures that do not directly quantify bias or inaccuracy in survey estimates. We then estimate TSE for recipients and

[21] There are two prominent exceptions to the high administrative data PIK rate. For Medicaid, 2.3 percent of the administrative records do not have a PIK. While Medicaid agencies generally require an SSN as a condition of eligibility, states sometimes enroll individuals without an SSN and give them another unique identifier (this occurs in California). For retirement distributions, 1.3 percent of recipients and 0.8 percent of dollars are un-PIKed, but this is likely an overstatement due to a feature of the 1099-R data. The raw 1099-R data are at the level of an information return, of which the average recipient has 2.5. We must collapse to the individual level, but our only identifier is PIK, so we cannot perform this collapse for un-PIKed observations. Thus, the number of un-PIKed records corresponds to forms rather than individuals.

dollars across our income sources, which provide a direct measure of bias, before explaining why these TSE estimates (while valuable) still leave important questions unanswered. Addressing these questions requires decomposing TSE into its components.

4.1. Indirect Indicators of Survey Error

A first set of indicators emphasized in the survey methodology literature are unit non-response rates – i.e., the fraction of sampled units that provide no survey information at all. Appendix Figure A.2 documents that the CPS supplement non-response rate hovered around 10% between 1990 and 2009, then rose steadily to roughly 20% in recent years. But supplement non-response is only measured among those responding to the Basic Monthly CPS, and Basic non-response itself was typically 5-10% prior to 2010 before increasing sharply to above 30% in 2022. Taken together, by 2022, roughly 45% of the original Basic Monthly CPS sample did not provide responses to *any* CPS ASEC question.

A second set of indicators are item non-response rates, which measure the fraction of respondents who complete the survey but do not answer a particular question. Appendix Figure A.3 shows that item non-response has risen over time for nearly all income sources in the CPS ASEC, quintupling for pensions and OASDI. By 2022, item non-response rates reached nearly 8% for Medicaid, 6% for pensions, and 4% for SNAP (with most other sources around 2% or lower). Unlike the sharp post-2010 rise in unit non-response, these increases in item non-response have been more gradual and sustained over the full 1990–2022 period.

Beyond non-response, a growing literature documents substantial measurement error among respondents, often summarized by false negative rates (the share of true recipients who do not report receipt) and false positive rates (the share of true non-recipients who report receipt). Appendix Table A.3 compiles estimates from prior studies alongside estimates we calculate for respondents in this paper. Prior studies consistently find high false negative rates, but they are limited to a small number of states, years, and programs. Our linked data allow us to report error rates for ten major income sources, covering the entire U.S. for all sources except SNAP. In line with prior work, we find substantial false negative rates in 2017, ranging from 14% for OASI to close to one-half or more for some sources (DI, UI, and pensions). False positive rates are much lower, in part because they are computed over the much larger set of true non-recipients.

4.2. Total Survey Error: A Direct Measure of Bias

The conventional metrics above raise clear concerns, but they measure the prevalence of error rather than its contribution to bias or inaccuracy. This fundamental gap motivates our TSE approach. To directly measure bias in survey estimates, we calculate TSE for average dollars received and the recipient share of the population across our income sources. TSE is the difference between the weighted survey estimate and the survey target constructed from published totals, using public-data adjustments for intentional coverage differences (Appendix Section A.3). Figures 1 and 2 summarize TSE in the CPS from 1984 to 2022, expressed as a share of the survey target for our measures of recipients (seven income sources) and dollars (eight income sources).[22]

In levels, TSE is almost always negative for both recipients and dollars (with SSI dollars an exception in recent years), indicating that survey means are systematically biased downward. For nearly half of the variables the bias is 40% or more. Yet, there is substantial heterogeneity across income sources. TSE tends to be more modest for SSA programs (specifically OASDI and OASI, for which the net bias is below 15% for recipients and dollars), while it is much larger for income sources such as pensions, UI, and SNAP (for which the net understatement exceeds 40%). These magnitudes imply that survey-based analyses of the income available to retirees, the take-up of SNAP, and the extent to which UI cushions income losses during unemployment – to name just a few – may suffer from large biases. Outside of the SSA programs, TSE for recipients and dollars are often quite similar in magnitude. Because recipient error operates purely on the extensive margin (whether receipt is captured at all), this similarity foreshadows our finding that the extensive-margin understatement of receipt – and not the intensive-margin misreporting of amounts – is the major driver of TSE for dollars.

Trends over time also generally point toward worsening TSE, again with heterogeneity across programs. Early in our period, the understatement is relatively similar across sources: for recipients, TSE is around 20% or less for all sources, and for dollars it is below 30% for seven of the eight sources. Among recipients, TSE has worsened for nearly all programs – including SSI, *despite* SSI dollars showing reduced understatement in recent years. This pattern is consistent with increasing confusion between SSI and the larger OASDI program, which we confirm directly in

[22] The estimates for recipients here cover seven income sources rather than the ten in the decomposition (Sections 5 and 6), because they rely on published aggregate totals, and comparable published recipient counts are unavailable for UI and pensions and incomplete for housing assistance. The decomposition instead uses linked microdata and covers all ten.

our linked data (Section 6). Over time, survey error has also tended to diverge across programs. The increase in understatement is modest for OASDI and OASI, more variable for UI and VA disability, and steadier and more pronounced for DI and SNAP (nearly tripling for SNAP between 1984 and 2022). The divergence across programs implies that comparisons of receipt or benefit levels over time or across programs in the CPS may be distorted in ways that have worsened.

Overall, these TSE estimates establish that bias is substantial and generally increasing. While TSE provides a direct measure of bias that is comparable across time and programs, it remains to be seen how the indirect measures discussed in Section 4.1 relate to TSE. Simple comparisons reveal several puzzles. For SSI, VA disability, and Medicaid, much of worsening TSE for recipient shares occurs before the mid- to late-2000s – well before the sharp rise in unit and supplement non-response rates after 2010. Conversely, item non-response rose dramatically over time for OASDI and pensions, yet these sources show comparatively small increases in TSE. These discrepancies arise from the fact that the indirect measures track the prevalence of different error sources but not their per-unit consequences. Both of these aspects can matter for the overall distortion – for example, TSE can increase either because non-response becomes more prevalent or because the bias per non-respondent increases. TSE is a direct measure of bias that already combines prevalence and per-unit error, which lets us separate rates from consequences within a component over time. Since TSE is defined consistently across components, it also facilitates clean comparisons of component-by-component contributions. In the remainder of this paper, we decompose TSE into comparable components to rigorously assess the roles of different error sources in explaining levels and trends in survey error.

5. Decomposing Total Survey Error at a Point in Time

We begin by decomposing TSE for each of our ten income sources in a single recent year (2017) into coverage error, supplement non-response error, item non-response error, and measurement error. Within each component, we further separate extensive-margin contributions (whether receipt is captured at all) from intensive-margin contributions (amounts reported conditional on correctly reporting receipt). We then complement these net decompositions with absolute error measures that capture the extent to which error makes survey values less accurate.

5.1. Decomposition of Bias: Net Survey Error

Figure 3 summarizes the main net decomposition results for 2017. For each income source, it displays the contribution of each error component to total net survey error, expressed as a percentage of the survey target.[23] Each income source has two bars (one for recipients and one for dollars) except Medicaid and housing assistance, for which we focus on recipients only. At the bottom of each bar, we report the TSE in parentheses. We start by walking through the results for an income source, pensions (top left of Figure 3), that exhibits a typical pattern of error. The negative net TSE indicates substantial understatement of both recipients and dollars. The dominant contributor is measurement error, accounting for 79-87% of the total net understatement. Non-response error components are also negative, but smaller in magnitude. Coverage error is slightly positive, indicating mild over-coverage of the retirement-income population.

Turning to patterns across income sources, we find that overall net bias is large and at least 40% of the survey target in seven of eighteen variables. We first consider the eight sources (pensions, OASDI, OASI, DI, UI, VA disability, SNAP, and Medicaid) that exhibit a "typical" pattern and then discuss SSI and housing assistance as special cases. For the eight typical income sources, measurement error is the dominant component and is always negative. For recipient shares, it constitutes the majority of TSE for six of the eight sources and exceeds two-thirds for four sources. For mean dollars, it accounts for the majority of TSE for every source and exceeds two-thirds of TSE for five of the seven non-Medicaid sources. Non-response, both supplement and item, consistently generates negative bias (except item non-response for Medicaid) but of a smaller magnitude than measurement error. Supplement non-response typically contributes more than item non-response, which is consistent with its higher prevalence. Supplement non-response ranges between 11-21% of TSE for recipients and 16-33% of TSE for dollars, while item non-response never exceeds 6% of TSE for recipients and ranges between 3-11% of TSE for dollars. That non-response plays a smaller role than measurement error is perhaps unsurprising since respondents ultimately comprise 73-80% of survey observations across income sources, while supplement non-respondents are about 19% and item non-respondents are another 1-8% (Appendix Table A.4).

[23] Some components can be positive while others are negative, so a component's share of net TSE can exceed 100%.

However, these aggregate contributions reflect both the prevalence and per-unit severity of error within each group. On a per-unit basis, net error for supplement non-respondents is less than that for respondents for three income sources, and never more than 50% higher than that for respondents for any source except SNAP (Appendix Table A.5). Item non-respondents have somewhat higher net error per unit, ranging from 1.1 to 2.3 times that of respondents. As a result, supplement non-response contributes to aggregate bias primarily through its prevalence (roughly 19% of the sample), whereas item non-response (far less prevalent at 2-4% of the sample) generates more net bias per unit. Per-unit error is not only higher among item non-respondents, but it also varies more across income sources than for supplement non-respondents. This makes item non-response *rates* a less useful indicator of the bias that item non-respondents produce.

Finally, coverage error is modest. Pension recipients appear slightly over-covered, and coverage error for OASDI and OASI dollars is also mildly positive. In contrast, most other sources show slight under-coverage. For DI and SNAP, under-coverage is more pronounced for recipients than for dollars, consistent with low-dollar recipients being underrepresented in the survey; for SNAP, this may potentially reflect households with more churn and instability (Meyer and Wu 2021). Prior work has documented meaningful coverage gaps for specific populations, such as Black men (Sabety and Spitzer 2026) and young children (O'Hare et al. 2016), and some of the under-coverage we find for programs like DI and SNAP may reflect similar populations missing from the survey frame. That said, coverage error tends to be a less important contributor to overall bias than the other error components.

These decomposition results yield several key takeaways. First, net understatement in the survey is driven primarily by misreporting among respondents, so strategies that improve respondent reporting (or that impute underreported receipt among respondents) are likely to yield the largest reductions in bias. Second, because imputations are based on reports, this bias spreads to imputed values. Therefore, improving reporting would also reduce imputation bias. Supplement non-respondents generate approximately as much bias per unit as respondents, suggesting that the imputation procedure for supplement non-response matches reported means.[24] For item non-

[24] Note that an imputation procedure based on reports can only lead to less bias per unit than reporting if imputed values are positively selected in terms of reporting quality. But such a relation of imputed values to reporting accuracy would make (characteristics that predict) non-response predict receipt artificially and thus seems undesirable.

respondents, imputations generate more bias per unit, suggesting that item imputations fail to entirely capture the systematic relationship between non-response and receipt.

SSI and housing assistance behave differently in net terms from the other eight sources. While SSI exhibits a large negative TSE for recipients (−25%), it is the only source with positive TSE for dollars (a small overstatement of 2%). As we confirm in the next subsection, this pattern is driven in large part by confusion between SSI and OASDI in survey reporting, given that both programs are administered by the SSA and consistent with prior evidence (Huynh, Rupp, and Sears 2000; Gathright and Crabb 2014). When some true OASDI recipients misclassify themselves as SSI recipients in the survey, it can generate the patterns we see because OASDI benefits are larger on average than SSI benefits. SSI also displays a large coverage-error component (−17%), but this should be interpreted cautiously as it partly reflects the survey's limitations in capturing child SSI receipt.[25] For housing assistance, the results show that our method becomes less informative but can still be useful when the administrative data are incomplete. The TSE of +12% is partly due to the fact that our administrative records miss non-HUD units. As a result, our estimate of TSE is the sum of "true" TSE and recipients of non-HUD programs who report housing assistance receipt – a limitation that does not apply to the other programs we study. Nevertheless, as we show in the next subsections, the finer decomposition into false negatives and false positives remains informative even with incomplete administrative data.

5.2. Extensive vs. Intensive Margin Errors

We now examine whether the (almost always negative) bias arises primarily from (1) errors in receipt (the extensive margin) or (2) errors in amounts among true reporting recipients (the intensive margin). We focus on average dollars received because recipiency rate error operates purely on the extensive margin. Figure 4 presents this breakdown among the three respondent types (supplement non-respondents, item non-respondents, and respondents) for dollars, with Tables 3 and 4 providing the underlying numbers for recipients and dollars, respectively.[26]

[25] Because children can receive SSI in administrative data but only those aged 15+ can report it in the survey, linked child recipients who cannot be assigned to a survey adult are placed in the coverage error residual (see Section 3). Survey-linked SSI aggregates from prior work suggest the actual coverage error is roughly half what we report here.
[26] Appendix Tables A.6 and A.8 show the net error analogs of Tables 3 and 4, where the estimates are shown as a percentage of total survey error (rather than the survey target).

Among respondents who make up most of the sample and thus most of TSE, extensive-margin underreporting (false negatives) explains the bulk of net bias. For seven of the eight income sources (excluding binary Medicaid and housing assistance receipt), false negatives account for more than 80% of net measurement error for dollars. For five sources (OASDI, OASI, DI, UI, VA disability), false negatives are so pronounced that they exceed 100% of net measurement error, with other errors (false positives and errors in amounts) partially offsetting them in net terms. Intensive-margin errors in amounts are not small for pensions and SNAP but still represent less than a quarter of total net error. SSI is the clear exception to this pattern, as the bias from respondent false positives (+36%) exceeds that from respondent false negatives (−23%), consistent with confusion between SSI and OASDI. In fact, 53% of SSI false positives are also OASDI false negatives, suggesting that OASDI (mostly DI) recipients often report their benefits as SSI.

For supplement and item non-respondents, net bias is again almost entirely an extensive-margin phenomenon, because the vast majority of true recipients are imputed to receive nothing (Table 4). Average errors in amounts conditional on correctly imputed receipt are small. The negative contribution from false negatives exceeds the positive contribution from false positives for essentially all sources (SSI among supplement non-respondents is the exception), explaining why non-respondent bias is consistently negative on net. For some sources, the relative differences are small; e.g., for supplement non-respondents, the bias from OASI false negatives (−4.3%) only slightly outweighs that from OASI false positives (+3.3%). But for many sources (e.g., pensions, UI, and SNAP), the bias from false negatives is more than twice that from false positives.

Housing assistance is a special case because our administrative data cover only HUD programs, which capture a subset of true housing assistance recipients. Even so, our approach remains useful in the presence of incomplete administrative data: our estimate of the bias from false negatives is a valid estimate of how much HUD assistance the CPS misses. At 22%, it indicates that a sizeable fraction of recipients fail to report housing assistance receipt. On the other hand, those who report receipt without being linked to a HUD record do not clearly indicate false positive bias, since they may be recipients of non-HUD housing assistance. Nevertheless, our apparent false positive estimate is informative about the size of non-HUD assistance, which is otherwise difficult to assess (Olsen 2018). Assuming that underreporting is similar for HUD and non-HUD housing assistance, and that true false positives are rare, our results imply a non-HUD caseload in the range of prior estimates (i.e., 2.2-2.5 million households, see Olsen 2018).[27]

Overall, our key takeaway is that extensive-margin underreporting (false negatives) is the dominant source of bias in dollar amounts, and this pattern is remarkably consistent across income sources and respondent types. Departures from this pattern (e.g., for SSI) are more consistent with misclassification across programs than with reporting excess take-up. Although false positives contribute much less to net bias than false negatives, they are not negligible and partly offset downward biases, making absolute error larger than net error.

5.3. Decomposition of Inaccuracy: Absolute Survey Error

We next decompose total absolute survey error, complementing the net decompositions by measuring overall inaccuracy at the unit level. Figure 5 provides the absolute-error analog of Figure 3 for 2017 (in which all components are positive by design), and Tables 3 and 5 report the corresponding magnitudes for the intensive- and extensive-margin subcomponents.[28] Absolute error ranges from 26-96% of the survey target for the recipient share and 45-124% for dollars received. It exceeds 80% of the survey target for eight of the eighteen variables we consider, indicating that making the survey error-free would require changes equal to 80% of the true total recipients or dollars in these cases. For five income sources, eliminating all dollar error would require changing nearly as many dollars as (if not more than) the total received. Unlike net error, total absolute error is uniformly larger for dollars received than for recipients, reflecting the additional intensive-margin (amounts) error channel that applies only to dollars. SSI demonstrates most strikingly how net error can be a misleading guide to data quality. In net terms, SSI dollars appear nearly unbiased (+1.9%) but in absolute terms SSI has the largest dollar TSE of 124% – revealing extensive offsetting errors that net measures obscure. This is consistent with substantial misclassification and misallocation of SSI dollars across observations, which is precisely the kind of noise that makes a survey inaccurate without affecting net bias.

[27] See Section A.4 of the Data Appendix for details. The 2.2-2.5 million range corresponds to two combinations of inputs: our overall false negative rate paired with a false positive rate around 0.5%, or our respondent false negative rate paired with no false positives.
[28] Appendix Tables A.7 and A.9 show the absolute error analogs of Tables 3 and 5, where the estimates are shown as a percentage of total survey error (rather than the survey target).

Measurement error remains the largest contributor to absolute error in the recipient share for seven of ten sources, and in dollars for all income sources. Our absolute error decomposition reveals that intensive-margin errors in amounts sometimes become more salient when analyzed in absolute terms. For example, among OASDI respondents, net error in amounts is essentially zero (−0.1%), but the absolute error in amounts is 16.0%. This result implies a substantial degree of (approximately mean-zero) error in reported OASDI amounts among reporting recipients.

Perhaps most markedly, non-respondents contribute disproportionately to survey inaccuracy despite being a minority of observations. Figure 5 shows that supplement non-respondents account for 23-40% of absolute recipient error and 18-37% of absolute dollar error, well above their roughly 19% share of the sample and indicating substantially higher severity per unit. Their contributions are especially large for VA disability (40% of absolute recipient error and 37% of absolute dollar error), SNAP (29% and 33%, respectively), and Medicaid (31% of absolute recipient error). Item non-respondents also play a much larger role in absolute error than in net error, contributing 4-17% of absolute recipient error and 3-9% of absolute dollar error. As a result, measurement error among respondents, while still large, typically represents a smaller share of total absolute error than of total net error.

Appendix Table A.5 quantifies these per-unit differences by comparing per-unit absolute error across respondent types. On a per-unit basis, supplement non-respondent absolute error is 1.4 to 3.4 times that of respondents for recipients and 1.3 to 2.5 times that of respondents for dollars. Absolute error per unit for item non-respondents is higher still, with multiples of 2.2 to 9.6 for recipients and 2.2 to 3.6 for dollars. This stands in contrast to the net error comparison in Section 5.1, where supplement non-respondent per-unit net error was often comparable to that of respondents. Thus, while supplement non-respondents contribute about as much bias per unit as respondents (Section 5.1), they contribute substantially more inaccuracy per unit. Item non-respondents are worse on both dimensions, and because their per-unit error varies more across income sources, item non-response rates are a particularly poor predictor of the inaccuracy they produce.

The larger role of non-respondents in absolute error stems from extensive-margin misclassification being more severe among them. This is visible in the bias decomposition (Table 4) and in a direct comparison of false negative and false positive rates by respondent type (Table 6). False negative rates are lowest among respondents, rise sharply for item non-respondents, and are highest for supplement non-respondents – exceeding 68% for eight sources and 81% for five (DI, SSI, UI, VA disability, and housing assistance). False positive rates are also lowest among respondents and higher among non-respondents, with item non-respondents particularly prone to false positives in some cases (notably Medicaid). Non-respondents are thus more likely both to be recorded as not receiving an income source they do receive and as receiving one they do not. Because opposite-signed errors partially offset in net terms but cumulate in absolute terms, the relative contribution of non-response increases markedly when we shift from net to absolute error.

These absolute-error results underscore that low net bias can mask substantial inaccuracy from offsetting components, making net bias an incomplete guide for empirical work (as SSI starkly suggests). More broadly, imputations for non-respondents can add considerable noise even when their contribution to net bias is smaller. While measurement error among respondents is the dominant source of net bias, non-respondent imputations are an important source of inaccuracy. Indeed, relative to their sample share, they are the dominant source of inaccuracy.

6. Decomposing Total Survey Error Over Time

We next apply our framework to analyze whether and how survey accuracy has changed since 1996 for the four income sources (pensions, OASDI, SSI, and housing assistance) with linked data in both 1996 and 2017. Figures 1 and 2 showed that overall TSE was relatively similar across income sources at the beginning of our period but has since diverged considerably. The direction of change over time is fairly heterogeneous across income sources: TSE improves slightly for pensions, worsens for OASDI, worsens for SSI recipients (while moving from negative to slightly positive for SSI dollars), and becomes slightly less positive for housing assistance recipients. Indeed, the four sources we analyze here are among the least problematic in terms of changes in net TSE. Yet moving from net to absolute error (Figure 7) shows clearly that the CPS has become less accurate over time. Across all four income sources, absolute error increases sharply from 1996 to 2017. SSI again illustrates the point starkly, with total net error for SSI dollars moving from a 13% understatement in 1996 to a 2% overstatement in 2017 even as total absolute error rises from 79% to 124%.[29] Notably, survey error can increase in absolute terms even as it shrinks or reverses in net terms.

[29] A key driver is rising confusion between SSI and OASDI, with the share of SSI false positives that are also OASDI false negatives nearly doubling from 28% in 1996 to 53% in 2017.

Section 6.1 uses our decomposition and measure of absolute error to identify which sources drive the unambiguous deterioration of survey accuracy and why it is not reflected in the ambiguous net error trends. The net trends differ across sources because errors offset to differing degrees, and only the finer decomposition reveals that apparent improvements often reflect rising offsetting errors rather than greater accuracy. Section 6.2 then argues that regularly implementing our decomposition would improve upon commonly reported measures of survey accuracy.

