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

January 30, 2026

Abstract

Survey estimates are an essential foundation of research, policy, and public opinion. Strands of literature have documented serious error from four broad sources (survey coverage, unit non-response, item non-response, and measurement error), that can often be subdivided into false positives, false negatives, and errors in amounts. Yet the absence of a measure of error that is comparable across sources and over time limits efforts by survey producers to improve accuracy and by users to estimate and interpret survey statistics. We analyze and decompose survey error in an empirical Total Survey Error (TSE) framework that measures error in terms of bias, yielding a unified metric of direct relevance to survey producers and users that is comparable across error sources and time and inexpensive to routinely implement. We apply our approach to comprehensively measure survey error in the CPS ASEC, the official source of income and poverty statistics in the U.S., using linked administrative records for ten income sources spanning more than two decades. Our results show severe underreporting of income and program receipt that is driven primarily by measurement error among survey respondents on the extensive margin (false negatives). As a result, improving reporting among respondents may be the most promising margin for remedy. Non-respondent imputations are particularly noisy, with positive and negative errors that tend to offset in net terms. Moreover, we find strong evidence that absolute error (noise) has risen over time, even for income sources with relatively stable net bias. Overall, our approach provides a new method for measuring and decomposing survey error, clarifying which sources of bias require corrections and where improvements would be most consequential.

---

* Any views expressed are those of the authors and not those of the U.S. Census Bureau. 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 audiences at AEA/ASSA, NTA, and the Upjohn Institute for helpful comments. We are grateful to the Alfred P. Sloan Foundation, Russell Sage Foundation, Charles Koch Foundation, Menard Family Foundation, National Institute on Aging, and American Enterprise Institute for their support of the Comprehensive Income Dataset Project. 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 as the main instrument to collect information on a variety of measures, from concrete quantities like income and expenditures to subjective expectations and opinions. Governments allocate funding, produce official statistics (e.g., on poverty and unemployment), and score legislation based on household survey data. Pollsters run surveys to gauge public opinion and predict election outcomes. Researchers rely on surveys for representative samples and are increasingly designing their own surveys (see Stantcheva 2023). However, surveys are widely known to suffer from several main sources of error: households who are missed entirely by the survey (coverage/unit non-response error), respondents who do not answer specific questions (item non-response error), and respondents who give inaccurate answers (measurement error).[1] Separate strands of literature have extensively examined these main sources of error, each providing measures of survey accuracy such as response rates. This specialization creates two fundamental challenges. First, it is difficult to compare severity across error sources, because 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? Second, it is difficult to compare the severity of survey error over time, because the measures are indirect in that they quantify a symptom of the problem rather than the *distortion* to survey estimates. For example, a decline in response rates over time may not increase bias if non-response is random.

Survey producers, data users, and policymakers urgently need measures of survey error that are comparable across time and source and directly relate to bias, which is what matters for data quality. Without such measures, governments and policymakers cannot assess the reliability of official statistics and whether changes over time reflect true changes or error dynamics. Survey producers have several tools at their disposal to improve accuracy (e.g. Stantcheva 2023), but they need measures of whether and how much these tools reduce the distortion in their estimates to identify which survey design changes would cost-effectively reduce survey error. Survey users also need to measure and understand the sources of bias in their estimates to optimally choose their samples, corrections, and robustness checks to minimize the impact of survey error (e.g. Celhay, Meyer, and Mittag 2021).

[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.

In this paper, we develop an empirical approach to comprehensively measure survey error in terms of the bias it causes, directly quantifying the consequences of error for survey statistics. Our measurement system is based on the Total Survey Error (TSE) framework, which establishes a general and comparable metric of survey error but is rarely empirically implemented (Groves and Lyberg 2010). Extending Meyer and Mittag (2021), our approach uses data linkage to measure both net and absolute error. We decompose overall error into coverage error, item non-response error, measurement error, and a new component – ASEC non-response error – that measures a type of unit non-response error. We further divide most 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. Importantly, we demonstrate that our measures are simple and inexpensive to produce at high frequency for many variables, making the approach ideal for continuously and thoroughly monitoring survey accuracy.