6.1. Decomposing Changes in Net and Absolute Survey Error Over Time

Figure 6 decomposes net TSE for the four sources in 1996 and 2017, separately for recipients and dollars, to show why trends in net error are heterogeneous even as response rates and absolute error point to an unambiguous decline in survey accuracy. Much of the heterogeneity in net TSE reflects the fact that measurement error among respondents, still the dominant component in most cases, falls for some sources and rises for others. By contrast, two smaller components move in a common direction, with net errors from supplement and item non-respondents growing considerably over time (often more than doubling) though they remain smaller than measurement error in levels. Coverage error varies idiosyncratically across income sources but is generally smaller. Taken together, these patterns suggest that net trends can mask a clear deterioration in non-respondent error. This result can explain our finding from Section 4.2 that dramatic increases in non-response rates are often not visible in TSE, as they can be drowned out by changes in measurement error. So while measurement error is the largest source of error, non-response error is ultimately what drives the systematic decline in survey accuracy for these four income sources.

The same decomposition can explain why changes in net TSE conceal the uniform deterioration in accuracy shown by absolute error (Figure 7). For SSI dollars and housing assistance, net TSE falls because some error components flip sign and offset increased error elsewhere. In fact, error from almost all components grows in absolute terms (with the exception being a small decrease in item non-response error for housing assistance). Thus, the net reduction in TSE does not stem from an improvement in any error source, but a more benign alignment of offsetting errors. Even for pensions and OASDI, the change in net TSE is partly due to an increase in offsetting biases. These analyses explain why net reporting appears to improve in places (for pensions and SSI recipients), even as absolute error uniformly rises.

We next turn to changes in extensive- and intensive-margin error, again focusing on dollar amounts (Tables 8 and 9 for net and absolute error, respectively) for which this distinction is meaningful, with Table 7 reporting the corresponding extensive-margin patterns for recipients.[30] Overall, Tables 7-9 paint a bleak picture, showing in most cases that all extensive- and intensive-margin errors increase. For non-response error, all measures worsen for both recipients and dollars across all income sources.[31] Among supplement non-respondents, false positive bias more than doubles for every source (and roughly triples for SSI), and false negative bias more than doubles across the board. Item non-respondents show the same directional pattern, with rising false-positive and false-negative bias for all sources, though smaller in magnitude. On the intensive margin, net errors in amounts remain relatively small and idiosyncratic for non-respondents, but absolute errors in amounts rise for both supplement and item non-respondents across all income sources, reinforcing the view that imputations have become noisier over time. Changes in extensive-margin misclassification thus drive the deterioration in non-respondent accuracy, which on net shows up as rising negative bias because false negatives dominate false positives for most sources (SSI is an exception because of program confusion). In short, while measurement error is the main source of bias, non-response adds substantial and increasing noise. Coverage error also generally becomes worse, but improves for OASDI both in terms of recipients and dollars.

The only error component that does not systematically deteriorate is measurement error. While net measurement error increases for OASDI and SSI dollars, it stays constant for housing assistance and declines for pensions and the SSI recipient share. For pensions, both false positive and false negative errors fall, improving net and absolute error for recipients and dollars alike. Similarly, both false positive and false negative errors decrease for housing assistance. But this improvement is not reflected in net error, since the two survey improvements offset each other so that there is no change in net bias. Only our measure of absolute error shows that the reporting of housing assistance actually improved. On the contrary, for OASDI, our decomposition reveals that the marked increase in false negative bias masks a small reduction in false positive bias, so that net and absolute errors grow despite a reduction in overreporting.[32]

[30] Appendix Tables A.12-A.15 show the net and absolute error analogs of Tables 7-9, where the estimates are shown as a percentage of total survey error (rather than the survey target).
[31] The increase in net error in amounts is generally smaller and it minimally decreases in two cases.
[32] Even though absolute measurement error declines for pensions and housing assistance, it still increases overall due to a substantial increase in both supplement and item non-response error. This finding again underlines the pronounced increase in non-response error.

Notably, these error reductions are partly compositional. As non-response rises, respondents become a smaller fraction of the sample and so cause less bias simply because there are fewer of them. The false negative and false positive rates in Table 10 show that reporting becomes worse for pensions, OASDI, and SSI.[33] For pensions and SSI recipients, the error reduction due to increased non-response dominates so that reporting error declines. For OASDI, reporting accuracy decreases so severely that the smaller sample of respondents in 2017 causes more bias than the larger sample in 1996. The only notable exception is housing assistance, where reporting genuinely improves as shown by a reduction in both error rates.

Finally, changes in the accuracy of amounts do not appear to be a major source of deteriorating survey accuracy. Table 8 shows that net error in reported dollar amounts decreases for all three income sources with amounts. However, these improvements should be treated with a grain of salt – for pensions and OASDI, Table 9 shows that focusing on net error hides a decrease in accuracy. For SSI, both net and absolute error in amounts decline, but it is unclear whether these changes reflect changes in reporting or changes in who reports receipt, since for SSI we see a dramatic deterioration due to extensive margin measurement error.

Taken together, our decomposition and measure of absolute error show that trends in net error alone can be misleading, because offsetting changes obscure substantial declines in survey accuracy. Even though we see some modest improvements in net bias and a few cases of genuine improvement in survey reporting, survey accuracy has unambiguously decreased for every income source and measure in terms of absolute error.

6.2. The Value of Routinely Decomposing Bias and Inaccuracy

Our results above reveal a clear deterioration in surveys over time, especially in extensive-margin errors among non-respondents and in inaccuracy overall. For survey users conducting time-series analyses or comparing statistics over time, understanding how survey accuracy changes is essential for judging robustness and, where feasible, bias-correcting trends. For survey producers, tracking survey accuracy over time shows where it is slipping (rising non-response and worsening errors among non-respondents) and can help target improvements cost-effectively. Trends in survey error are typically discussed in terms of the non-response rates (and, less frequently, error

[33] OASDI and SSI unambiguously become worse, while for pensions there is a small reduction in the false negative rate that is more than offset by an increase in false positives.

rates) presented in Section 4.1. Non-response rates have indeed risen, with supplement non-respondents more than doubling as a share of the sample (from 8% in 1996 to 19% in 2017) and item non-respondents also increasing over time though remaining far less prevalent (Appendix Table A.4). All else equal, this rising prevalence mechanically increases the contribution of non-respondents to aggregate error. As we saw, that contribution is difficult to detect without a decomposition like ours to isolate non-response error from the dominating trends in measurement error. But there also remains the possibility of a second layer of deterioration – namely, that per-unit errors may also be rising within respondent type, which standard measures like non-response rates do not capture. If per-unit error changes, then non-response rates would poorly capture survey accuracy even when isolating a specific error component. Our approach captures both layers.

As we saw in Section 5, per-unit non-response error varies across programs for both net and absolute error, but far less for supplement than for item non-response. We find a similar pattern for comparisons over time. For supplement non-respondents, per-unit absolute error is fairly stable, changing by less than 10% for pension and OASDI dollars and rising by about 20% for SSI dollars (Appendix Table A.11, Panel C). This stability suggests that the aggregate increase in supplement non-response error is driven more by the rising prevalence of non-respondents than by worsening error per non-respondent. In other words, the rise in inaccuracy caused by supplement non-respondents is roughly proportional to the rise in the supplement non-response rate, which makes this non-response rate a useful indicator of the impact of non-response on survey error. In contrast, for item non-respondents, per-unit absolute error roughly doubles for pension dollars (+101%) but is modest or declining for OASDI (+10%) and SSI (−8%), suggesting deterioration for specific sources. Therefore, the item non-response rates in Appendix Figure A.3 do not appear to be a useful indicator of the inaccuracy from item non-response. These patterns are similar but more pronounced for net error. Changes in net error per unit are larger for both types of non-response, and even the supplement non-response rate misses meaningful trends in bias because changes in per-unit severity play a greater role for net than for absolute error.

7. Implications for Improving Survey Statistics

7.1. Implications for Survey Users

Our findings carry implications for how researchers and policymakers should use and interpret CPS ASEC data and, by extension, other survey data that may exhibit similar error patterns. We first consider the treatment of non-respondents. While they constitute a minority of all observations, non-respondents cause (at least) as much bias per unit as respondents and add substantially more noise to the data. Specifically, supplement non-respondents add roughly as much bias per unit as respondents,[34] but they introduce noise into the data that can be particularly problematic for multivariate analyses. The scope to improve imputation methods is therefore limited, as additional predictors can at best close the small gap in bias relative to respondents and would need substantial predictive power to reduce the noise from prediction error. This combination of low bias and high noise suggests that *excluding* supplement non-respondents and *reweighting* the data is a more promising strategy. When data are missing at random, both imputation and reweighting are generally consistent under reasonable assumptions. We indeed find a small reduction in net error when we exclude supplement non-respondents and reweight the CPS public-use sample. But unlike imputation, reweighting shifts weight toward the less noisy reports and away from the noisy imputations, so it should mechanically reduce absolute error as well.

One may reasonably conjecture that this advantage of reweighting over imputation may extend to unit non-response more generally. The CPS already adjusts for non-response to the Basic Monthly CPS by reweighting to match key demographics. That we find evidence of good survey coverage, at least prior to the pandemic, suggests that treating this unit non-response as missing conditionally at random does not introduce substantial bias. While more speculative, this conjecture is in line with recent analyses of pre-pandemic unit non-response (e.g., Bee, Gathright, and Meyer 2015; Rothbaum and Bee 2021).

In a similar vein, one could use our methods to examine whether other weighting strategies, such as incorporating information from linked data (Rothbaum and Bee 2021; Rothbaum et al. 2021), further reduce survey error and whether similar gains extend to item non-respondents. However, item non-response is a harder problem to address, as it causes substantially more inaccuracy than supplement non-response and more bias per unit than respondents – making its imputation biased. Since the imputation procedures for supplement and item non-response are very similar, this difference indicates that the data missing due to item non-response are not missing conditionally at random, in line with the findings of Celhay, Meyer, and Mittag (2021) for SNAP

[34] In other words, the imputation procedure for ASEC non-respondents appears roughly unbiased, suggesting that the assumption that the data are missing conditionally at random is a reasonable approximation. See Rothbaum and Bee (2021) and Bee and Rothbaum (2025) for evidence that this proportionality may no longer hold after 2019.

and cash assistance. Better imputation methods are therefore likely to improve the problem but unlikely to solve it. Since both reweighting and imputation cause bias in this setting, the question of whether imputations for item non-respondents should be used is less clear than for supplement non-respondents (see Celhay, Meyer, and Mittag 2021 for a more extensive discussion).

Our approach can also help address measurement error, which we find to be the largest source of bias. For the general public and consumers of survey statistics, the severe understatement of program receipt and dollars received suggests caution in interpreting survey analyses of receipt and low incomes. Receipt rates and dollars received are likely understated and should be corrected upwards, potentially making use of our estimates. Biases in analyses over time are less clear-cut but could be corrected using estimates of the error trends. However, the rise in absolute error we document likely translates into more complicated biases in multivariate estimates, such as take-up models. Consequently, when working with the error-ridden data, survey users should consider correcting their estimates upwards (perhaps using linked administrative data) or imputing additional receipt in the survey data (see e.g., Mittag 2019).

7.2. Implications for Survey Producers

Our methods allow survey producers to comprehensively measure the consequences of survey design changes ex ante (in pilot studies) or ex post (by measuring error before and after a survey design change). Our measures directly indicate whether and where the survey has improved and whether there were any side effects such as increased inaccuracy or spillovers. Combined with information on costs, this can further allow survey producers to identify the design changes that reduce survey error most efficiently. For example, during our study period, the CPS ASEC moved to a dual-pass system and redesigned its questions on retirement income, both introduced with the 2014 ASEC and fielded to three-eighths of the sample in a split-sample test (Semega and Welniak 2013; Edwards and Creamer 2019). While detailed assessments of these reforms would require estimating and decomposing TSE right before and after the reforms, our results for a longer time window suggest at best modest effects. In a similar vein, our approach could be used to evaluate design changes the literature considers promising, such as additional prompts to aid reporting (Juster and Smith 1997; Lynn et al. 2006). It is equally suited to evaluate design choices that may trade off respondent burden against accuracy, such as moving to a single or summary income question (Micklewright and Schnepf 2010; Han, Meyer, and Sullivan 2020; Crossley, Fisher, and

Hussein 2023) or lengthening the reference period, as the SIPP did in 2014 when it moved from three interviews a year to one covering the previous calendar year (NAS 2018).

Our approach could also assist survey producers seeking to mitigate non-response more generally. Substantial resources have been invested to increase response rates (e.g., National Research Council 2013; Pew Research Center 2017). Yet several studies have found that the additional respondents recruited through such efforts only modestly improve survey representativeness (Keeter et al. 2006; Groves and Peytcheva 2008), consistent with our finding of generally good survey coverage. Our methods let producers assess whether these reforms can (cost-effectively) reduce bias, rather than just response rates. Our approach is particularly valuable in this context because our measure of survey error is comparable across components, which allows survey producers to quantify the potential trade-off between increasing response rates and reducing response accuracy. Several studies have argued that this trade-off exists because marginal respondents may report inaccurately.[35] Our approach provides a way to evaluate this trade-off empirically by quantifying how efforts to increase response rates affect both non-response error (rather than just rates) and measurement error. Our finding that a supplement non-respondent adds about as much bias to the survey as a respondent, in line with Kreuter, Müller, and Trappmann (2010), suggests that converting marginal unit non-respondents to respondents would do little to reduce sample selection, but also that raising survey response rates is unlikely to substantially increase reporting bias.

Yet, current survey practice institutionalizes the monitoring of non-response in a way it does not for measurement error. As noted earlier, non-response has standardized prevalence metrics (e.g., AAPOR 2023) that are tied to official quality standards and benchmarks for surveys (e.g., OMB 2006, especially Standard 1.3 and Appendix A). Response rates are also routinely used as headline indicators of survey data quality and concerns over declining response rates anchor discussions of the modern survey error crisis (e.g., Groves 2006, 2011; Brick and Williams 2013). Our results show that similar standards and routine reporting are needed for measurement error, and our approach supplies the relevant indicators. Importantly, our measures of non-response and measurement error are comparable, which enables a joint and comprehensive assessment that unifies the discussion of these two important survey error problems.

[35] See, e.g., Olson (2006) and Kreuter, Müller, and Trappmann (2010) for unit non-response and Bollinger and David (2001) and Hokayem, Bollinger, and Ziliak (2015) for item non-response.

The alarming extent of measurement error in program receipt raises the question of whether it would be preferable to directly supply information that is accurately recorded elsewhere, or an imputed version of it (see, e.g., Mittag 2019), by linking it to the survey. Because item non-response is not missing at random and measurement error is non-classical, this approach would simultaneously address the two most difficult survey error problems. Our results also point to the form and source of measurement error and hence suggest survey design adjustments when accurate measures cannot be substituted or imputed. Since extensive margin reporting is the dominant source of error, improving reporting by true recipients appears to be the most promising margin for reducing survey error. This could be achieved by probing further among likely recipients (e.g., those who report receipt of other programs or live in areas with high program participation). Reducing underreporting among respondents is especially promising since it would also improve the imputations built on reports and to which the net understatement extends. Therefore, reducing underreporting would also reduce a second important source of bias – non-response error.

As with non-response error, our approach is well suited to evaluating measures aimed at reducing measurement error because it captures the full effect of an intervention, including spillovers to imputation error. Our approach captures both net and absolute error, which matters because inaccuracy can be high even when net bias is low (as illustrated by SSI). Measuring both net and absolute error also distinguishes changes and corrections that genuinely reduce bias (e.g., by improving reporting among true recipients) from those that merely offset existing biases by introducing new errors. This distinction is particularly important for evaluating survey redesigns aimed at improving aggregate statistics (e.g., Hisnanick, Loveless, and Chesnut 2007) and corrections that adjust survey data to match aggregate income totals (e.g., Scholz, Moffitt, and Cowan 2009; Zedlewski and Giannarelli 2015; Mittag 2019; Ziliak 2026).

8. Conclusions

In this paper, we extend and implement a framework for measuring survey error through an empirical Total Survey Error decomposition. By expressing every error source in terms of its consequences for bias and inaccuracy, our framework provides a general measure that is of direct relevance and comparable over time and across sources. We apply our approach at scale, to the recipient share and mean dollars received for ten income sources or programs over two decades, demonstrating that survey accuracy can be monitored comprehensively, continuously, and at low cost. Our results reveal important insights about survey error and directly point to potential improvements in the CPS ASEC, one of the most important household surveys for both research and official statistics.

We find that the survey substantially understates both recipients and mean dollars for nearly all income sources. The understatement is large, at 40 percent or more for nearly half of the variables. The primary driver is measurement error rather than non-response, reversing the conventional focus on the latter as the main threat to survey accuracy. Within measurement error, failure to report receipt by true recipients constitutes the main source of bias. Both supplement and item non-response error also cause consistently negative bias, but the net bias is much smaller especially for item non-response error. The largest potential reductions in bias are therefore available from improving respondent reporting rather than from raising response rates, the margin that has tended to draw the most attention. And because imputations are based on donor responses, improving respondent reporting would propagate into improved imputed values among non-respondents. Yet, in absolute terms, non-respondents add disproportionate noise, so that relative to their share of the sample they are the largest source of inaccuracy. On the positive side, coverage prior to supplement non-response appears good for recipients of our ten programs and income sources.

Survey accuracy has also deteriorated over time. Net error has evolved unevenly across the income sources we examine, but this non-uniform trend masks a uniform increase in absolute error for every source, as both false positive and false negative rates have climbed. For nearly half of the variables, the total recipients or dollars that would need to change to eliminate all errors is at least 80% of the true total. Although measurement error remains the dominant source of bias, the sharpest deterioration over time has occurred among supplement and item non-respondents. Their errors have grown both because non-response has become more prevalent and because error among imputations has risen over time. Both non-response problems have thus become not only more common but also more problematic.

There are various extensions of our approach that could help both survey users and producers adapt it to their settings. For example, in cases where coverage error is more relevant than in our application, it may be of interest to further decompose it into sources such as frame error and unit non-response and to examine to what extent reweighting serves as a remedy (Krohn 2024). Coverage error can be further disaggregated by geography or demographic group. For survey users interested in multivariate analyses, it may be useful to augment our measure of inaccuracy with the degree to which survey error is related to their covariates. Our approach also extends naturally to the many other variables for which a measure of truth can be linked, as discussed in the introduction. It even remains useful when truth is more difficult to obtain than in our setting. Our housing assistance results show that our approach can deliver useful insights even when the administrative data are incomplete. The approach can also be applied to cases where the administrative data contain error. Net error components can still be estimated as long as that error cancels in the required aggregates. Even when it does not, the administrative records may be sufficiently more accurate than the survey to approximate the error components well (which seems likely in cases where the survey contains as much error as we document here).

For other surveys and variables where no measure of truth can be obtained, our results offer a basis for extrapolation. Income and program receipt are not the most sensitive questions in surveys, not more so than drug use, criminal behavior, or voting. Nor are they subject to framing and interpretation to the degree that questions on expectations and opinions are, although respondents are not always sure what a question covers, as our SSI results and the evidence on pension income (Bee and Mitchell 2017) indicate. Many of the income sources are received monthly and thus easier to recall than infrequent and unevenly timed events. What sets the variables we study apart is simply that good validation data exist. Where validation data have been available for other variables such as educational attainment (e.g., Black et al. 2003) and disease prevalence (Martin et al. 2000), error rates have likewise been high. Asking about more sensitive, ambiguous, or poorly recalled topics is harder still, and the limited evidence from such questions points in a similar direction. The patterns we highlight may therefore be as (if not more) severe elsewhere, which users of surveys would be wise to note.

Surveys measure far more than the income sources we examine, from employment and election predictions to attitudes and expectations. The income sources examined here are an ideal first application, because the accuracy of the linked data and the alignment of definitions and timing make our results clear. However, we do not yet know how the extent and components of survey error in income sources compare to these other domains. Establishing that requires applying a systematic measurement approach like ours to those cases as well. Surveys are one of the central instruments of measurement in the social sciences, but this paper shows that the instrument needs calibrating. Even in a flagship survey asking clear, well-tested questions about regular events, errors are substantial and often larger than its users assume. The tools we develop and apply in this paper make that calibration possible, turning survey error from an acknowledged caveat into a measurable and decomposable quantity.

References

Alwin, Duane F. 1991. "Research on Survey Quality." *Sociological Methods & Research*, 20(1): 3-29.

Alwin, Duane F. 2007. *Margins of Error: A Study of Reliability in Survey Measurement*. New York: Wiley.

American Association for Public Opinion Research (AAPOR). 2023. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys, 10th ed*. Alexandria, VA: AAPOR.

Atkeson, Lonna R., and R. Michael Alvarez (eds.). 2018. *The Oxford Handbook of Polling and Survey Methods*. Oxford University Press.

Bee, C. Adam, and Joshua W. Mitchell. 2017. "Do Older Americans Have More Income Than We Think?" SEHSD Working Paper 2017-39. Washington, DC: U.S. Census Bureau.

Bee, C. Adam, Graton Gathright, and Bruce D. Meyer. 2015. "Unit Non-Response Bias in the Current Population Survey." Working Paper.

Bee, C. Adam, Joshua Mitchell, Nikolas Mittag, Jonathan Rothbaum, Carl Sanders, Lawrence Schmidt, and Matthew Unrath. 2023. "National Experimental Wellbeing Statistics Version 1." SEHSD Working Paper 2023-02. Washington, DC: U.S. Census Bureau.

Bee, Adam, and Jonathan Rothbaum. 2025. "Using Administrative Data to Evaluate Nonresponse Bias in the 2025 Current Population Survey Annual Social and Economic Supplement." Research Matters (U.S. Census Bureau Blog). Washington, DC: U.S. Census Bureau.

Biemer, Paul P. 2010. "Total Survey Error: Design, Implementation, and Evaluation." *Public Opinion Quarterly*, 74(5): 817–848.

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.

Bollinger, Christopher R. and Martin H. David. 2001. "Estimation with Response Error and Nonresponse: Food-Stamp Participation in the SIPP." *Journal of Business & Economic Statistics*, 19(2): 129-141.

Bollinger, Christopher R., Barry T. Hirsch, Charles M. Hokayem, and James P. Ziliak. 2019. "Trouble in the Tails? What We Know about Earnings Nonresponse Thirty Years after Lillard, Smith, and Welch." *Journal of Political Economy*, 127(5): 2143-2185.

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

Brick, J. Michael, and Douglas Williams. 2013. "Explaining Rising Nonresponse Rates in Cross-Sectional Surveys." *ANNALS of the American Academy of Political and Social Science*, 645(1): 36-59.

Celhay, Pablo A., Bruce D. Meyer, and Nikolas Mittag. 2021. "Errors in Reporting and Imputation of Government Benefits and Their Implications." NBER Working Paper No. 29184.

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." SEHSD Working Paper No. 2016-18. Washington, DC: U.S. Census Bureau.

Coibion, Olivier, and Yuriy Gorodnichenko. 2026. "Schumpeter Lecture 2025: The New Causal Macroeconomics of Surveys and Experiments." *Journal of the European Economic Association*, 24(1): 58-81.

Corinth, Kevin, Bruce D. Meyer, and Derek Wu. 2022. "The Effect of Public Transfer Income on Measures of Poverty and Income Inequality." *AEA Papers and Proceedings*, 112: 98-102.

Crossley, Thomas F., Paul Fisher, and Omar Hussein. 2023. "Assessing Data from Summary Questions about Earnings and Income." *Labour Economics*, 81: 102331.

Czajka, John L., Karen Cunnyngham, and Randy Rosso. 2015. "Simulated Versus Actual SNAP Unit Composition in Survey Households in Two States." Working Paper.

de Leeuw, Edith D., and Wim de Heer. 2002. "Trends in Household Survey Nonresponse: A Longitudinal and International Comparison." In *Survey Nonresponse*, edited by Robert M. Groves, Don A. Dillman, John L. Eltinge, and Roderick J. A. Little, 41-54. New York: Wiley.

Edwards, Ashley, and John Creamer. 2019. "Updating the Current Population Survey Processing System and Bridging Differences in the Measurement of Poverty." SEHSD Working Paper 2019-05. Washington, DC: U.S. Census Bureau.

Gathright, Graton, and Tyler Crabb. 2014. "Reporting of SSA Program Participation in SIPP." Working Paper.

Gelman, Andrew. 2021. "Failure and Success in Political Polling and Election Forecasting." *Statistics and Public Policy*, 8(1): 67-72.

Greene, William H. 2017. *Econometric Analysis, 8th ed.* Upper Saddle River, NJ: Pearson.

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

Groves, Robert M. 2006. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *Public Opinion Quarterly*, 70(5): 646-675.

Groves, Robert M. 2011. "Three Eras of Survey Research." *Public Opinion Quarterly*, 75(5): 861-871.

Groves, Robert M., Don A. Dillman, John L. Eltinge, and Roderick J. A. Little (Eds.). 2002. *Survey Nonresponse*. New York: Wiley.

Groves, Robert M., and Lars Lyberg. 2010. "Total Survey Error: Past, Present, and Future." *Public Opinion Quarterly*, 74(5): 849-879.

Groves, Robert M., and Emilia Peytcheva. 2008. "The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis." *Public Opinion Quarterly*, 72(2): 167-189.

Han, Jeehoon, Bruce D. Meyer, and James X. Sullivan. 2020. "Income and Poverty in the COVID-19 Pandemic." *Brookings Papers on Economic Activity*, 51(2): 85–118.

Hisnanick, John J., Tracy Loveless, and John Chesnut. 2007. "Evaluation Report Covering Receipt of Food Stamps." 2006 American Community Survey Content Test Report H.6. Washington, DC: U.S. Census Bureau.

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

Hokayem, Charles, Trivellore Raghunathan, and Jonathan Rothbaum. 2015. "The Impact of Income Nonresponse on Estimates of the Distribution of Earnings and Household Income from the Current Population Survey: A Multiple Imputation Approach." Working Paper.

Huynh, Minh, Kalman Rupp, and James Sears. 2000. "The Assessment of Survey of Income and Program Participation (SIPP) Benefit Data Using Longitudinal Administrative Records." SEHSD Working Paper 2000-09. Washington, DC: U.S. Census Bureau.

Juster, F. Thomas, and James P. Smith. 1997. "Improving the Quality of Economic Data: Lessons from the HRS and AHEAD." *Journal of the American Statistical Association*, 92(440): 1268–1278.

Keeter, Scott, Courtney Kennedy, Michael Dimock, Jonathan Best, and Peyton Craighill. 2006. "Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey." *Public Opinion Quarterly*, 70(5): 759-779.

Kreuter, Frauke, Gerrit Müller, and Mark Trappmann. 2010. "Nonresponse and Measurement Error in Employment Research: Making Use of Administrative Data." *Public Opinion Quarterly*, 74(5): 880-906.

Krohn, Ethan. 2024. "Earnings Through the Stages: Using Tax Data to Test for Sources of Error in CPS ASEC Earnings and Inequality Measures." CES Discussion Paper 24-52. Washington, DC: U.S. Census Bureau.

Little, Roderick J. A., and Donald B. Rubin. 2002. *Statistical Analysis with Missing Data, 2nd ed.* Hoboken, NJ: Wiley.

Lynn, Peter, Annette Jäckle, Stephen P. Jenkins, and Emanuela Sala. 2006. "The Effects of Dependent Interviewing on Responses to Questions on Income Sources." *Journal of Official Statistics*, 22(3): 357–384.

Martin, Linda M., Marilyn Leff, Ned Calonge, Carol Garrett, and David E. Nelson. 2000. "Validation of Self-Reported Chronic Conditions and Health Services in a Managed Care Population." *American Journal of Preventive Medicine*, 18(3): 215-218.

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

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

Meyer, Bruce D., and Nikolas Mittag. 2021. "An Empirical Total Survey Error Decomposition Using Data Combination." *Journal of Econometrics*, 224(2): 286-305.

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

Meyer, Bruce D., and Derek Wu. 2021. "Do Surveys Miss the Poor? Evidence from Administrative SNAP Records." 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.

Micklewright, John, and Sylke V. Schnepf. 2010. "How Reliable are Income Data Collected with a Single Question?" *Journal of the Royal Statistical Society*: Series A, 173(2): 409–429.

Mittag, Nikolas. 2019. "Correcting for Misclassification Bias in Survey Data Using Response Variance Estimates." *American Economic Journal: Economic Policy*, 11(4): 229-252.

Mulry, Mary H. 2007. "Summary of Accuracy and Coverage Evaluation for Census 2000." Working Paper.

Mulry, Mary H., and Bruce D. Spencer. 2001. "Accuracy and Coverage Evaluation: Overview of Total Error Modeling and Loss Function Analysis." DSSD Census 2000 Procedures and Operations Memorandum Series B-19. Washington, DC: U.S. Census Bureau.

National Academies of Sciences, Engineering, and Medicine (NAS). 2017. "The Value of National Statistics in the United States." In *Principles and Practices for a Federal Statistical Agency: Sixth Ed.*, 9–16. Washington, DC: The National Academies Press.

National Academies of Sciences, Engineering, and Medicine (NAS). 2018. *The 2014 Redesign of the Survey of Income and Program Participation: An Assessment*. Washington, DC: The National Academies Press.

National Research Council. 2013. *Nonresponse in Social Science Surveys: A Research Agenda.* Edited by Roger Tourangeau and Thomas J. Plewes. Washington, DC: The National Academies Press.

Office of Management and Budget (OMB). 2006. *Standards and Guidelines for Statistical Surveys.* Washington, DC: Executive Office of the President.

O'Hare, William, J. Gregory Robinson, Kirsten West, and Thomas Wood. 2016. "Comparing the U.S. Decennial Census Coverage Estimates for Children from Demographic Analysis and Coverage Measurement Surveys." *Population Research and Policy Review*, 35(5): 685-704.

Olsen, Edgar O. 2018. "Reported and Actual Housing Assistance Based on CPS and AHS Variables." Unpublished Note, January 30, 2018.

Olson, Kristen. 2006. "Survey Participation, Nonresponse Bias, Measurement Error Bias, and Total Bias." *Public Opinion Quarterly*, 70(5): 737-758.

Pew Research Center. 2017. *What Low Response Rates Mean for Telephone Surveys.* Washington, DC: Pew Research Center.

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.

Rothbaum, Jonathan. 2015. "Comparing Income Aggregates: The Importance of Accounting for Taxes and Transfers in CPS and National Income and Product Account Estimates." SEHSD Working Paper 2015-01. Washington, DC: U.S. Census Bureau.

Rothbaum, Jonathan, and C. Adam Bee. 2021. "Coronavirus Infects Surveys, Too: Nonresponse Bias During the Pandemic in the CPS ASEC." SEHSD Working Paper 2021-01. Washington, DC: U.S. Census Bureau.

Rothbaum, Jonathan, Jonathan Eggleston, Adam Bee, Mark Klee, and Brian Mendez-Smith. 2021. "Addressing Nonresponse Bias in the American Community Survey During the Pandemic Using Administrative Data." American Community Survey Research and Evaluation Report Memorandum Series ACS21-RER-04. Washington, DC: U.S. Census Bureau.

Sabety, Adrienne, and Ariella Spitzer. 2026. "Missing Black Men? The Impact of Non-Reporting on Estimates of Labor Market Outcomes for Black Men." NBER Working Paper No. 34724.

Scholz, John Karl, Robert Moffitt, and Benjamin Cowan. 2009. "Trends in Income Support." In *Changing Poverty, Changing Policies*, edited by Maria Cancian and Sheldon Danziger, 203–241. New York: Russell Sage Foundation.

Seeskin, Zachary H., and Bruce D. Spencer. 2018. "Correcting for Population Displacement in Census Coverage Estimates." *Journal of Survey Statistics and Methodology*, 6(4): 502-523.

Semega, Jessica L., and Edward Welniak, Jr. 2013. "Evaluating the 2013 CPS ASEC Income Redesign Content Test." Proceedings of the 2013 Federal Committee on Statistical Methodology Research Conference.

Stantcheva, Stefanie. 2023. "How to Run Surveys: A Guide to Creating Your Own Identifying Variation and Revealing the Invisible." *Annual Review of Economics*, 15: 205-234.

U.S. Census Bureau. 2002. "Current Population Survey: Design and Methodology." Technical Paper 63RV. Washington, DC: U.S. Census Bureau.

U.S. Census Bureau. 2019. "Current Population Survey: Design and Methodology." Technical Paper 77. Washington, DC: U.S. Census Bureau.

U.S. Census Bureau. 2021. "Imputation of Unreported Data Items." Current Population Survey, Technical Documentation. Last Revised October 8, 2021.

U.S. Census Bureau. 2025. "Current Population Survey, 2025 Annual Social and Economic (ASEC) Supplement: Technical Documentation." Washington, DC: U.S. Census Bureau.

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

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

Zedlewski, Sheila, and Linda Giannarelli. 2015. "TRIM: A Tool for Social Policy Analysis." Washington, DC: Urban Institute.

Ziliak, James. 2026. "Using Survey-Based Regression Imputation to Address Transfer Underreporting in the CPS ASEC: Implications for Inequality and Poverty 2001-2024." Presentation Slides.

Tables and Figures

Table 1. Income Sources Analyzed in this Paper

	Survey Data (CPS)		Linked Administrative Data		
Income Source	**Unit of Analysis**	**Available in 1996**	**Source**	**States Covered**	**Available in 1996**
Pensions	Individuals 15+	Yes	Form 1099-R (IRS)	All	Yes
OASDI	Individuals 15+	Yes	PHUS & MBR (SSA)	All	Yes
OASI	Individuals 15+	No	PHUS & MBR (SSA)	All	Yes
DI	Individuals 15+	No	PHUS & MBR (SSA)	All	Yes
SSI	Individuals 15+	Yes	SSR (SSA)	All	Yes
UI	Individuals 15+	Yes	Form 1099-G (IRS)	All	No
VA Disability	Individuals 15+	Yes	US VETS (VA)	All	No
SNAP	Households	Yes	State Agencies	5 States	No
Medicaid	All Individuals	Yes	MSIS (CMS)	All	No
Housing Assistance	Households	Yes	PIC & TRACS (HUD)	All	Yes

Notes: This table lists the ten income sources that we focus on as part of our total survey error decomposition. For each income source, we first list based on the survey data the unit of analysis in sample and whether a given income source is available in the CPS ASEC in 1996 (the first year that we analyze). Note that OASI and DI are not available in the CPS in 1996 because the survey question on reasons for receiving OASDI (needed to identify OASI vs. DI) is only asked starting in 2000. We also list for each income source the administrative data source, the states covered by the administrative data we use, and whether that administrative source is available in 1996.

Table 2. Total Recipients and Dollars that Should Be Covered by the Survey Sample, 2017

	Pensions	OASDI	OASI	DI	SSI	UI	VA Disability	SNAP	Medicaid	Housing Assistance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Recipients (millions)										
Recipients in Admin Microdata	53.27	66.09	54.36	11.73	9.00	7.54	4.78	3.44	89.84	5.40
Subtraction of Recipients Not Covered by Household Samples (%)										
Decedents	2.81%	1.95%	2.20%	0.71%	0.78%	0.26%	1.44%	1.35%	1.25%	1.83%
Estimated Recipients in GQs	0.95%	3.01%	3.01%	3.01%	3.01%	0.91%	1.61%	7.21%	2.94%	3.61%
Recipients in Overseas & Territories	--	1.57%	1.52%	1.83%	0.01%	--	--	--	--	--
Subtraction of Recipients Not Covered by Linked Data (%)										
Unlinkable Admin Records	1.33%	--	--	--	--	0.06%	0.04%	0.10%	2.31%	0.18%
Estimated Residents of Other States	--	--	--	--	--	--	--	1.01%	--	--
Total Subtraction	5.09%	6.52%	6.67%	5.61%	3.80%	1.23%	3.10%	9.67%	6.49%	5.62%
Total Recipients Covered	50.56	61.78	50.71	11.11	8.66	7.45	4.63	3.11	84.01	5.10
B. Dollars ($ millions)										
Payments in Admin Microdata	1,203,000	941,554	798,728	142,826	54,509	31,530	72,160	7,197	--	--
Subtraction of Payments Not Covered by Household Samples (%)										
Payments to Decedents	1.95%	2.04%	2.28%	0.74%	0.71%	0.23%	1.62%	0.63%	--	--
Estimated Payments to GQs	0.68%	2.35%	2.35%	2.35%	4.46%	0.81%	1.52%	6.84%	--	--
Payments to Overseas & Territories	--	1.57%	1.52%	1.83%	0.01%	--	--	--	--	--
Subtraction of Payments Not Covered by Linked Data (%)										
Unlinkable Admin Records	0.80%	--	--	--	--	0.04%	0.05%	0.03%	--	--
Estimated Payments to Out-of-State Residents	--	--	--	--	--	--	--	1.22%	--	--
Total Subtraction	3.43%	5.96%	6.08%	5.00%	5.18%	1.08%	3.18%	8.73%	--	--
Total Payments Covered	1,161,754	885,459	749,761	136,057	51,683	31,190	69,865	6,568	--	--

Notes: This table shows aggregate recipients and dollars from our administrative data, as well as the adjustments that we apply to account for intentional coverage differences with the survey. Cells marked "--" are not calculated. All amounts are in millions. For OASDI, OASI, DI, and SSI, universe administrative data are unavailable because the PHUS, MBR, and SSR contain only PIKed individuals matched to the CPS; estimates for these sources therefore rely on publicly available administrative aggregates.

Table 3. Total Survey Error for Recipients and Its Components (as a % of Survey Targets), 2017

	Pensions	OASDI	OASI	DI	SSI	UI	VA Disability	SNAP	Medicaid	Housing Assistance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Coverage Error	2.3	-3.9	-2.3	-11.5	-17.9	-4.5	-3.2	-25.1	-3.7	7.1
B. Supplement Non-Response Error										
False Positives	3.9	3.2	3.3	8.2	10.4	7.5	9.3	5.9	7.7	9.8
False Negatives	-12.9	-4.9	-4.5	-12.3	-11.4	-18.0	-17.3	-14.2	-12.5	-15.2
Total Net Error Due to Suppl. Non-Response	-9.0	-1.7	-1.2	-4.1	-0.9	-10.4	-8.1	-8.3	-4.7	-5.5
Total Absolute Error Due to Suppl. Non-Response	16.8	8.1	7.8	20.5	21.8	25.5	26.6	20.1	20.2	25.0
C. Item Non-Response Error										
False Positives	1.2	0.5	0.6	1.5	1.5	0.6	2.1	--	5.8	4.1
False Negatives	-4.1	-0.9	-0.8	-2.6	-2.7	-2.5	-2.8	--	-5.3	-5.1
Total Net Error Due to Item Non-Response	-2.8	-0.3	-0.2	-1.1	-1.3	-1.8	-0.7	--	0.5	-1.0
Total Absolute Error Due to Item Non-Response	5.3	1.4	1.3	4.0	4.2	3.1	4.8	--	11.2	9.2
D. Measurement Error										
False Positives	5.8	3.0	4.1	18.9	21.4	6.2	2.0	2.6	3.5	32.9
False Negatives	-42.1	-11.6	-11.3	-34.3	-26.5	-40.1	-30.1	-20.5	-26.3	-21.9
Total Net Error Due to Measurement Error	-36.3	-8.6	-7.1	-15.4	-5.1	-33.8	-28.1	-17.9	-22.8	11.0
Total Absolute Error Due to Measurement Error	47.9	14.5	15.4	53.2	47.9	46.3	32.1	23.1	29.8	54.9
Total Net Error	-45.9	-14.6	-10.7	-32.0	-25.2	-50.6	-40.0	-51.3	-30.8	11.6
Total Absolute Error	72.3	27.9	26.8	89.3	91.7	79.4	66.7	68.3	64.9	96.2

Notes: Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as shares of coverage-adjusted administrative aggregates. For each reporting group (supplement non-respondents, item non-respondents, and respondents), we report false positives, false negatives, and their sum, defined as net error due to mis-imputed or mis-reported receipt status, which corresponds to total error due to the reporting type. Absolute error reports the same quantity using absolute values.

Table 4. Total Net Survey Error for Dollar Amounts and Its Components (as a % of Survey Targets), 2017

	Pensions	OASDI	OASI	DI	SSI	UI	VA Disability	SNAP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Coverage Error	5.9	1.2	1.7	-0.9	-17.0	-4.5	-2.0	-6.2
B. Supplement Non-Response Error								
False Positives	3.5	3.1	3.3	8.7	14.9	8.0	9.0	4.8
False Negatives	-11.0	-4.7	-4.3	-13.5	-11.9	-17.6	-17.6	-19.1
Net Error due to Mis-Imputed Receipt Status	-7.5	-1.6	-1.0	-4.8	3.0	-9.6	-8.7	-14.3
Net Error in Imputed Amounts	-0.4	0.2	0.3	-0.3	0.6	0.3	-0.1	-1.0
Total Net Error Due to Suppl. Non-Response	-7.9	-1.4	-0.8	-5.0	3.5	-9.3	-8.8	-15.3
C. Item Non-Response Error								
False Positives	1.1	0.5	0.5	1.6	1.9	0.8	2.2	--
False Negatives	-4.3	-0.8	-0.7	-2.7	-3.0	-2.2	-3.6	--
Net Error due to Mis-Imputed Receipt Status	-3.2	-0.3	-0.2	-1.1	-1.1	-1.4	-1.4	--
Net Error in Imputed Amounts	-0.5	0.0	-0.1	0.0	-0.1	--	--	--
Total Net Error Due to Item Non-Response	-3.8	-0.4	-0.2	-1.1	-1.2	-1.4	-1.4	--
D. Measurement Error								
False Positives	4.1	2.4	3.7	17.8	36.1	7.4	1.7	2.5
False Negatives	-31.7	-9.9	-9.2	-36.1	-23.3	-32.8	-27.1	-21.6
Net Error due to Mis-Reported Receipt Status	-27.6	-7.4	-5.5	-18.3	12.8	-25.4	-25.3	-19.0
Net Error in Reported Amounts	-10.6	-0.1	0.0	-0.3	3.8	-1.8	-0.5	-6.3
Total Net Error Due to Measurement Error	-38.2	-7.5	-5.5	-18.6	16.6	-27.2	-25.9	-25.3
Total Net Error	-44.0	-8.1	-4.9	-25.7	1.9	-42.3	-38.0	-46.9

Notes: Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as shares of coverage-adjusted administrative aggregates. For each reporting group (supplement non-respondents, item non-respondents, and respondents), we report false positives, false negatives, and their sum, defined as net error due to mis-imputed or mis-reported receipt status. Error in imputed or reported amounts reflects dollar discrepancies for correct reporting recipients. Total net error equals the sum of net receipt-status error and error in amounts.

Table 5. Total Absolute Survey Error for Dollar Amounts and Its Components (as a % of Survey Targets), 2017

	Pensions	OASDI	OASI	DI	SSI	UI	VA Disability	SNAP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Coverage Error	5.9	1.2	1.7	-0.9	-17.0	-4.5	-2.0	-6.2
B. Supplement Non-Response Error								
False Positives	3.5	3.1	3.3	8.7	14.9	8.0	9.0	4.8
False Negatives	-11.0	-4.7	-4.3	-13.5	-11.9	-17.6	-17.6	-19.1
Absolute Error due to Mis-Imputed Receipt Status	14.5	7.8	7.6	22.2	26.8	25.7	26.6	23.9
Absolute Error in Imputed Amounts	4.9	5.1	5.2	1.3	1.8	0.9	1.4	2.6
Total Absolute Error Due to Suppl. Non-Response	19.4	12.9	12.8	23.6	28.6	26.6	28.1	26.5
C. Item Non-Response Error								
False Positives	1.1	0.5	0.5	1.6	1.9	0.8	2.2	--
False Negatives	-4.3	-0.8	-0.7	-2.7	-3.0	-2.2	-3.6	--
Absolute Error due to Mis-Imputed Receipt Status	5.4	1.3	1.2	4.3	4.8	2.9	5.9	--
Absolute Error in Imputed Amounts	3.0	1.2	1.2	0.3	0.6	--	--	--
Total Absolute Error Due to Item Non-Response	8.4	2.4	2.4	4.6	5.4	2.9	5.9	--
D. Measurement Error								
False Positives	4.1	2.4	3.7	17.8	36.1	7.4	1.7	2.5
False Negatives	-31.7	-9.9	-9.2	-36.1	-23.3	-32.8	-27.1	-21.6
Absolute Error due to Mis-Reported Receipt Status	35.8	12.3	13.0	53.8	59.4	40.1	28.8	24.1
Absolute Error in Reported Amounts	26.7	16.2	16.2	10.3	13.7	18.7	20.9	14.2
Total Absolute Error Due to Measurement Error	62.5	28.5	29.2	64.1	73.1	58.9	49.7	38.3
Total Absolute Error	96.2	45.1	46.1	93.2	124.0	92.9	85.6	71.1

Notes: Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as shares of coverage-adjusted administrative aggregates. For each reporting group (supplement non-respondents, item non-respondents, and respondents), we report false positives, false negatives, and their sum, defined as absolute error due to mis-imputed or mis-reported receipt status. Error in imputed or reported amounts reflects dollar discrepancies for correct reporting recipients. Total absolute error equals the sum of net receipt-status error and error in amounts.