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 and poverty statistics in the United States, and it is fielded as a supplement to the Basic Monthly CPS – a long-running household survey that produces that nation's headline labor-force statistics and has served as a model for many surveys in the United States and abroad. We measure survey error in the number of recipients and dollars received for ten income sources (private pensions, six cash transfers, and the three largest in-kind transfers in the U.S.) in 2017 and going back to 1996 in some cases. Implementing the method at scale – two outcomes each for ten income sources over many years - allows us to systematically compare error patterns across income sources and thereby provide a comprehensive assessment of survey data quality.

Importantly, these income sources provide the ideal conditions to measure survey error. We link the CPS ASEC to administrative tax records from the Internal Revenue Service and to program records from four federal agencies and five state agencies, providing accurate information on receipt and amounts for the entire population at the individual level. This method 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.[2] While our methods can be simplified for

[2] We do not argue that the linked measure is entirely error free, as errors may arise from incomplete administrative records, linkage errors, or from an inexact match to survey concepts. However, we consider these errors to be sufficiently rare to be of little or no practical relevance in our case. 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

cases where only aggregate information is available or where the individual or aggregate statistics only approximate truth, the accuracy and detail of our data allows us to explore the potential of our approach in a benchmark setting. At the same time, income and program receipt are substantively important because they are core inputs into measures of economic well-being (including poverty and inequality) and into analyses of the reach and take-up of government programs. Less concrete measures like subjective expectations and opinions are used just as frequently (or even more so) and may suffer from similar or greater errors, but are harder to validate.

We start by constructing measures of what the survey intends to yield (survey targets) for our ten income sources to determine the overall bias in surveys, which helps to assess which statistics are reliable and how others are distorted by survey error. Our results confirm that the survey misses substantial shares of recipients and dollars received for nearly all income sources, 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). For example, the survey misses more than a third of all dollars for five programs and more than 30% of all recipients for six programs. We then decompose TSE into four broad components (coverage error, ASEC non-response error, item non-response error, and measurement error) to understand the sources of overall bias in a recent year (2017). Understanding which error components are large helps survey users better handle error-ridden data and helps survey producers understand which parts of the survey need improvement. We find that measurement error makes up the majority of the overall bias for nine income sources (for dollars) and seven income sources (for recipients), suggesting that imputing additional receipt – as discussed in Mittag (2019) – may help. Our separate estimation of ASEC non-response error – where the CPS imputes entire ASEC records for households that respond to the monthly CPS but not the ASEC – and to a lesser degree item non-response error reveals that imputations also cause meaningful negative bias. Our findings suggest that excluding ASEC non-respondents and re-weighting the sample may substantially reduce error, which is particularly relevant for academic studies emphasizing causal effects over representative estimates.

aligned concepts. We also focus on income sources where these problems are least relevant (as opposed to e.g. earnings, asset income, or business income).

Next, we further decompose ASEC 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. This finer decomposition points survey producers to specific causes and potential remedies. We find that failure to report receipt constitutes the dominant source of bias. Even among respondents, for whom errors in amounts are more applicable, false negatives account for more than 100% of the overall bias in dollars for five income sources. This result points to reducing extensive margin underreporting (e.g., by probing more among likely recipients) as the most promising strategy to reduce bias. This matters doubly, as failure to report receipt translates into bias in imputed values (which are based on reports).

As a key extension, we provide a decomposition of absolute error to measure another important dimension of survey accuracy – noise – which is particularly relevant for multivariate analyses. This distinction reveals aspects that are crucial for overall data accuracy, because low net error can stem either from accurate data or from errors that cancel in aggregate. It is also important to distinguish net and absolute error for analyses of bias. While net survey error may be the ideal indicator for univariate analyses such as receipt rates, bias in multivariate analyses typically depends on measures of absolute error. For instance, bias in models with a misclassified dependent variable depends on the sum of the misclassification probabilities (Meyer and Mittag 2017), and bias in linear regressions generally depends on variances (Greene 2017). Our estimates of absolute error show that net error often masks substantial noise from offsetting error components. This is particularly concerning for imputations, which add substantial noise despite their smaller contribution to net error. False negative rates for imputations of ASEC non-respondents are more than twice as high as for respondents across five income sources, while false positive rates for item non-respondents are more than three times as high across eight sources.