Table 6. False Negative and False Positive Rates for Income Sources, 2017

	Pensions	OASDI	OASI	DI	SSI	UI	VA Disability	SNAP	Medicaid	Housing Assistance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. False Negative Rates (%)										
Overall	57.7	18.0	16.9	55.5	49.4	63.4	51.9	46.3	45.8	39.5
For Supplement Non-Respondents	74.7	31.5	28.9	83.7	81.1	95.6	93.7	77.4	67.8	85.6
For Item Non-Respondents	64.7	26.9	22.9	81.8	77.0	--	--	--	47.4	68.3
For Respondents	53.5	15.0	14.3	48.5	41.0	54.0	39.9	36.2	39.5	26.8
B. False Positive Rates (%)										
Overall	2.6	2.0	1.9	1.3	1.1	0.4	0.2	2.2	5.9	2.0
For Supplement Non-Respondents	4.8	4.9	4.0	1.9	1.9	1.2	0.9	7.3	14.7	2.2
For Item Non-Respondents	8.8	7.5	6.1	2.7	2.9	1.0	1.4	--	24.6	14.8
For Respondents	1.8	1.2	1.3	1.1	0.9	0.2	0.0	0.8	1.7	1.7

Notes: This table shows false negative rates (the share of true recipients who do not report receipt in the survey) and false positive rates (the share of true non-recipients who do report receipt in the survey) calculated for each income source and reporting type. Sample consists of individuals or households in the linked CPS sample for reference year 2017, dropping non-PIKed units and adjusting survey weights using inverse probability weights. Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules.

Table 7. Total Survey Error for Recipients and Its Components (as a % of Survey Targets), 1996 and 2017

	Pensions		OASDI		SSI		Housing Assistance	
	1996	2017	1996	2017	1996	2017	1996	2017
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Coverage Error	0.0	2.3	-9.4	-3.9	-13.7	-17.9	3.4	7.1
B. Supplement Non-Response Error								
False Positives	2.1	3.9	1.5	3.2	4.3	10.4	4.5	9.8
False Negatives	-5.6	-12.9	-1.6	-4.9	-4.0	-11.4	-5.3	-15.2
Total Net Error Due to Suppl. Non-Response	-3.5	-9.0	0.0	-1.7	0.3	-0.9	-0.8	-5.5
Total Absolute Error Due to Suppl. Non-Response	7.7	16.8	3.1	8.1	8.3	21.8	9.8	25.0
C. Item Non-Response Error								
False Positives	0.5	1.2	0.2	0.5	1.3	1.5	2.9	4.1
False Negatives	-1.2	-4.1	-0.3	-0.9	-1.4	-2.7	-1.0	-5.1
Total Net Error Due to Item Non-Response	-0.7	-2.8	0.0	-0.3	-0.1	-1.3	2.0	-1.0
Total Absolute Error Due to Item Non-Response	1.7	5.3	0.5	1.4	2.7	4.2	3.9	9.2
D. Measurement Error								
False Positives	6.2	5.8	3.4	3.0	13.1	21.4	38.9	32.9
False Negatives	-49.1	-42.1	-7.5	-11.6	-22.4	-26.5	-27.8	-21.9
Total Net Error Due to Measurement Error	-42.9	-36.3	-4.1	-8.6	-9.3	-5.1	11.1	11.0
Total Absolute Error Due to Measurement Error	55.3	47.9	10.9	14.5	35.5	47.9	66.8	54.9
Total Net Error	-47.1	-45.9	-13.6	-14.6	-22.8	-25.2	15.7	11.6
Total Absolute Error	64.7	72.3	23.9	27.9	60.1	91.7	83.8	96.2

Notes: All reported values are expressed as shares of coverage-adjusted administrative aggregates. For each reporting group (supplement non-respondents, item non-respondents, and respondents), we report false positives, false negatives, and their sum, defined as net error due to mis-imputed or mis-reported receipt status, which corresponds to total error due to the reporting type. Absolute error reports the same quantity using absolute values.

Table 8. Total Net Survey Error for Dollar Amounts and Its Components (as a % of Survey Targets), 1996 and 2017

	Pensions		OASDI		SSI	
	1996	2017	1996	2017	1996	2017
	(1)	(2)	(3)	(4)	(5)	(6)
A. Coverage Error	-4.4	5.9	-2.1	1.2	-12.6	-17.0
B. Supplement Non-Response Error						
False Positives	1.6	3.5	1.3	3.1	4.8	14.9
False Negatives	-4.9	-11.0	-1.4	-4.7	-3.9	-11.9
Net Error due to Mis-Imputed Receipt Status	-3.3	-7.5	-0.1	-1.6	0.8	3.0
Net Error in Imputed Amounts	-0.5	-0.4	-0.3	0.2	-0.1	0.6
Total Net Error Due to Suppl. Non-Response	-3.8	-7.9	-0.4	-1.4	0.8	3.5
C. Item Non-Response Error						
False Positives	0.3	1.1	0.2	0.5	1.3	1.9
False Negatives	-0.8	-4.3	-0.2	-0.8	-1.3	-3.0
Net Error due to Mis-Imputed Receipt Status	-0.5	-3.2	0.0	-0.3	-0.1	-1.1
Net Error in Imputed Amounts	0.0	-0.5	0.0	0.0	--	-0.1
Total Net Error Due to Item Non-Response	-0.5	-3.8	0.0	-0.4	-0.1	-1.2
D. Measurement Error						
False Positives	4.4	4.1	2.9	2.4	13.9	36.1
False Negatives	-39.3	-31.7	-6.0	-9.9	-21.0	-23.3
Net Error due to Mis-Reported Receipt Status	-34.9	-27.6	-3.1	-7.4	-7.1	12.8
Net Error in Reported Amounts	-13.5	-10.6	-0.5	-0.1	5.8	3.8
Total Net Error Due to Measurement Error	-48.5	-38.2	-3.6	-7.5	-1.4	16.6
Total Net Error	-57.2	-44.0	-6.0	-8.1	-13.2	1.9

Notes: Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as shares of coverage-adjusted administrative aggregates. For each reporting group (supplement non-respondents, item non-respondents, and respondents) we report false positives, false negatives, and their sum, defined as net error due to mis-imputed or mis-reported receipt status. Error in imputed or reported amounts reflects dollar discrepancies for correct reporting recipients. Total net error equals the sum of net receipt-status error and error in amounts.

Table 9. Total Absolute Survey Error for Dollar Amounts and Its Components (as a % of Survey Targets), 1996 and 2017

	Pensions		OASDI		SSI	
	1996	2017	1996	2017	1996	2017
	(1)	(2)	(3)	(4)	(5)	(6)
A. Coverage Error	-4.4	5.9	-2.1	1.2	-12.6	-17.0
B. Supplement Non-Response Error						
False Positives	1.6	3.5	1.3	3.1	4.8	14.9
False Negatives	-4.9	-11.0	-1.4	-4.7	-3.9	-11.9
Absolute Error due to Mis-Imputed Receipt Status	6.5	14.5	2.7	7.8	8.7	26.8
Absolute Error in Imputed Amounts	1.6	4.9	2.2	5.1	1.2	1.8
Total Absolute Error Due to Suppl. Non-Response	8.1	19.4	5.0	12.9	9.9	28.6
C. Item Non-Response Error						
False Positives	0.3	1.1	0.2	0.5	1.3	1.9
False Negatives	-0.8	-4.3	-0.2	-0.8	-1.3	-3.0
Absolute Error due to Mis-Imputed Receipt Status	1.0	5.4	0.4	1.3	2.6	4.8
Absolute Error in Imputed Amounts	0.5	3.0	0.3	1.2	--	0.6
Total Absolute Error Due to Item Non-Response	1.5	8.4	0.7	2.4	2.6	5.4
D. Measurement Error						
False Positives	4.4	4.1	2.9	2.4	13.9	36.1
False Negatives	-39.3	-31.7	-6.0	-9.9	-21.0	-23.3
Absolute Error due to Mis-Reported Receipt Status	43.6	35.8	9.0	12.3	34.9	59.4
Absolute Error in Reported Amounts	22.2	26.7	14.1	16.2	18.5	13.7
Total Absolute Error Due to Measurement Error	65.8	62.5	23.0	28.5	53.4	73.1
Total Absolute Error	79.8	96.2	30.8	45.1	78.5	124.0

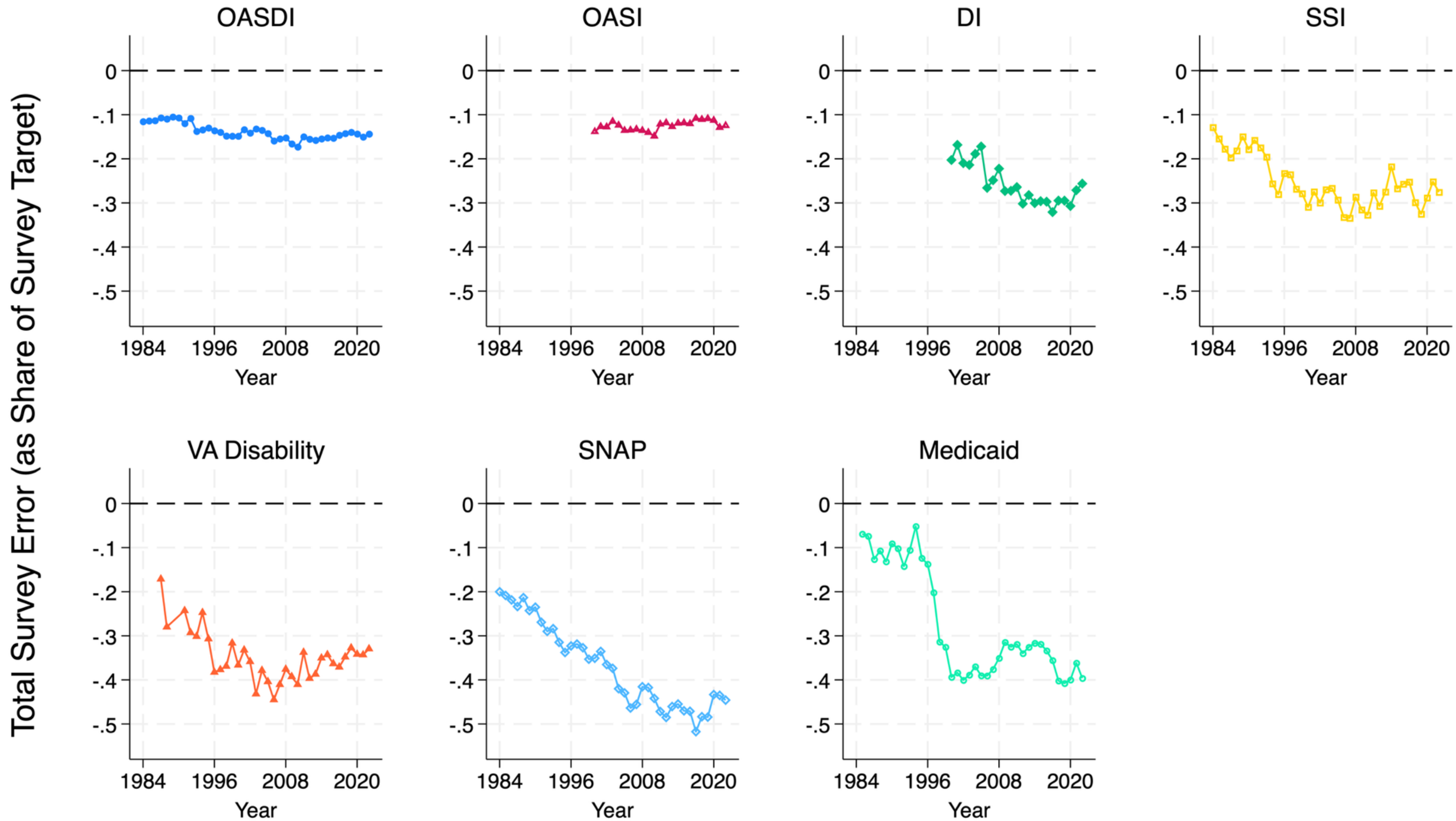
Notes: Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as shares of coverage-adjusted administrative aggregates. For each reporting group (supplement non-respondents, item non-respondents, and respondents) we report false positives, false negatives, and their sum, defined as absolute error due to mis-imputed or mis-reported receipt status. Error in imputed or reported amounts reflects dollar discrepancies for correct reporting recipients. Total absolute error equals the sum of net receipt-status error and error in amounts.

Table 10. False Negative and False Positive Rates for Income Sources, 1996 and 2017

	Pensions		OASDI		SSI		Housing Assistance	
	1996	2017	1996	2017	1996	2017	1996	2017
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. False Negative Rates (%)								
Overall	55.9	57.7	10.2	18.0	32.2	49.4	33.0	39.5
For Supplement Non-Respondents	77.5	74.7	23.9	31.5	69.6	81.1	70.5	85.6
For Item Non-Respondents	65.4	64.7	26.6	26.9	--	77.0	40.4	68.3
For Respondents	54.0	53.5	9.0	15.0	28.3	41.0	29.8	26.8
B. False Positive Rates (%)								
Overall	1.7	2.6	1.3	2.0	0.6	1.1	2.1	2.0
For Supplement Non-Respondents	5.0	4.8	5.0	4.9	1.8	1.9	2.5	2.2
For Item Non-Respondents	7.9	8.8	6.8	7.5	5.5	2.9	12.6	14.8
For Respondents	1.3	1.8	1.0	1.2	0.5	0.9	1.9	1.7

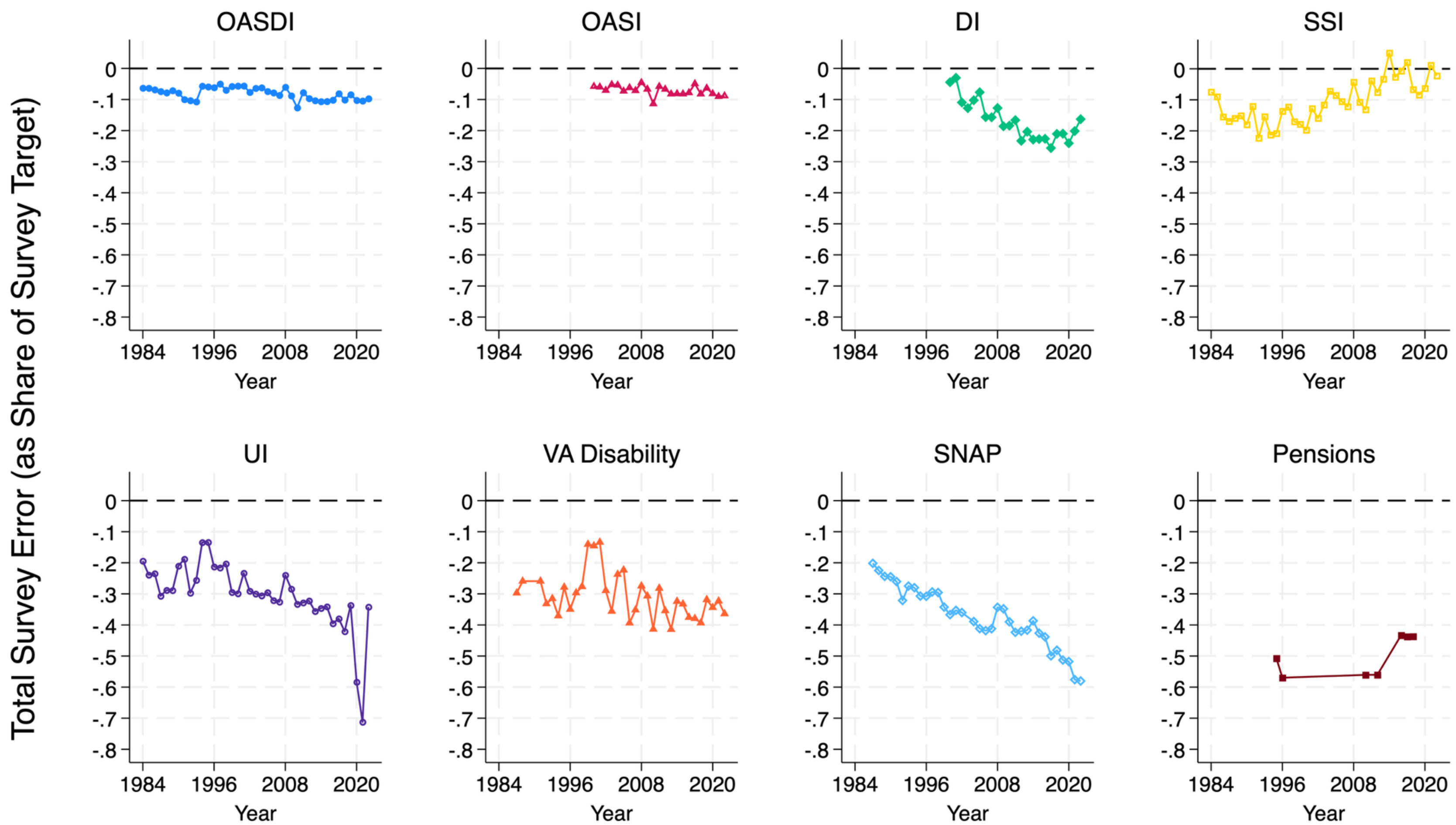
Notes: This table shows false negative rates (the share of true recipients who do not report receipt in the survey) and false positive rates (the share of true non-recipients who do report receipt in the survey) calculated for each income source and reporting type. Sample consists of individuals or households in the linked CPS sample for reference years 1996 and 2017, dropping non-PIKed units and adjusting survey weights using inverse probability weights. Cells marked “--” are suppressed to comply with Census Bureau disclosure avoidance rules.

Figure 1. Trends in Total Net Survey Error for Recipients Over Time, 1984-2022



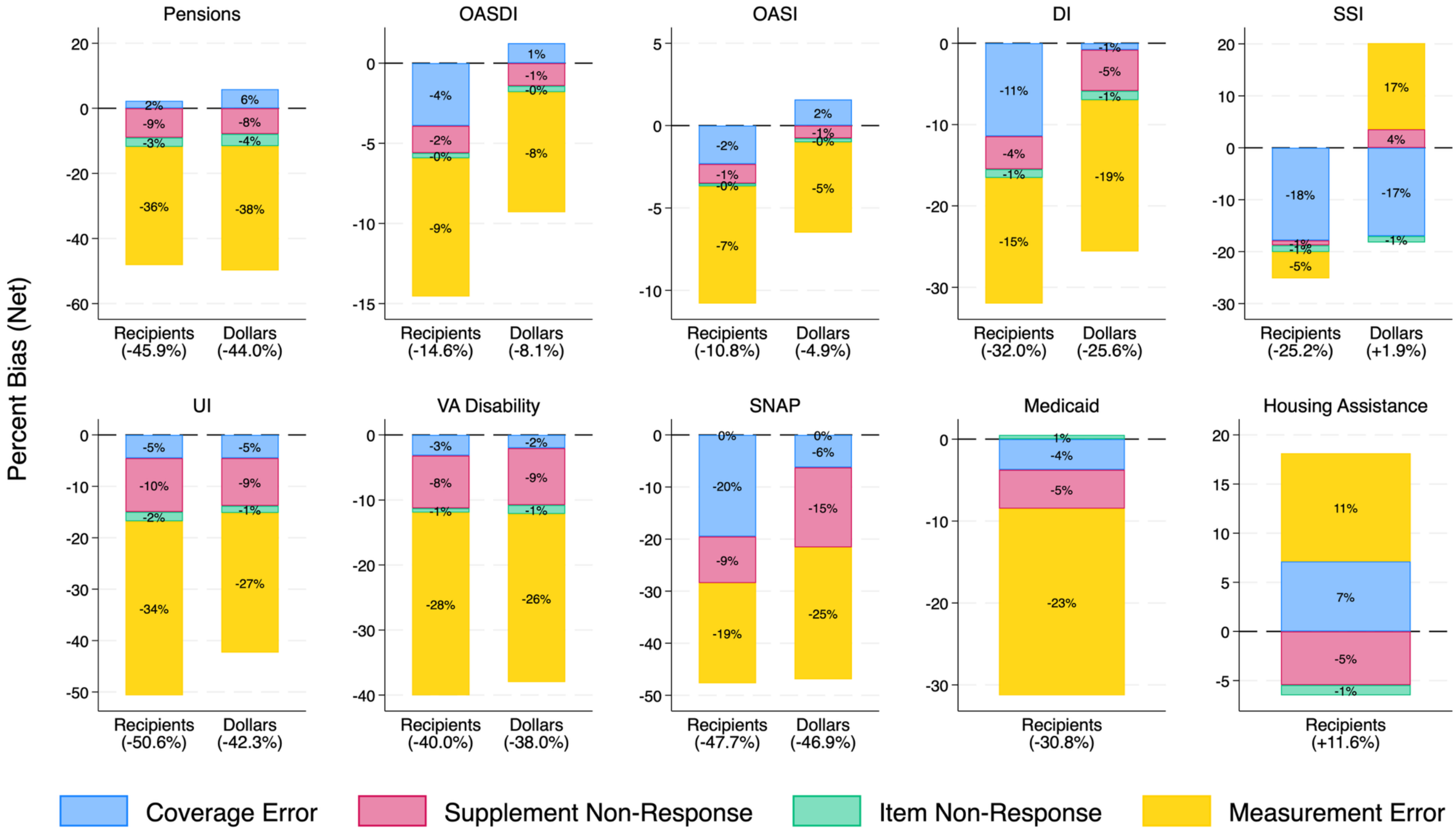
Notes: These figures show total net survey error for annual recipients in the CPS ASEC over time. Net TSE here is calculated as the difference between the weighted survey estimate (based on the public-use CPS ASEC) and the survey target constructed from publicly available administrative aggregates (after making public-data adjustments for intentional coverage differences), as a share of the survey target.

Figure 2. Trends in Total Net Survey Error for Dollars Over Time, 1984-2022



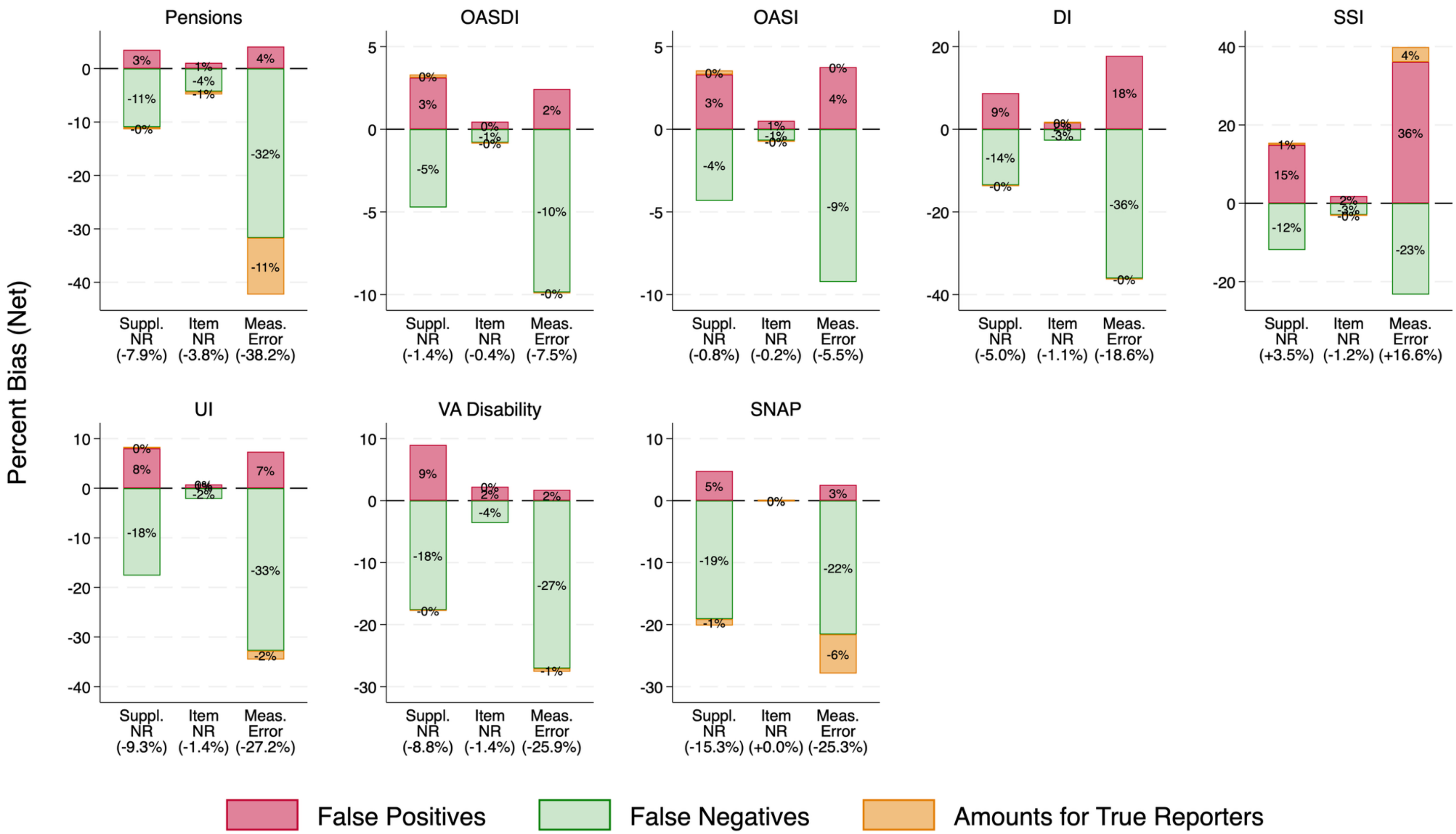
Notes: These figures show total net survey error for annual dollars in the CPS ASEC over time. Net TSE here is calculated as the difference between the weighted survey estimate (based on the public-use CPS ASEC) and the survey target constructed from publicly available administrative aggregates (after making public-data adjustments for intentional coverage differences), as a share of the survey target.

Figure 3. Decomposition of Total Net Survey Error for Recipients and Dollars, 2017



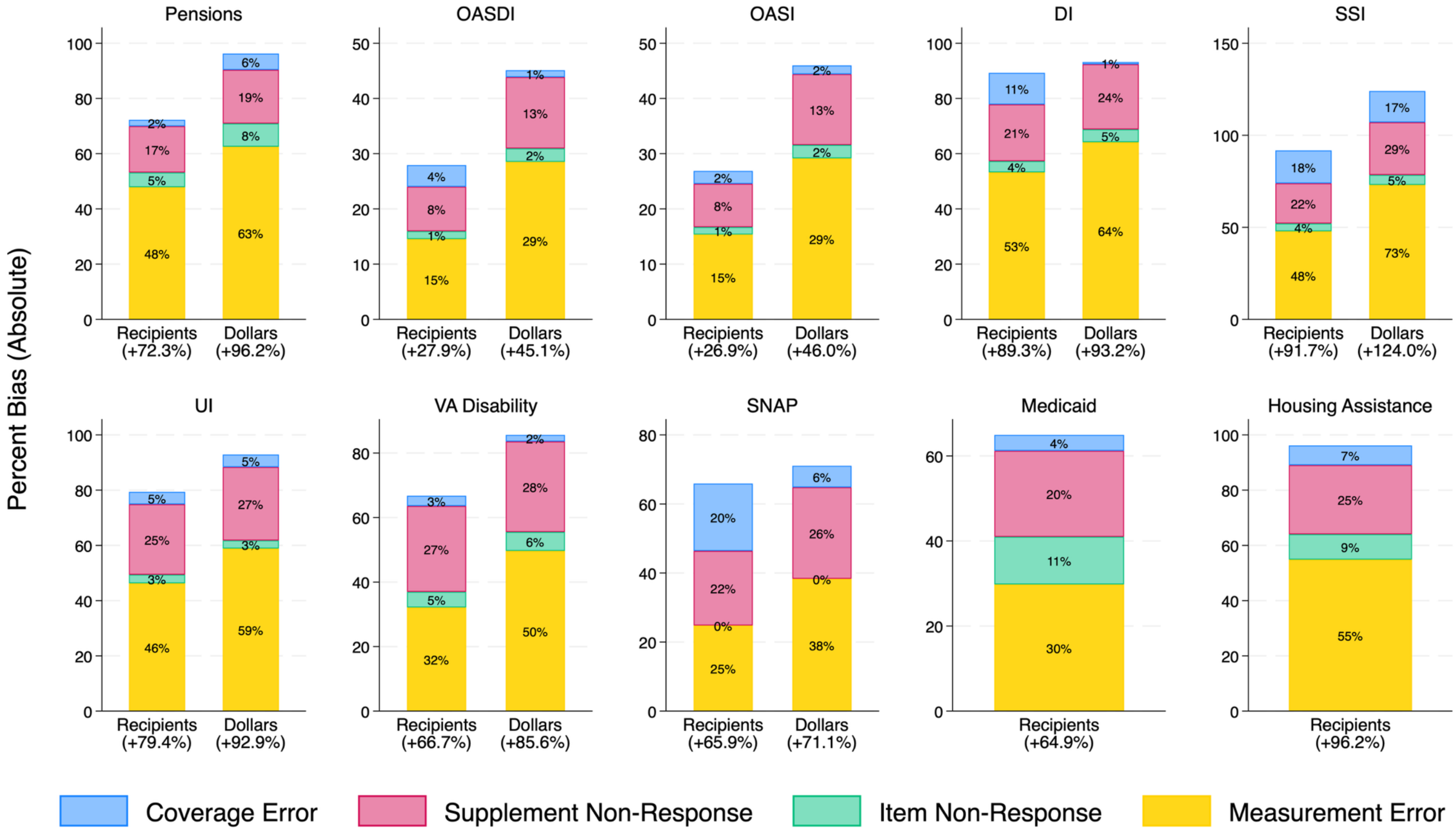
Notes: These figures decompose total net survey error for recipients and dollars into separate components for supplement non-response, item non-response, measurement error among respondents, and coverage error (residual category) in 2017. Errors are reported as a share of the survey target. We examine both recipients and dollars for all sources except Medicaid, for which we only examine recipients. Sample consists of individuals or households in the linked CPS sample for reference years 1996 and 2017, dropping non-PIKed units and adjusting survey weights using inverse probability weights.

Figure 4. Decomposition of Total Net Survey Error for Dollars by Extensive and Intensive Margin Components, 2017



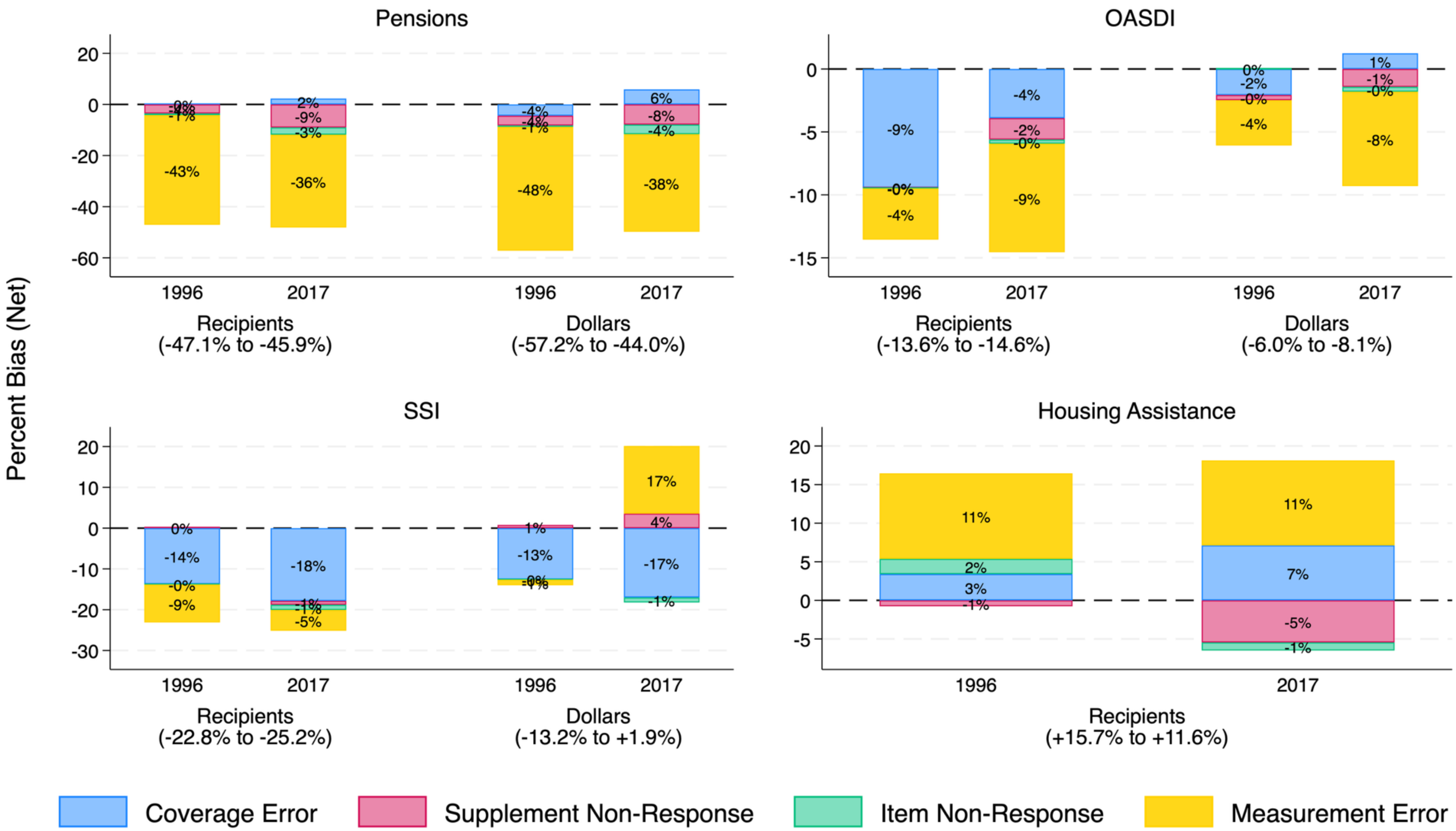
Notes: These figures decompose total net survey error for dollars (separately by supplement non-response, item non-response, and measurement error among respondents) into extensive-margin components (false positives and false negatives) and intensive-margin errors in amounts among true reporting recipients in 2017. Errors are reported as a share of the survey target. Sample consists of individuals or households in the linked CPS sample for reference years 1996 and 2017, dropping non-PIKed units and adjusting survey weights using inverse probability weights

Figure 5. Decomposition of Total Absolute Survey Error for Recipients and Dollars, 2017



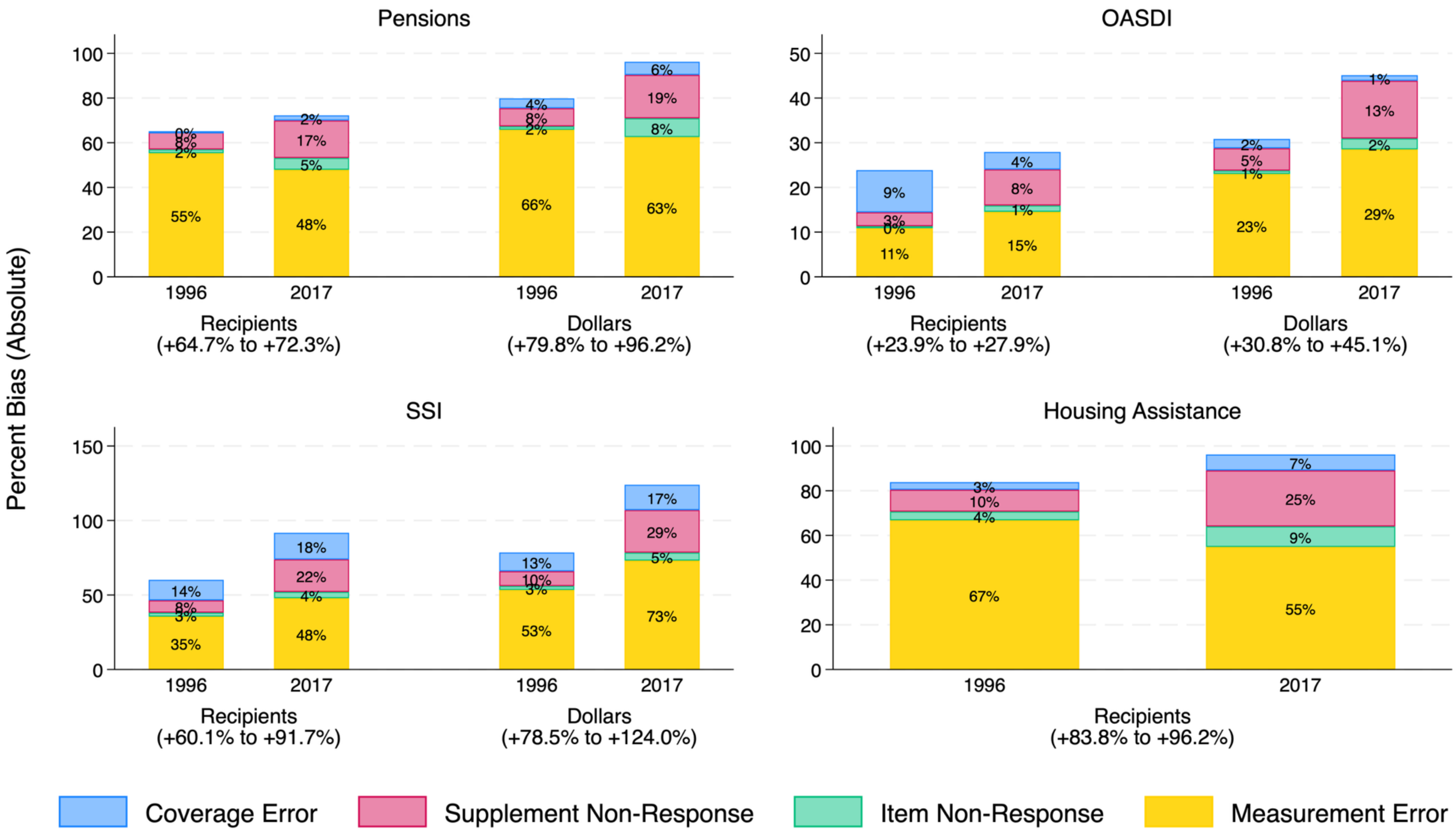
Notes: These figures decompose total absolute survey error for recipients and dollars into separate components for supplement non-response, item non-response, measurement error among respondents, and coverage error (residual category) in 2017. Errors are reported as a share of the survey target. We examine both recipients and dollars for all sources except Medicaid, for which we only examine recipients. Sample consists of individuals or households in the linked CPS sample for reference years 1996 and 2017, dropping non-PIKed units and adjusting survey weights using inverse probability weights.

Figure 6. Decomposition of Total Net Survey Error for Recipients and Dollars, 1996 and 2017



Notes: These figures decompose total net survey error for recipients and dollars into separate components for supplement non-response, item non-response, measurement error among respondents, and coverage error (residual category) in 1996 and 2017. Errors are reported as a share of the survey target. Sample consists of individuals or households in the linked CPS sample for reference years 1996 and 2017, dropping non-PIKed units and adjusting survey weights using inverse probability weights.

Figure 7. Decomposition of Total Absolute Survey Error for Recipients and Dollars, 1996 and 2017



Notes: These figures decompose total absolute survey error for recipients and dollars into separate components for supplement non-response, item non-response, measurement error among respondents, and coverage error (residual category) in 1996 and 2017. Errors are reported as a share of the survey target. Sample consists of individuals or households in the linked CPS sample for reference years 1996 and 2017, dropping non-PIKed units and adjusting survey weights using inverse probability weights.

Appendix Tables and Figures

Table A.1. Recipients and Dollars in Reweighted PIKed CPS Sample and Full CPS Sample, 1996 and 2017

	Pensions	OASDI	OASI	DI	SSI	UI	VA Disability	SNAP	Medicaid	Housing Assistance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					2017					
A. Recipients (millions)										
Recipients in Reweighted PIKed Sample	27.36	52.79	45.23	7.55	6.48	3.68	2.78	1.51	58.17	5.69
Recipients in Full Sample	27.39	52.77	45.22	7.54	6.47	3.67	2.78	1.50	58.41	5.68
Reweighted Sample as a % of Full Sample	99.91%	100.04%	100.02%	100.15%	100.10%	100.24%	99.83%	100.66%	99.59%	100.20%
B. Dollars ($ millions)										
Payments in Reweighted PIKed Sample	651,140	814,084	712,877	101,241	52,662	17,985	43,311	3,486	--	--
Payments in Full Sample	649,787	813,950	712,750	101,201	52,761	18,219	43,294	3,494	--	--
Reweighted Sample as a % of Full Sample	100.21%	100.02%	100.02%	100.04%	99.81%	98.71%	100.04%	99.76%	--	--
					1996					
A. Recipients (millions)										
Recipients in Reweighted PIKed Sample	17.93	37.55	--	--	5.24	--	--	--	--	5.00
Recipients in Full Sample	17.94	37.54	--	--	5.20	--	--	--	--	4.99
Reweighted Sample as a % of Full Sample	99.98%	100.03%	--	--	100.63%	--	--	--	--	100.15%
B. Dollars ($ millions)										
Payments in Reweighted PIKed Sample	651,140	814,084	712,877	101,241	52,662	17,985	43,311	3,486	--	--
Payments in Full Sample	649,787	813,950	712,750	101,201	52,761	18,219	43,294	3,494	--	--
Reweighted Sample as a % of Full Sample	100.21%	100.02%	100.02%	100.04%	99.81%	98.71%	100.04%	99.76%	--	--

Notes: Cells marked “--” are not calculated. SNAP recipients and dollars are for five states only. See Section A.2 of the Data Appendix for details of the CPS ASEC variables used for each income source.

Table A.2. Total Recipients and Dollars that Should Be Covered by the Survey Sample, 1996 and 2017

	Pensions		OASDI		SSI		Housing Assistance	
	1996	2017	1996	2017	1996	2017	1996	2017
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Recipients (millions)								
Recipients in Admin Microdata	36.10	53.27	46.92	66.09	7.10	9.00	4.76	5.40
Subtraction of Recipients Not Covered by Household Samples (%)								
Decedents	2.84%	2.81%	2.54%	1.95%	1.14%	0.78%	1.10%	1.83%
Estimated Recipients in GQs	2.05%	0.95%	3.35%	3.01%	3.35%	3.01%	5.30%	3.61%
Recipients in Overseas & Territories	--	--	1.54%	1.57%	0.01%	0.01%	--	--
Subtraction of Recipients Not Covered by Linked Data (%)								
Unlinkable Admin Records	1.24%	1.33%	--	--	--	--	2.75%	0.18%
Total Subtraction	6.13%	5.09%	7.43%	6.52%	4.50%	3.80%	9.15%	5.62%
Total Recipients Covered	33.89	50.56	43.44	61.78	6.78	8.66	4.32	5.10
B. Dollars ($ millions)								
Payments in Admin Microdata	495,700	1,203,000	347,119	941,554	28,250	54,509	495,7	1,203,000
Subtraction of Payments Not Covered by Household Samples (%)								
Payments to Decedents	1.62%	1.95%	2.76%	2.04%	0.96%	0.71%	1.62%	1.95%
Estimated Payments to GQs	1.61%	0.68%	4.02%	2.35%	7.80%	4.46%	1.61%	0.68%
Payments to Overseas & Territories	--	--	1.54%	1.57%	0.01%	0.01%	--	--
Subtraction of Payments Not Covered by Linked Data (%)								
Unlinkable Admin Records	0.85%	0.80%	--	--	--	--	0.85%	0.80%
Total Subtraction	4.08%	3.43%	8.31%	5.96%	8.77%	5.18%	4.08%	3.43%
Total Payments Covered	475,462	1,161,754	318,263	885,459	25,773	51,683	475,462	1,161,754

Notes: Cells marked "--" are not calculated. All amounts are in millions. For Pensions, the "Recipients in Admin Microdata" row is de-duplicated by subtracting PIKed duplicates from the universe administrative count. For OASDI and SSI, universe administrative data are unavailable because the PHUS and SSR contain only PIKed individuals matched to the CPS; estimates for these sources therefore rely on publicly available administrative aggregates. See Section A.2 of the Data Appendix for details on these data sources and adjustments.

Table A.3. Misreporting Rates Among Respondents from Prior Studies and This Paper

Study	Program	Survey	Sample	False Negative Rates	False Positive Rates
Meyer, Mittag, George (2022)	SNAP	ACS	IL + MD	38%	0.5%
Survey Years: 2001-2005		CPS	IL + MD	45%	0.6%
		SIPP	IL + MD	22%	1.2%
Celhay, Meyer, Mittag (2021)	SNAP	ACS	NY	25%	1.1%
Survey Years: 2007-2012		CPS	NY	37%	1.2%
		SIPP	NY	18%	1.3%
	TANF + GA	ACS	NY	59%	1.3%
		CPS	NY	59%	0.3%
		SIPP	NY	46%	0.4%
This Paper	Pensions	CPS	U.S.	54%	1.8%
Survey Year: 2018	OASDI	CPS	U.S.	15%	1.2%
	OASI	CPS	U.S.	14%	1.3%
	DI	CPS	U.S.	49%	1.1%
	SSI	CPS	U.S.	41%	0.9%
	UI	CPS	U.S.	54%	0.2%
	VA Disab.	CPS	U.S.	40%	0.0%
	SNAP	CPS	5 States	36%	0.8%
	Medicaid	CPS	U.S.	40%	1.7%
	Housing	CPS	U.S.	27%	1.7%

Notes: This table contains false negative and false positive rates for various programs, surveys, samples, and years calculated based on linked survey-administrative data from prior studies as well as this paper.

Table A.4. Sample Shares for Respondents and Non-Respondents by Income Source, 1996 and 2017

	Pensions	OASDI	OASI	DI	SSI	UI	VA Disability	SNAP	Medicaid	Housing Assistance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. 2017										
Supplement Non-Respondents	19.0	19.0	19.0	19.0	19.0	19.0	19.0	21.3	18.4	18.7
Item Non-Respondents	3.9	2.5	2.5	2.5	1.8	1.8	2.6	0.7	9.1	1.4
Respondents	77.1	78.5	78.5	78.5	79.2	79.2	78.4	78.0	72.5	79.9
B. 1996										
Supplement Non-Respondents	7.9	7.9			7.9					8.2
Item Non-Respondents	1.4	0.8			0.8					1.1
Respondents	90.8	91.3			91.3					90.7

Notes: This table shows the share of the sample accounted for by each reporting type (supplement non-respondents, item non-respondents, and respondents) for each income source in 1996 and 2017. The sample consists of individuals or households in the linked CPS sample, dropping non-PIKed units and adjusting survey weights using inverse probability weights.

Table A.5. Per Unit Measures of Survey Error for Non-Respondents (as Multiple of that for Respondents), 2017

	Pensions	OASDI	OASI	DI	SSI	UI	VA Disability	SNAP	Medicaid	Housing Assistance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Recipients, Net Error										
Supplement Non-Respondents	1.01	0.82	0.70	1.10	0.74	1.28	1.19	1.70	0.81	-2.14
Item Non-Respondents	1.52	1.10	0.88	2.24	11.22	2.34	0.75	--	-0.17	-5.19
B. Dollars, Net Error										
Supplement Non-Respondents	0.84	0.77	0.6	1.11	0.88	1.43	1.4	2.21	--	--
Item Non-Respondents	1.97	1.67	1.14	1.86	-3.18	2.26	1.63	--	--	--
C. Recipients, Net Error										
Supplement Non-Respondents	1.42	2.31	2.09	1.59	1.90	2.30	3.42	3.19	2.67	1.95
Item Non-Respondents	2.19	3.03	2.65	2.36	3.86	2.95	4.51	--	2.99	9.56
D. Dollars, Absolute Error										
Supplement Non-Respondents	1.26	1.87	1.81	1.52	1.63	1.88	2.33	2.53	--	--
Item Non-Respondents	2.66	2.64	2.58	2.25	3.25	2.17	3.58	--	--	--

Notes: Cells marked “--” are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as the per-unit error for the relevant non-respondent group as a multiple/ratio of the per-unit error for respondents (measurement error).

Table A.6. Total Net Survey Error for Recipients and Its Components (as a Conditional % of TSE), 2017

	Pensions	OASDI	OASI	DI	SSI	UI	VA Disability	SNAP	Medicaid	Housing Assistance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Coverage Error	-5.0	26.9	21.1	35.8	71.0	8.9	7.9	48.9	12.2	61.2
B. Supplement Non-Response Error										
False Positives	-8.5	-21.9	-31.0	-25.7	-41.5	-14.8	-23.1	-11.5	-25.1	84.2
False Negatives	28.1	33.5	42.0	38.4	45.2	35.5	43.3	27.6	40.6	-131.0
Total Net Error Due to Suppl. Non-Response	19.6	11.7	11.0	12.7	3.7	20.6	20.2	16.2	15.4	-46.9
C. Item Non-Response Error										
False Positives	-2.7	-3.7	-5.4	-4.7	-5.8	-1.3	-5.2	--	-19.0	35.0
False Negatives	8.9	6.0	7.0	8.0	10.9	4.9	6.9	--	17.3	-43.9
Total Net Error Due to Item Non-Response	6.2	2.3	1.6	3.3	5.1	3.6	1.7	--	-1.7	-9.0
D. Measurement Error										
False Positives	-12.6	-20.3	-38.4	-59.0	-85.0	-12.3	-5.1	-5.0	-11.4	283.1
False Negatives	91.8	79.5	104.7	107.2	105.2	79.1	75.2	39.9	85.5	-188.5
Total Net Error Due to Measurement Error	79.2	59.2	66.3	48.2	20.2	66.8	70.2	34.9	74.1	94.6
Total Net Error	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Notes: Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as shares of Total Net Survey Error. For each reporting group (supplement non-respondents, item non-respondents, and respondents), we report false positives, false negatives, and their sum, defined as net error due to mis-imputed or mis-reported receipt status. Each component A-D sums to total net error of 100%.

Table A.7. Total Absolute Survey Error for Recipients and Its Components (as a Conditional % of TSE), 2017

	Pensions	OASDI	OASI	DI	SSI	UI	VA Disability	SNAP	Medicaid	Housing Assistance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Coverage Error	3.2	14.0	8.5	12.8	19.5	5.7	4.7	36.8	5.8	7.4
B. Supplement Non-Response Error										
False Positives	5.4	11.4	12.4	9.2	11.4	9.5	13.9	8.6	11.9	10.2
False Negatives	-17.8	-17.5	-16.8	-13.8	-12.4	-22.6	-26.0	-20.8	-19.2	-15.8
Total Absolute Error Due to Suppl. Non-Response	23.2	28.9	29.2	23.0	23.8	32.1	39.9	29.4	31.1	26.0
C. Item Non-Response Error										
False Positives	1.7	2.0	2.2	1.7	1.6	0.8	3.1	--	9.0	4.2
False Negatives	-5.6	-3.1	-2.8	-2.9	-3.0	-3.1	-4.1	--	-8.2	-5.3
Total Absolute Error Due to Item Non-Response	7.3	5.1	5.0	4.5	4.6	3.9	7.2	--	17.2	9.5
D. Measurement Error										
False Positives	8.0	10.6	15.4	21.2	23.3	7.8	3.0	3.8	5.4	34.2
False Negatives	-58.3	-41.4	-41.9	-38.5	-28.9	-50.5	-45.1	-30.0	-40.5	-22.8
Total Absolute Error Due to Measurement Error	66.3	52.0	57.3	59.6	52.2	58.3	48.2	33.8	45.9	57.0
Total Absolute Error	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Notes: Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as shares of Total Absolute Survey Error. For each reporting group (supplement non-respondents, item non-respondents, and respondents), we report false positives, false negatives, and their sum, defined as absolute error due to mis-imputed or mis-reported receipt status. Each component A-D sums to total absolute error of 100%.

Table A.8. Total Net Survey Error for Dollar Amounts and Its Components (as a Conditional % of TSE), 2017

	Pensions	OASDI	OASI	DI	SSI	UI	VA Disability	SNAP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Coverage Error	-13.4	-15.4	-34.0	3.5	-898.1	10.7	5.3	13.3
B. Supplement Non-Response Error								
False Positives	-8.0	-38.7	-68.2	-34.0	785.8	-19.0	-23.7	-10.2
False Negatives	25.0	58.7	89.2	52.7	-628.8	41.6	46.4	40.7
Net Error due to Mis-Imputed Receipt Status	17.0	20.0	21.1	18.7	157.0	22.6	22.8	30.5
Net Error in Imputed Amounts	0.9	-2.4	-5.2	1.0	30.5	-0.7	0.3	2.2
Total Net Error Due to Suppl. Non-Response	17.9	17.6	15.9	19.7	187.4	21.9	23.0	32.7
C. Item Non-Response Error								
False Positives	-2.5	-5.7	-10.5	-6.2	97.9	-1.8	-5.9	--
False Negatives	9.8	10.1	14.1	10.7	-157.3	5.1	9.5	--
Net Error due to Mis-Imputed Receipt Status	7.3	4.3	3.6	4.5	-59.4	3.3	3.6	--
Net Error in Imputed Amounts	1.2	0.4	1.5	0.0	-5.6	--	--	--
Total Net Error Due to Item Non-Response	8.6	4.8	5.1	4.5	-65.0	3.3	3.6	--
D. Measurement Error								
False Positives	-9.4	-30.1	-77.4	-69.2	1904.7	-17.4	-4.6	-5.4
False Negatives	72.1	122.5	190.5	140.6	-1230.3	77.4	71.2	46.0
Net Error due to Mis-Reported Receipt Status	62.7	92.3	113.1	71.4	674.4	60.0	66.6	40.6
Net Error in Reported Amounts	24.2	0.7	0.0	1.0	201.3	4.2	1.4	13.4
Total Net Error Due to Measurement Error	86.9	93.0	113.1	72.4	875.7	64.2	68.0	54.0
Total Net Error	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Notes: Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as shares of Total Net Survey Error. For each reporting group (supplement non-respondents, item non-respondents, and respondents), we report false positives, false negatives, and their sum, defined as net error due to mis-imputed or mis-reported receipt status. Error in imputed or reported amounts reflects dollar discrepancies for correct reporting recipients. Each component A-D sums to total net error of 100%.

Table A.9. Total Absolute Survey Error for Dollar Amounts and Its Components (as a Conditional % of TSE), 2017

	Pensions	OASDI	OASI	DI	SSI	UI	VA Disability	SNAP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Coverage Error	6.1	2.8	3.6	1.0	13.7	4.9	2.4	8.8
B. Supplement Non-Response Error								
False Positives	3.6	6.9	7.2	9.4	12.0	8.7	10.5	6.7
False Negatives	-11.4	-10.5	-9.4	-14.5	-9.6	-19.0	-20.6	-26.9
Absolute Error due to Mis-Imputed Receipt Status	15.0	17.4	16.6	23.8	21.6	27.6	31.1	33.6
Absolute Error in Imputed Amounts	5.1	11.2	11.2	1.4	1.4	1.0	1.7	3.6
Total Absolute Error Due to Suppl. Non-Response	20.2	28.6	27.8	25.3	23.0	28.6	32.8	37.3
C. Item Non-Response Error								
False Positives	1.1	1.0	1.1	1.7	1.5	0.8	2.6	--
False Negatives	-4.5	-1.8	-1.5	-2.9	-2.4	-2.3	-4.2	--
Absolute Error due to Mis-Imputed Receipt Status	5.6	2.8	2.6	4.7	3.9	3.1	6.8	--
Absolute Error in Imputed Amounts	3.1	2.6	2.7	0.3	0.4	--	--	--
Total Absolute Error Due to Item Non-Response	8.7	5.4	5.3	4.9	4.3	3.1	6.8	--
D. Measurement Error								
False Positives	4.3	5.4	8.2	19.0	29.1	7.9	2.0	3.6
False Negatives	-32.9	-21.9	-20.1	-38.7	-18.8	-35.3	-31.6	-30.4
Absolute Error due to Mis-Reported Receipt Status	37.2	27.3	28.2	57.7	47.9	43.2	33.6	33.9
Absolute Error in Reported Amounts	27.8	36.0	35.1	11.1	11.0	20.2	24.4	20.0
Total Absolute Error Due to Measurement Error	65.0	63.2	63.3	68.8	58.9	63.4	58.0	53.9
Total Absolute Error	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Notes: Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as shares of Total Absolute Survey Error. For each reporting group (supplement non-respondents, item non-respondents, and respondents), we report false positives, false negatives, and their sum, defined as absolute error due to mis-imputed or mis-reported receipt status. Error in imputed or reported amounts reflects dollar discrepancies for correct reporting recipients. Each component A-D sums to total absolute error of 100%.

Table A.10. Per Unit Measures of Survey Error Subcomponents for Non-Respondents (as Multiple of that for Respondents), 2017

	Pensions	OASDI	OASI	DI	SSI	UI	VA Disability	SNAP	Medicaid	Housing Assistance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Recipients: Supplement Non-Respondents										
False Positives	2.73	4.41	3.33	1.79	2.03	5.04	19.19	8.31	8.67	1.27
False Negatives	1.24	1.75	1.65	1.48	1.79	1.87	2.37	2.54	1.87	2.97
B. Recipients: Item Non-Respondents										
False Positives	4.09	5.23	4.60	2.49	3.08	4.26	31.66	--	13.20	7.11
False Negatives	1.93	2.44	2.22	2.38	4.48	2.74	2.81	--	1.61	13.29
C. Dollars: Supplement Non-Respondents										
False Positives	3.46	5.34	3.68	2.02	1.72	4.51	21.85	7.03	--	--
False Negatives	1.41	1.96	1.93	1.55	2.13	2.24	2.68	3.24	--	--
Errors in Amounts (Net)	0.15	-8.26	--	4.13	0.66	-0.69	0.83	0.58	--	--
Errors in Amounts (Absolute)	0.74	1.30	1.33	0.52	0.55	0.20	0.28	0.67	--	--
D. Dollars: Item Non-Respondents										
False Positives	5.30	6.54	4.24	2.82	2.32	4.76	39.02	--	--	--
False Negatives	2.68	2.54	2.39	2.35	5.67	2.95	4.01	--	--	--
Errors in Amounts (Net)	0.93	0.00	--	0.00	-1.16	--	--	--	--	--
Errors in Amounts (Absolute)	2.22	2.33	2.33	0.91	1.93	--	--	--	--	--

Notes: Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as the per-unit error for the relevant non-respondent group as a multiple/ratio of the per-unit error for respondents (measurement error).

Table A.11. Per Unit Measures of Survey Error for Non-Respondents and Respondents, Percent Change from 1996 and 2017

	Pensions	OASDI	SSI	Housing Assistance
	(1)	(2)	(3)	(4)
A. Recipients, Net Error				
Supplement Non-Respondents	6.9	--	CS	201.5
Item Non-Respondents	43.6	--	477.8	CS
Respondents	-0.3	144.0	-36.8	12.5
B. Dollars, Net Error				
Supplement Non-Respondents	-13.6	45.5	81.9	--
Item Non-Respondents	172.8	--	433.3	--
Respondents	7.2	-142.3	1466.9	--
C. Recipients, Absolute Error				
Supplement Non-Respondents	-9.3	8.6	9.2	11.9
Item Non-Respondents	11.9	-10.4	-30.9	85.3
Respondents	2.0	54.7	55.5	-6.7
D. Dollars, Absolute Error				
Supplement Non-Respondents	-0.4	7.3	20.1	--
Item Non-Respondents	101	9.7	-7.7	--
Respondents	11.9	44.1	57.8	--

Notes: Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as the percent change in per-unit error from 1996 to 2017 (using 1996 as the baseline) for the relevant respondent group, where "negative" means that the error is getting smaller and "positive" means that the error is getting bigger. We write "CS" when the error changes sign between 1996 and 2017.

Table A.12. Total Net Survey Error for Recipients and Its Components (as a Conditional % of TSE), 1996 and 2017

	Pensions		OASDI		SSI		Housing Assistance	
	1996	2017	1996	2017	1996	2017	1996	2017
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Coverage Error	0.0	-5.0	69.5	26.9	60.2	71.0	21.7	61.2
B. Supplement Non-Response Error								
False Positives	-4.4	-8.5	-11.4	-21.9	-18.9	-41.5	28.7	84.2
False Negatives	11.9	28.1	11.5	33.5	17.4	45.2	-33.5	-131.0
Total Net Error Due to Suppl. Non-Response	7.5	19.6	0.1	11.7	-1.5	3.7	-4.9	-46.9
C. Item Non-Response Error								
False Positives	-1.1	-2.7	-1.5	-3.7	-5.6	-5.8	18.7	35.0
False Negatives	2.5	8.9	1.8	6.0	6.1	10.9	-6.1	-43.9
Total Net Error Due to Item Non-Response	1.4	6.2	0.3	2.3	0.5	5.1	12.6	-9.0
D. Measurement Error								
False Positives	-13.2	-12.6	-25.1	-20.3	-57.5	-85.0	248.1	283.1
False Negatives	104.3	91.8	55.1	79.5	98.3	105.2	-177.5	-188.5
Total Net Error Due to Measurement Error	91.1	79.2	30.0	59.2	40.8	20.2	70.6	94.6
Total Net Error	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Notes: Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as shares of Total Net Survey Error. For each reporting group (supplement non-respondents, item non-respondents, and respondents), we report false positives, false negatives, and their sum, defined as net error due to mis-imputed or mis-reported receipt status. Each component A-D sums to total net error of 100%.

Table A.13. Total Absolute Survey Error for Recipients and Its Components (as a Conditional % of TSE), 1996 and 2017

	Pensions		OASDI		SSI		Housing Assistance	
	1996	2017	1996	2017	1996	2017	1996	2017
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Coverage Error	0.0	3.2	39.5	14.0	22.8	19.5	4.1	7.4
B. Supplement Non-Response Error								
False Positives	3.2	5.4	6.5	11.4	7.2	11.4	5.4	10.2
False Negatives	-8.6	-17.8	-6.5	-17.5	-6.6	-12.4	-6.3	-15.8
Total Absolute Error Due to Suppl. Non-Response	11.8	23.2	13.0	28.9	13.8	23.8	11.6	26.0
C. Item Non-Response Error								
False Positives	0.8	1.7	0.9	2.0	2.1	1.6	3.5	4.2
False Negatives	-1.8	-5.6	-1.0	-3.1	-2.3	-3.0	-1.1	-5.3
Total Absolute Error Due to Item Non-Response	2.6	7.3	1.9	5.1	4.4	4.6	4.6	9.5
D. Measurement Error								
False Positives	9.6	8.0	14.3	10.6	21.8	23.3	46.4	34.2
False Negatives	-75.9	-58.3	-31.3	-41.4	-37.2	-28.9	-33.2	-22.8
Total Absolute Error Due to Measurement Error	85.5	66.3	45.6	52.0	59.0	52.2	79.7	57.0
Total Absolute Error	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Notes: Cells marked “--” are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as shares of Total Absolute Survey Error. For each reporting group (supplement non-respondents, item non-respondents, and respondents), we report false positives, false negatives, and their sum, defined as absolute error due to mis-imputed or mis-reported receipt status. Each component A-D sums to total absolute error of 100%.

Table A.14. Total Net Survey Error for Dollar Amounts and Its Components (as a Conditional % of TSE), 1996 and 2017

	Pensions		OASDI		SSI	
	1996	2017	1996	2017	1996	2017
	(1)	(2)	(3)	(4)	(5)	(6)
A. Coverage Error	7.7	-13.4	34.3	-15.4	95.3	-898.1
B. Supplement Non-Response Error						
False Positives	-2.8	-8.0	-21.5	-38.7	-36.1	785.8
False Negatives	8.5	25.0	23.9	58.7	29.7	-628.8
Net Error due to Mis-Imputed Receipt Status	5.8	17.0	2.4	20.0	-6.4	157.0
Net Error in Imputed Amounts	0.9	0.9	4.3	-2.4	0.4	30.5
Total Net Error Due to Suppl. Non-Response	6.7	17.9	6.7	17.6	-6.0	187.4
C. Item Non-Response Error						
False Positives	-0.5	-2.5	-3.8	-5.7	-9.7	97.9
False Negatives	1.3	9.8	3.1	10.1	10.1	-157.3
Net Error due to Mis-Imputed Receipt Status	0.9	7.3	-0.7	4.3	0.4	-59.4
Net Error in Imputed Amounts	0.0	1.2	0.1	0.4	--	-5.6
Total Net Error Due to Item Non-Response	0.9	8.6	-0.7	4.8	0.4	-65.0
D. Measurement Error						
False Positives	-7.6	-9.4	-48.6	-30.1	-105.0	1904.7
False Negatives	68.7	72.1	99.9	122.5	159.0	-1230.3
Net Error due to Mis-Reported Receipt Status	61.1	62.7	51.3	92.3	54.0	674.4
Net Error in Reported Amounts	23.6	24.2	8.3	0.7	-43.8	201.3
Total Net Error Due to Measurement Error	84.7	86.9	59.7	93.0	10.3	875.7
Total Net Error	100.0	100.0	100.0	100.0	100.0	100.0

Notes: Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as shares of Total Net Survey Error. For each reporting group (supplement non-respondents, item non-respondents, and respondents), we report false positives, false negatives, and their sum, defined as net error due to mis-imputed or mis-reported receipt status. Error in imputed or reported amounts reflects dollar discrepancies for correct reporting recipients. Each component A-D sums to total net error of 100%.

Table A.15. Total Absolute Survey Error for Dollar Amounts and Its Components (as a Conditional % of TSE), 1996 and 2017

	Pensions		OASDI		SSI	
	1996	2017	1996	2017	1996	2017
	(1)	(2)	(3)	(4)	(5)	(6)
A. Coverage Error	5.5	6.1	6.7	2.8	16.0	13.7
B. Supplement Non-Response Error						
False Positives	2.0	3.6	4.2	6.9	6.1	12.0
False Negatives	-6.1	-11.4	-4.7	-10.5	-5.0	-9.6
Absolute Error due to Mis-Imputed Receipt Status	8.1	15.0	8.9	17.4	11.1	21.6
Absolute Error in Imputed Amounts	2.0	5.1	7.3	11.2	1.5	1.4
Total Absolute Error Due to Suppl. Non-Response	10.1	20.2	16.2	28.6	12.6	23.0
C. Item Non-Response Error						
False Positives	0.3	1.1	0.7	1.0	1.6	1.5
False Negatives	-1.0	-4.5	-0.6	-1.8	-1.7	-2.4
Absolute Error due to Mis-Imputed Receipt Status	1.3	5.6	1.4	2.8	3.3	3.9
Absolute Error in Imputed Amounts	0.6	3.1	1.1	2.6	--	0.4
Total Absolute Error Due to Item Non-Response	1.9	8.7	2.4	5.4	3.3	4.3
D. Measurement Error						
False Positives	5.5	4.3	9.5	5.4	17.7	29.1
False Negatives	-49.2	-32.9	-19.6	-21.9	-26.8	-18.8
Absolute Error due to Mis-Reported Receipt Status	54.7	37.2	29.1	27.3	44.4	47.9
Absolute Error in Reported Amounts	27.8	27.8	45.6	36.0	23.6	11.0
Total Absolute Error Due to Measurement Error	82.4	65.0	74.7	63.2	68.0	58.9
Total Absolute Error	100.0	100.0	100.0	100.0	100.0	100.0

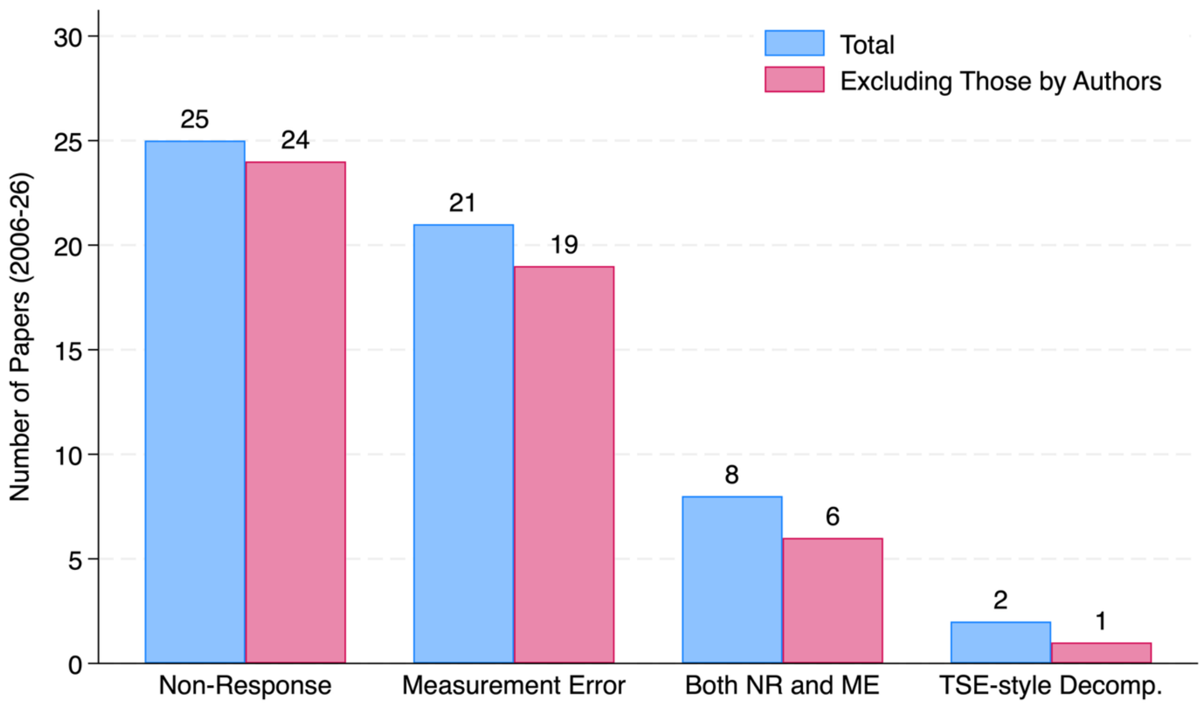
Notes: Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as shares of Total Absolute Survey Error. For each reporting group (supplement non-respondents, item non-respondents, and respondents), we report false positives, false negatives, and their sum, defined as absolute error due to mis-imputed or mis-reported receipt status. Error in imputed or reported amounts reflects dollar discrepancies for correct reporting recipients. Each component A-D sums to total absolute error of 100%.

Table A.16. Per Unit Measures of Survey Error Subcomponents for Non-Respondents and Respondents, Percent Change from 1996 and 2017

	Pensions	OASDI	SSI	Housing Assistance
	(1)	(2)	(3)	(4)
A. Recipients: Supplement Non-Respondents				
False Positives	-22.8	-11.3	0.6	-4.5
False Negatives	-4.2	27.3	18.5	25.8
B. Recipients: Item Non-Respondents				
False Positives	-13.8	-20.0	-48.7	11.1
False Negatives	22.6	-4.0	-14.3	300.7
C. Recipients: Respondents				
False Positives	10.2	2.6	88.3	-4.0
False Negatives	1.0	79.9	36.4	-10.6
D. Dollars: Supplement Non-Respondents				
False Positives	-9.0	-0.9	29.1	--
False Negatives	-6.7	39.6	26.9	--
Errors in Amounts (Net)	-66.7	CS	CS	--
Errors in Amounts (Absolute)	27.3	-3.6	-37.6	--
E. Dollars: Item Non-Respondents				
False Positives	31.6	-20.0	-35.0	--
False Negatives	92.9	28.0	2.6	--
Errors in Amounts (Net)	CS	--	--	--
Errors in Amounts (Absolute)	115.4	28.0	--	--
F. Dollars: Respondents				
False Positives	9.7	-3.7	199.4	--
False Negatives	-5.0	91.9	27.9	--
Errors in Amounts (Net)	-7.5	-76.7	-24.5	--
Errors in Amounts (Absolute)	41.6	33.6	-14.6	--

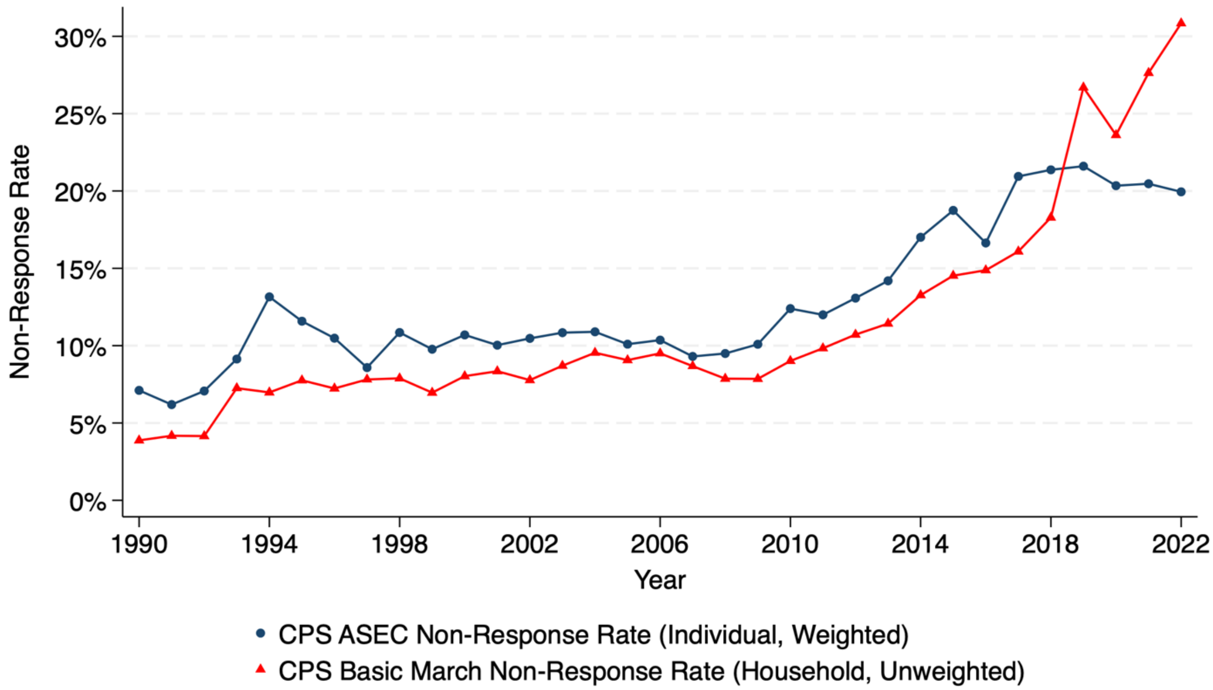
Notes: Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as the percent change in per-unit error from 1996 to 2017 (using 1996 as the baseline) for the relevant respondent group, where "negative" means that the error is getting smaller and "positive" means that the error is getting bigger. We write "CS" when the error changes sign between 1996 and 2017.

Figure A.1. Distribution of Empirical Papers that Quantify Survey Error, 2006-2026



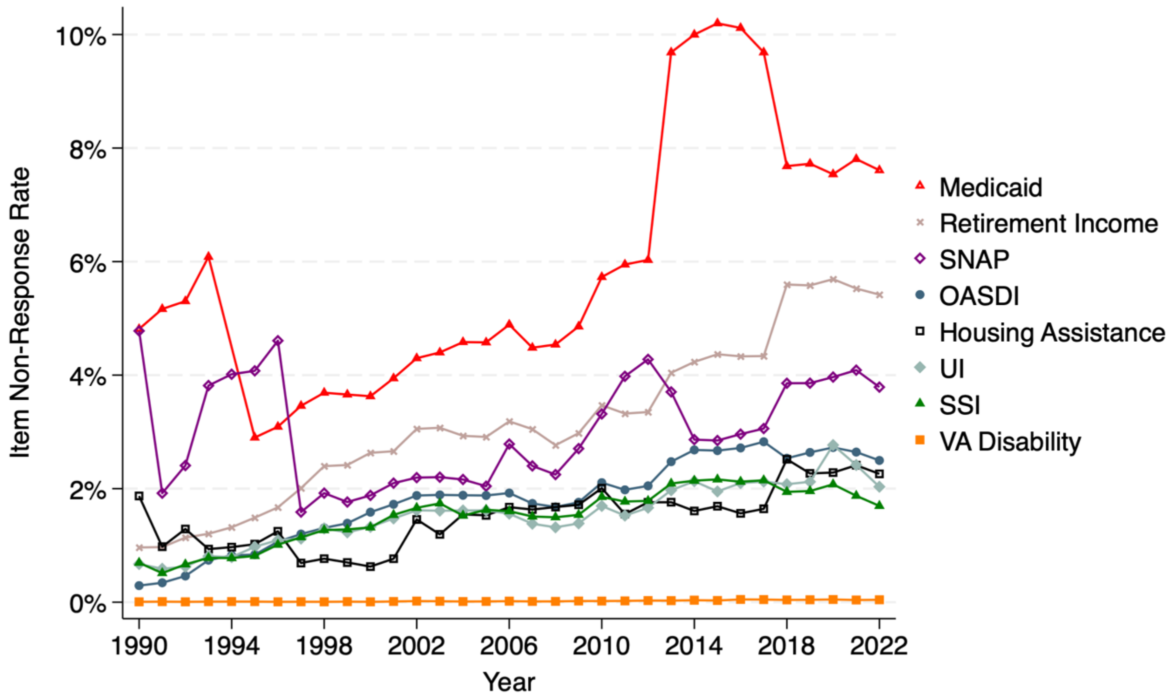
Notes: The figure classifies empirical studies published or circulated over between 2006 and 2026 that quantify one or more sources of error in household or population surveys. We searched leading economics journals (*AER*, *QJE*, *JPE*, *ReStud*, *Econometrica*, *ReStat*, *JEP*, *AEJ: Applied Economics*, *AEJ: Economic Policy*, *JPubE*, *JHR*, *JOLE*, and the *Journal of Econometrics*), statistics and survey-methodology journals (*JASA*, *Public Opinion Quarterly*, *Journal of Official Statistics*, *Journal of Survey Statistics and Methodology*, *Survey Methodology*, and *Sociological Methods & Research*), and major working paper series (the U.S. Census Bureau CES, SEHSD, CARRA, and ACS series; NBER; IZA; and RePEc/IDEAS). We include studies that quantify survey error using validation evidence (administrative records, benchmark aggregates, reinterviews, randomized incentives, survey-mode experiments, or comparable designs) and exclude papers that mention survey error only as a caveat, are purely theoretical, or discuss the total survey error framework conceptually without empirical quantification. Each study is assigned to one of four categories: those that quantify non-response error, those that quantify measurement error, those that quantify both, and those that implement a full total survey error decomposition. Studies that quantify both non-response and measurement error appear only in the "Both" category and are not double-counted in the single-source categories. The right bar of each pair excludes studies authored by the present authors.

Figure A.2. Trends in Unit Non-Response Rates in Basic Monthly CPS and ASEC, 1990-2022



Notes: This figure shows unit non-response rates calculated in the Basic Monthly CPS (red circles) at the household level (unweighted), as well as supplement non-response rates calculated in the CPS ASEC (blue circles) at the individual-level (incorporating ASEC survey weights). These rates are calculated based on publicly available CPS data.

Figure A.3. Trends in Item Non-Response Rates in CPS ASEC, 1990-2022



Notes: This figure shows item non-response rates calculated in the CPS ASEC, where the unit of analysis corresponds to the level at which an income source is reported (individual for all income sources except for SNAP and housing assistance, which are at the household level). All rates are weighted using survey weights. These rates are calculated based on publicly available CPS data.

Data Appendix

A.1. Survey Questions and Variables from CPS ASEC

This appendix section documents the CPS ASEC questions and variable names used to identify recipients and dollar amounts for each income source.[37] Respondents are classified as recipients of a given income source if they answer "yes" to any relevant receipt question. Dollar amounts are summed across all applicable variables.

Pensions

We define pension receipt as receipt of any retirement, survivor, or disability pensions, excluding OASDI. Variable names change beginning in the 2019 CPS ASEC. The table below summarizes variables, years, and questionnaire text. A respondent is identified as a pension recipient if they answer "yes" to any receipt question. Total pension income equals the sum across all pension-related amount variables.

CPS ASEC Years	Variable Names	Questionnaire
1985-2019	*ret_yn1/2* *ret_val1/2*	Other than Social Security or VA benefits, did you receive any pension or retirement income in [year]? How much did you receive from this income source in [year]?
1985-2023	*sur_yn1/2* *sur_val1/2*	Did you receive any survivor benefits in [year] such as widow's pensions, estates, trusts, insurance annuities, or any other survivor benefits other than Social Security or VA benefits? How much did you receive from your survivor benefits in [year]?

[37] Questions are obtained from Appendix E (Facsimile of ASEC Supplement Questionnaire) of the 2023 CPS ASEC Technical Documentation.

1985-2023	*dis_yn1/2* *dis_val1/2*	Did you receive any income in [year] as a result of your health problem other than Social Security Disability or VA Benefits? How much did you receive before deductions in these payments in [year]?
2019-2023	*pen_yn1/2* *pen_val1/2*	During [year] did you receive any pension income from a previous employer or union, other than Social Security or VA benefits? Do not include distributions or withdrawals from IRAs, 401(k)s, or similar accounts. How much did you receive in these pensions in [year]?
2019-2023	*dst_yn1/2* *dst_yn1/2_yng* *dst_val1/2* *dst_val1/2_yng*	At any time during [year] did you have any retirement accounts such as a 401(k), 403(b), IRA, or other account designed specifically for retirement savings? How much was your withdrawal or distribution from this account in [year]?

OASDI, OASI and DI

The variables *ss_yn* and *ss_val* identify OASDI recipients and dollars, respectively. Respondents are asked: "During [year] did you receive any Social Security payments from the U.S. Government?" and, if yes, "How much did you receive in Social Security payments in [year]?" Beginning in the 2001 survey, respondents are also asked: "What were the reasons you were getting Social Security in [year]?" If *rsnss1* or *rsnss2* = 2 (disabled), all Social Security dollars and receipt are assigned to DI; otherwise, they are assigned to OASI.

SSI

The variables *ssi_yn* and *ssi_val* identify SSI recipients and dollars, respectively. Respondents are asked: "During [year] did you receive: any SSI payments, that is, Supplemental Security Income?" and, if yes, "How much did you receive in Supplemental Security Income payments in [year]?"

UI

The variables *uc_yn* and *uc_val* identify UI recipients and dollars, respectively. Respondents are asked: "At any time during [year] did you receive any State or Federal unemployment compensation?" and, if yes, "How much did you receive in State or Federal unemployment compensation during [year]?"

VA DC

The variables *va_yn*, *va_val* and *vet_typ1* identify VA DC recipients and dollars. Respondents are asked: "At any time during [year] did you receive: Any Veterans' (VA) payments?" then, "What type of Veterans' payment did you receive?" and, "How much did you receive before deductions in [type of veteran's payment] in [year]?" Respondents are classified as VA DC recipients if *va_yn* = 1 and *vet_typ1* = 1 (service-connected disability compensation).

SNAP

The variables *hfoodsp* and *hfdval* identify household SNAP receipt and benefits amounts. Respondents are asked: "At any time during [year], did you receive benefits from SNAP (the Supplemental Nutrition Assistance Program) or the Food Stamp program, or use a SNAP or food stamp benefit card?" and, if yes, "What is the value of the food assistance received in [year]?"

Medicaid

The variable *caid* identifies Medicaid coverage. Respondents are asked: "Were you covered by Medicaid last year?"

Housing Assistance

The variables *hpublic* and *hlorent* ask respondents: "Is this public housing, that is, is it owned by a local housing authority or other public agency?" and, if no, "Are you paying lower rent because the Federal, State, or local government is paying part of the cost?" Respondents answering "yes" to either question are classified as receiving Housing Assistance.

Housing subsidy amounts are imputed in the survey. The variable changes starting with the 2019 CPS ASEC (from *fhoussub* to *spm_caphousesub*) and the imputation method also changes. In earlier years, the amounts are based on data from the 1985 American Housing Survey. Beginning with the 2019 survey, amounts are estimated by statistical match to HUD PIC and TRACS using geography and household, and HUD program rules are then used to impute amounts based on the household income reported in the survey.

A.2. Adjusting for Non-Random Non-PIKing

This appendix section describes our adjustment for non-random non-PIKing in the CPS ASEC, discusses the identifying assumption it requires, and presents evidence on its performance.

Why Non-PIKing Requires an Adjustment

The Census Bureau's Person Identification Validation System (PVS) assigns a PIK by probabilistically matching identifying fields – such as name, date of birth, and address – against a reference file built from the SSA Numident and supplemented with additional administrative and geographic sources (Wagner and Layne 2014). In our setting, PVS is used to assign PIKs to CPS survey records so that individuals in the survey can be linked to administrative records. Because PIK assignment does not require individuals in the survey to opt into linkage, our linked sample is not subject to the consent-induced selection that arises in surveys requiring explicit linkage consent (e.g., Sakshaug and Kreuter 2012). Erroneous linkage among records assigned a PIK also appears to be rare (Layne, Wagner, and Rothhaas 2014).

The more consequential issue for us is failure to assign a PIK. In our survey sample, 81% of individuals in 1996 and 85% in 2017 receive a PIK (86% and 91% of households have at least one PIKed member). The main reasons for non-PIKing likely reflect limitations or discrepancies in the identifying information available for matching, although some may arise because an individual is absent from the reference files (i.e., has not obtained an SSN or ITIN). For example, a respondent may provide an incomplete or incorrectly reported name or date of birth. They may give a name that differs from the one on the reference file (e.g., a middle name rather than a first name) or have a name common enough to prevent unique matching (e.g., "Joe Smith"). They may also have an address that differs from the corresponding information in the reference file.

Critically, failure to receive a PIK is not random. Bond et al. (2014) show using the 2009-2010 ACS that young children, racial and ethnic minorities, recent immigrants, frequent movers, low-income and non-employed individuals, and residents of group quarters all receive PIKs at systematically lower rates. Since we cannot assess survey error for observations without a PIK, all of our analyses must be conducted on the PIKed sample, and unadjusted estimates from that sample would not be representative of the CPS population. For example, if low-income households are less likely to receive a PIK, then estimates of both program receipt and survey error in reporting receipt may be biased in the linked sample.

Inverse Probability Weighting

To restore representativeness, we apply inverse probability weighting (Wooldridge 2007). Specifically, we model the probability that unit i receives a PIK as a probit function of observed survey covariates X:

$$\Pr(PIK_i = 1 \mid X_i) = \Phi(X_i'\beta),$$

where Φ is the standard normal cumulative distribution function. We use a probit specification rather than a linear probability model because our object of interest is a predicted probability, and the probit bounds all predicted values between 0 and 1. We then divide the base survey weight for each PIKed observation by its estimated probability of being assigned a PIK, $\hat{p}_i$, yielding the IPW-adjusted weight:

$$w_i^{IPW} = \frac{w_i^{base}}{\hat{p}_i}.$$

The identifying assumption is that, conditional on X, failure to receive a PIK is unrelated to the survey-error outcomes that would be observed if the record could be linked. Although this assumption is inherently untestable, two features of our setting make it considerably more plausible. First, X is unusually rich, incorporating extensive demographic, geographic, household-structure, labor-market, and income-source information collected in the ASEC, including the main characteristics that Bond et al. (2014) identify as predicting PIK assignment. We provide more details on these covariates below. Second, we can assess the performance of the reweighting by examining whether the reweighted linked subsample reproduces weighted recipient and dollar totals for various income sources observed in the full survey sample.

We estimate the model separately at the individual and household level, and adjust at the level corresponding to the unit of analysis for a given income source. Specifically, we use the individual-level model for all sources except SNAP and housing assistance, which are received at the household level. The individual model is estimated on all individuals and predicts whether a given individual is PIKed. The household model is estimated on household heads in households with at least one PIKed member and predicts whether a household has any PIKed member. In the household model, individual-level covariates (e.g., age, sex, race) refer to the household head, and household-level measures are constructed as composites of individual characteristics.

Covariate Specification

The set of covariates for the household model includes characteristics of the household head together with household-level income and demographic measures. The set of covariates for the individual model includes most of the household-level covariates (so as to capture intra-household resource sharing) while adding person-specific measures such as individual income-source receipt and disability status. The one exception is income-source amounts by quartile: including both individual- and household-level amount indicators causes the individual model to fail to converge, so the individual specification includes only individual-level amount indicators. Specifically, the covariates comprise:

- *Household structure and size*: five mutually exclusive household types (single adult with children; single adult without children; multiple adults with children; multiple adults without children, all conditional on a non-elderly head; and households with a head aged 65 or older), plus counts of children, non-elderly adults, and elderly members.
- *Demographics*: five-year age bins (0–74, with 75+ omitted); a male indicator; interactions of broad age groups (under 25, 25–54, 55–64, 65+) with sex; mutually exclusive race-ethnicity indicators (Hispanic of any race, non-Hispanic Black, non-Hispanic White, non-Hispanic Asian); a foreign-born indicator; a non-citizen indicator; and interactions identifying foreign-born Hispanic individuals and non-citizen foreign-born Hispanic individuals, which are intended to capture the undocumented population.
- *Geography*: state fixed effects and a rural indicator.

- *Income*: indicators for household receipt of each income source (OASDI or OASI and DI, SSI, TANF, UI, SNAP, housing assistance, VA benefits, pension income, dividends, interest, rental income and royalties, self-employment, and wage and salary income); indicators for income-source amounts by quartile; an indicator for self-employment income exceeding 25% of household income; an indicator for negative self-employment income; and indicators for bins of equivalized household income and earnings relative to the SPM poverty threshold.
- *Human capital and labor market*: educational attainment indicators; an interaction of receipt of "high" income sources (interest, dividends, rental income and royalties, and substantial business income) with a bachelor's degree or higher; marital status; labor force participation; an indicator for working 35 or more hours per week; and public and private health insurance indicators.
- *Disability*: household and (in the individual model) individual disability indicators.

Validating the IPW Adjustment

The conditional-randomness assumption is not directly testable, since we do not observe administrative records for the individuals who fail to receive a PIK. It does, however, have an implication we can check, as reweighting the PIKed sample by the inverse of the estimated PIK probability should reproduce the base-weighted distribution of survey variables. Appendix Table A.1 reports these checks for the income sources we analyze in this paper. For each income source we compute weighted totals of recipients and dollars in the full CPS sample using the base weights and in the PIKed sample using IPW-adjusted weights, and we also report the latter as a share of the former. The estimates are consistently close to 100%. In 2017, reweighted dollar totals range from 98.7% (UI) to 100.2% (pensions) of full-sample totals, with every other source within 0.3% of 100%. Reweighted recipient totals range from 99.6% (Medicaid) to 100.7% (SNAP). In 1996, reweighted dollars are within half a percent of full-sample totals for pensions, OASDI, and SSI, and reweighted recipients are within two-thirds of a percent for those sources and for housing assistance. We also see no clear trend across years in how the reweighted PIKed totals compare to the full-sample totals, which supports treating the adjustment as performing comparably over our study period despite the individual PIK rate rising from 81% to 85%.

Coverage of Individuals without a Social Security Number (SSN)

One potential concern with our reweighting is that receipt of the income sources we study generally requires an SSN, and one might worry that the non-PIKed group consists disproportionately of individuals without an SSN (who are ineligible for these programs). In that case, weighting up PIKed individuals who have positive receipt propensities conditional on covariates to represent a group that cannot receive benefits would overstate receipt in the reweighted sample, and the extent of this problem could change over time.

Several considerations address this concern. First, as previously discussed, the main reasons for non-PIKing are missing or inconsistent identifying information rather than the absence of an SSN, and the population without either identifier is likely to be very small. Second, individuals with ITINs (Individual Taxpayer Identification Numbers) but no SSN can receive PIKs and appear frequently in the CPS. Our model is therefore not extrapolating to the non-SSN population from scratch, as we can observe a sizable group of PIKed individuals who lack SSNs, are ineligible for the programs we study, and report no receipt. Third, and perhaps most importantly, our estimates are defined relative to the administrative data – i.e., we measure the share of *recipients* of various income sources (as recorded in administrative records) who are represented by the CPS, not the share of all individuals residing in the U.S. Because program receipt requires an SSN, individuals without an SSN lie outside the recipient universe by construction. They can affect our estimates only through the representation of non-recipients in the reweighted sample, and any resulting bias is bounded by their small population share.

Appendix Table A.1 provides evidence consistent with this final margin being handled adequately. If individuals without SSNs were fully covered by the CPS and poorly represented by the reweighting, then the reweighted PIKed sample would overstate survey-reported totals of recipients and dollars relative to the full CPS sample. The reweighted totals instead sit very close to full-sample totals in both 1996 and 2017, with no clear pattern across years.

References

Bond, Brittany, J. David Brown, Adela Luque, and Amy O'Hara. 2014. "The Nature of the Bias When Studying Only Linkable Person Records: Evidence from the American Community Survey." CARRA Working Paper 2014-08. Washington, DC: U.S. Census Bureau.

Layne, Mary, Deborah Wagner, and Cynthia Rothhaas. 2014. "Estimating Record Linkage False Match Rate for the Person Identification Validation System." CARRA Working Paper 2014-02. Washington, DC: U.S. Census Bureau.

Sakshaug, Joseph W., and Frauke Kreuter. 2012. "Assessing the Magnitude of Non-Consent Biases in Linked Survey and Administrative Data." *Survey Research Methods*, 6(2): 113-122.

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

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

A.3. Sources and Adjustments for Public Aggregates

This appendix section documents the publicly available administrative data sources used to benchmark CPS ASEC reports of income, along with the adjustments applied to align these administrative aggregates with the CPS ASEC sample frame and reference period. All administrative income sources are constructed to correspond as closely as possible to CPS ASEC income concepts, which measure income received during the prior calendar year by the non-institutionalized U.S. civilian population.

Sources of Publicly Available Administrative Data

This section describes the sources and construction of administrative aggregates spanning calendar years 1984 to 2022. For each income source, we summarize the underlying data and the unit of observation (e.g., individual, case, household). Our goal for each income source is to recover 1) total benefit dollars paid out during a calendar year and 2) total recipients at any point during a calendar year (to align with the CPS ASEC's annual income reporting window). We therefore also describe any adjustments required to align the administrative aggregates with these concepts (e.g., monthly to annual recipient counts, fiscal year to calendar year, etc.).

OASDI, OASI, and DI

Total annual benefit dollars for OASDI, as well as OASI and DI separately, are obtained from the *Annual Statistical Supplement to the Social Security Bulletin 2024*, Table 4.A4. These totals are reported on a calendar-year basis and cover all benefit payments made during the year. Total annual individual recipient counts are not directly published. To approximate annual recipients, we construct counts by adding total annual terminations to December monthly recipient counts. December monthly beneficiaries provide a lower bound on annual recipients, while terminations capture individuals who received benefits earlier in the year but were no longer receiving benefits in December. Total annual OASDI terminations are taken from Table 6.F1 of the *Annual Statistical Supplement*. We then disaggregate total OASDI terminations into OASI and DI using the distributions reported in Tables 6.F1 and 6.F3 for each year.[38]

UI

Total annual UI payments are sourced from the Department of Labor's Employment & Training Administration Form 2112. We do not have publicly available administrative data on the number of individual UI recipients. The Department of Labor publishes counts of UI claims, but this is not comparable to the number of UI recipients as individuals can claim unemployment benefits multiple times in a single year.

VA Disability Compensation

Fiscal-year benefit dollars and recipient counts for VA Disability Compensation are obtained from two sources: the *VBA Annual Benefit Reports* (1999-2022) and the *Annual Report of the Secretary of Veterans Affairs* (1984-1998). These sources report totals on a federal fiscal-year basis (October-September). To align VA Disability Compensation aggregates with the CPS ASEC calendar-year reference period, we convert fiscal-year values (F) to calendar-year values (C) using a weighted average that reflects the overlap between fiscal and calendar years. Specifically, for calendar year t, we compute:

$$C_t = 0.75F_t + 0.25F_{t+1}.$$

This approximation assumes a roughly uniform flow of payments throughout the fiscal year.

[38] Table 6.F1 provides total benefits terminated, as well as benefits for retired workers, widowed mothers and fathers, widow(er)s, parents of deceased workers (all of which we allocate to OASI) and disabled workers (allocated to DI). We use Table 6.F3 to allocate spouses of retired workers and children or students of retired or deceased workers to OASI and spouses, children, and students of disabled workers to DI.

SSI

Total annual SSI benefit dollars are obtained from the *Annual Statistical Supplement to the Social Security Bulletin*, Table 7.B1 for 1991 to 2022 and Table 9.B1 for 1984 to 1990. These tables report calendar-year totals of federally-administered SSI payments.[39] Total annual SSI recipient counts are available from the *Annual Report of the SSI Program* for 1996 to 2022 (rounded to the nearest 100,000). For earlier years (1984-1995), only December monthly recipient counts are published. To recover annual recipient counts for these years, we calculate a multiplier (M) using 1996 data, the earliest year for which we have both annual and December monthly counts from the SSA:

$$\mathrm{M}_{1996} = \frac{\text{1996 Total Annual Recipients}}{\text{1996 December Monthly Recipients}} = 1.074$$

We then apply this multiplier to December monthly counts for 1984 to 1995 to estimate total annual recipients. This approach assumes stability over time in the ratio of annual to December recipients from 1984 to 1995; we observe that this ratio is roughly stable from 1996 to 2022.

Medicaid

Total annual Medicaid recipient counts are drawn from multiple sources. For fiscal years 1985 to 2013, we use the *Annual Statistical Supplement 2015*, Table 8.E1. For 2014 to 2022, we use the Medicaid and CHIP Payment and Access Commission (MACPAC) *Reports to the Congress on Medicaid and CHIP*, Exhibit 1. Because these sources report figures on a fiscal-year basis, we convert both recipient counts and dollar totals to calendar-year values using the same weighted-average conversion formula that we applied to VA Disability Compensation. This produces calendar-year Medicaid recipient counts that are comparable to CPS ASEC reports. Because the CPS ASEC does not ask respondents to report Medicaid benefit amounts, we use administrative Medicaid aggregates solely to benchmark recipient counts, not dollars.

SNAP

[39] However, CPS ASEC respondents may also include state-administered SSI payments (as the survey does not distinguish between the two types). As a result, our estimated SSI reporting rate using public data may be understated.

SNAP administrative aggregates are constructed using three complementary sources: (i) USDA Food and Nutrition Service SNAP Data Tables, (ii) the University of Kentucky Center for Poverty Research (UKCPR) *National Welfare Data*, and (iii) historical expenditure data provided by Professor James Ziliak (University of Kentucky), as documented in Ziliak, Gundersen, and Figlio (2000). For 1984 to 1989, caseload data come from UKCPR state-level fiscal-year files, while expenditure data come from Ziliak's historical series. For 1989 to 2022, both caseloads and expenditures are drawn from USDA SNAP Data Tables at the state-by-month level. To obtain aggregate SNAP dollars in a calendar year, we simply sum expenditures across all months in a calendar year. However, calculating aggregate SNAP recipient counts is more complex, as these USDA tables report average monthly caseloads rather than totals for each month. Additionally, SNAP recipient counts in administrative data are reported at the *assistance unit* level, whereas CPS ASEC aggregates correspond to *households* that received SNAP at any point during the year. To reconcile these differences, we use restricted-use microdata to estimate two multipliers: (i) a case-to-household conversion factor, and (ii) an average-monthly-to-annuafl conversion factor.[40] These multipliers are applied to the public SNAP recipient aggregates to construct annual household-level recipient counts comparable to the CPS ASEC.

Pensions

Aggregate pension dollars are constructed using disclosed 1099-R totals from restricted-use data, as no comprehensive publicly available administrative source exists that reports pension income dollars consistent with CPS ASEC concepts. While some studies have used the National Income and Product Accounts as an administrative source of pension dollars, these data report pension income in terms of accruals rather than distributions (as measured in the CPS ASEC). Pension dollars for reference years 1995 and 2016 are obtained from Corinth, Meyer, and Wu (2022), dollars for 2010 are from Meyer and Wu (2024), and 2018 from Corinth et al. (2021). Pension dollars for 2012 are obtained from Bee and Mitchell (2017). One caveat is that, for many

[40] For the case-to-household multiplier, we use survey-linked administrative data to identify the weighted average number of SNAP cases per survey household, in the five states for which we have consistent restricted-use data over our time horizon. In 2017, this multiplier is 1.07, which we apply to all years (until multipliers for each year are disclosed). Czajka et al. (2015) find a similar case-to-household multiplier of 1.08 in New York and 1.06 in Colorado. For the average-monthly-to-annual multiplier, we use the universe admin data to calculate average monthly cases and total annual cases over the same states; in 2017 this multiplier is 1.37 (which we apply to all years until annual multipliers are disclosed). Using national SIPP data, the USDA similarly finds an average-monthly-to-annual multiplier of 1.4.

of these sources, the dollar amounts are not totals from the universe of administrative 1099-R data but rather weighted linked totals from the CPS ASEC for the relevant reference year (which may introduce coverage error).

Adjustments to Administrative Data for Intentional Differences in Coverage

We adjust all administrative aggregates to match the CPS ASEC target population and reference period, generally following the procedures outlined in Rothbaum (2015). These adjustments account for differences in population coverage, geography, and survival between administrative data and the CPS ASEC.

Group Quarters

The CPS ASEC excludes institutionalized populations and active-duty military members living on post without their families. To account for this, we adjust administrative aggregates downward to remove benefits received by individuals living in group quarters (GQs). For reference years 2005 to 2022, we use the ACS to estimate the share of survey-reported recipients and dollars for each income source that accrue to individuals living in GQs. While the CPS technically includes non-institutional GQ populations, in practice this population is small in the CPS sample. Thus, we subtract benefits accruing to all ACS GQ individuals.[41] As the ACS income questions refer to the prior 12 months and the ACS is collected throughout the year, for reference year t we use ACS year $t + 1$. Administrative aggregates are reduced by the estimated GQ share: we assume that the proportion of individuals in GQs in the ACS is the same as the actual proportion of GQ individuals in the administrative data (rather than assuming the level of group quarters is the same across the ACS and administrative data, which would be a stronger condition).

For reference years prior to 2005, we apply the GQ shares estimated from the 2006 ACS, as earlier ACS waves did not include GQ populations. In addition, we cannot apply a GQ adjustment for SNAP using publicly available data, since SNAP receipt is not collected for individuals living in GQs in the ACS. Instead, we calculate a group quarters share using ACS-linked SNAP administrative data for 2017 (approximately 7% for both recipients and dollars), and apply this share to all years. For income sources where ACS income categories are broader than

[41] Additionally, beginning in 2017, the CPS ASEC excludes college dormitories, the largest single source of non-institutional GQ populations.

CPS categories (notably UI and VA Disability Compensation, which appear within "Other Income"), we apply the GQ share estimated for the broader ACS category to each corresponding CPS subcategory.

Overseas & Territories

The CPS ASEC sample frame excludes individuals living overseas without a regular U.S. residence and residents of U.S. territories. Accordingly, we remove payments made to these populations from administrative aggregates. For OASDI, OASI, and DI, we use the *Annual Statistical Supplement* Table 5.J1 to identify payments made to foreign residents and U.S. territories. We remove the corresponding share of total benefit dollars and apply the same proportional reduction to recipient counts, which could lead to an overstatement (understatement) of this adjustment for recipients if foreign residents tend to receive smaller (larger) benefits. For SSI, analogous adjustments are made using Tables 7.B1 (1991-2022) and 9.B1 (1984-1990). Because no separate overseas adjustment is available for pension income, we assume that the share of retirement payments made overseas mirrors that of OASI and apply the same proportional adjustment. For SNAP and UI, our administrative data source explicitly excludes payments to recipients outside of the 50 states and DC. For Medicaid, recipients in Puerto Rico and the U.S. Virgin Islands are excluded from the administrative data for fiscal years 1999-2013. For the remaining years of Medicaid, we obtain dollar expenditures for the U.S. states and territories from the MACPAC Report, Exhibit 16. We calculate the share of all Medicaid expenditures that go to the territories and apply this proportional reduction to recipient counts, which could lead to an overstatement (understatement) of this adjustment if territory residents tend to receive smaller (larger) benefits. For VA Disability Compensation, we assume that few payments are made overseas or to territories and do not adjust the administrative aggregate.

Decedents

The CPS ASEC records income received during the prior calendar year only for individuals who are alive at the time of the survey (February to April of the following year). Administrative aggregates, by contrast, include benefits paid to individuals who died during the reference year. We therefore adjust administrative totals downward to account for income received by decedents who are not subsequently surveyed in the CPS. We estimate this adjustment using CPS ASEC microdata and Department of Health and Human Services (HHS) life tables by age, sex, and race/ethnicity (non-Hispanic White, non-Hispanic Black, or Hispanic). For each individual in the CPS, we assign a survival probability based on these characteristics and reweight the data accordingly. We then compute, for each income source, the implied share of income lost due to excluding decedents and reduce the corresponding administrative aggregate by this share. For SNAP, this procedure is performed at the household level using the demographic characteristics of the household head.

Individuals in the CPS who are neither White, Hispanic, nor Black are assigned White survival probabilities. Prior to 2007, Hispanic/non-Hispanic are not designated in the life tables, so we assign Black survival probabilities to Black individuals and White survival probabilities to all other individuals. Additionally, HHS life tables published prior to 1997 are abridged and use broader age categories, which produces substantially larger decedent adjustments. To avoid introducing discontinuities, we apply the 1997 life tables retroactively to reference years 1984 to 1996.

A.4. Estimating Non-HUD Housing Assistance Recipients

We assume that (1) the linked administrative variable accurately captures HUD housing assistance receipt (i.e., survey observations that do not link to the admin data do not receive HUD assistance) and (2) true false positive and false negative rates are the same for HUD and non-HUD housing assistance.

We present calculations for the linked CPS sample; to keep things simple, we omit weights from the formulas but in practice one would use survey weights adjusted for non-PIKing. Let x_i^S denote reported housing assistance receipt, x_i^A receipt according to the linked administrative measure (i.e., true HUD receipt), and x_i^* true receipt of either HUD or non-HUD assistance. Thus, x_i^S is the reported version of x_i^*, and x_i^A accurately captures the HUD component of x_i^* but misses non-HUD recipients. We observe the two measures with error (x_i^S, x_i^A) but not the accurate measure x_i^*. We wish to estimate the true number of non-HUD assistance recipients who do not receive HUD assistance: $\sum x_i^*(1 - x_i^A)$. Our starting point is the number of survey recipients who do not receive HUD assistance (the "excess survey recipients"): $\sum x_i^S(1 - x_i^A)$, which we know from our results. Let $\Pr(FP) = \Pr(x_i^S = 1 | x_i^* = 0)$ and $\Pr(FN) = \Pr(x_i^S = 0 | x_i^* = 1)$ be the "true" false positive and false negative rates, respectively. Under our assumptions, we can estimate $\Pr(FN)$ from true HUD recipients, but we cannot estimate $\Pr(FP)$.

Note that the excess survey recipients can be written as the sum of true false positives and reporting non-HUD recipients:

$$\begin{aligned}\sum x_i^S\left(1-x_i^A\right) &= \sum x_i^S(1-x_i^*)\left(1-x_i^A\right) + \sum x_i^S x_i^*\left(1-x_i^A\right)\\ &= \sum x_i^S(1-x_i^*) + \sum x_i^S x_i^*\left(1-x_i^A\right),\end{aligned}$$

where the second line follows from the fact that $1-x_i^A=0$ implies $1-x_i^*=0$.

The first term (false positives) is the false positive rate times the number of true non-recipients. The second term (reporting non-HUD recipients) is the number of true non-HUD recipients times the probability of reporting receipt (i.e., one minus the false negative rate). Thus:

$$\sum x_i^S\left(1-x_i^A\right) = \Pr(FP)\sum(1-x_i^*) + (1-\Pr(FN))\sum x_i^*\left(1-x_i^A\right).$$

In the first RHS term, true non-recipients are the number of observations who do not receive HUD assistance minus the number of people who receive non-HUD assistance:

$$\sum x_i^S\left(1-x_i^A\right) = \Pr(FP)\left[\sum\left(1-x_i^A\right) - \sum x_i^*\left(1-x_i^A\right)\right] + (1-\Pr(FN))\sum x_i^*\left(1-x_i^A\right).$$

Solving for our statistic of interest, $\sum x_i^*\left(1-x_i^A\right)$, yields:

$$\sum x_i^*\left(1-x_i^A\right) = \frac{\sum x_i^S\left(1-x_i^A\right) - \Pr(FP)\sum\left(1-x_i^A\right)}{1-\Pr(FN)-\Pr(FP)}.$$

Except for the true false positive rate, all quantities on the right-hand side can be estimated from the linked data. If we make assumptions on plausible values (ranges) of the true false positive rate, we can calculate (ranges of) the number of non-HUD assistance recipients. Defining the false positive rate we "observe" in the linked data, $\Pr(FP^{LD}) = \Pr(x_i^S = 1 | x_i^A = 0)$, we can re-write this expression as:

$$\sum x_i^*\left(1-x_i^A\right) = \frac{[\Pr(FP^{LD}) - \Pr(FP)]\sum\left(1-x_i^A\right)}{1-\Pr(FN)-\Pr(FP)}$$

Note that by using the linked data, these calculations abstract from coverage error and thus provide an estimate of the number of non-HUD housing assistance recipients covered by the CPS. If the survey covers the same fraction of HUD and non-HUD units, then total non-HUD units can be estimated by dividing the expression above by the fraction of HUD units covered by the survey.

Note that if there were no coverage error, then $\sum(1 - x_i^A)$ would be the population size (N_{pop}) minus the survey target (μ, i.e. adjusted total recipients from HUD admin data), so that the expression becomes:

$$\sum x_i^*(1 - x_i^A) = \frac{[\Pr(FP^{LD}) - \Pr(FP)](N_{pop} - \mu)}{1 - \Pr(FN) - \Pr(FP)}.$$