Finally, we apply our decomposition annually to analyze how net and absolute error change over time, demonstrating the benefits of error measures that are comparable over time. We initially focus on four income sources for which we have administrative data in both 1996 and 2017—retirement income, Old-Age, Survivors, and Disability Insurance (OASDI), Supplemental Security Income (SSI), and housing assistance. These income sources happen to be 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 result is driven by two forces: 1) rising non-response rates shift the sample toward observations that are generally noisy at baseline and 2) noise

is increasing over time even conditional on respondent type. This finding shows that our distinction between net and absolute error is crucial as it distinguishes changes over time or due to survey design modifications that genuinely improve survey accuracy from changes that reduce bias by adding offsetting error. These error dynamics can also help survey users better deploy surveys for time-series analyses, just as our static measures of error can improve cross-sectional analyses. The signs and magnitudes of error and its components enable survey users to bias-correct estimates and assess robustness to likely changes in error. For survey producers and policymakers, understanding whether survey error improves or worsens – and which components drive trends – is essential for monitoring survey accuracy and improving data quality cost-effectively.

The remainder of the paper proceeds as follows. Section 2 explains how we implement the empirical total survey error decomposition and the extensions that we make to Meyer and Mittag (2021). Section 3 describes our data sources, linkage methods, and construction of survey targets. We discuss measures of overall error in Section 4 before decomposing net and absolute TSE at a point in time (for 2017) in Section 5. We examine the time-series of survey error and its components in Section 6. Section 7 concludes.

2. Methods: Empirical Total Survey Error Decomposition

Our approach rests on the ability to create measures of truth for key survey concepts by linking accurate records from the entire population to survey data. As Meyer and Mittag (2021) show, such linked data make it possible to estimate TSE and its components. The TSE Framework is the dominant paradigm in survey methodology to describe error properties of survey statistics. Groves and Lyberg (2010) provide a detailed discussion of the TSE Framework, its importance, and the numerous theoretical classifications of the sources of error in the TSE literature. We build on the decomposition of Meyer and Mittag (2021), who focused on average non-sampling error and use linked data to decompose TSE into error components corresponding to specific parts of survey design: generalized coverage error, item non-response error, and measurement error. We additionally estimate an error component arising from unit non-response to the ASEC part of the CPS.

These four components are each substantively important. Generalized coverage error includes error arising from an incorrect survey frame, unit non-response error, and the possibly

offsetting effects of adjustments such as sampling weight modifications.[3] A key part of coverage error arises from unit non-response, which has been in focus due to its rapidly increasing prevalence (Rothbaum and Bee 2021). Unit non-response is notoriously difficult to study (Bee, Meyer, and Gathright 2015), since no information beyond the sampled address is available about the missing respondents. The CPS allows us to separately estimate error from a type of unit non-respondent – households that respond to the Basic CPS but refuse to answer the ASEC.[4] For these households, the CPS imputes the entire ASEC, as we describe in the next section.[5] Item non-response error is arguably the most studied of the four components. Recent validation studies have shown that the common mitigation approach of imputation leads to high error rates at the household level (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 for variables such as income and education. Bound, Brown, and Mathiowetz (2001) provide an extensive review of measurement error in survey data.

To formally define these error components, let μ be the parameter of interest, which we want to estimate from random variables 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 receipt or the amount of dollars received for a specific program by household or individual i and μ is the receipt rate or average dollars received. We have two measures of x_i: x_i^A from the population data (such as our administrative records, which we consider to be accurate) 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 re-weighted to be representative of the population the survey actually covers. In our case, this amounts to assuming that multiplying the final survey weights by our inverse probability weight adjustment, $\widehat{w}_i$, addresses the problem of incomplete linkage. See Meyer and Mittag (2021) for discussion of this assumption. Finally, let

[3] 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 de Leeuw and de Heer (2002) and Mulry and Spencer (2001) for reviews of the extent of error and methods to address it.
[4] Meyer and Mittag (2021) include this error component in the measurement error component.
[5] In the absence of information from the Basic CPS, this error component would be an important part of unit non-response. The information from the Basic CPS allows us to link administrative records to these unit non-respondents (rather than just the sampled address), which provides a rare opportunity to examine unit non-response bias.

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 an estimate of the population mean obtained by weighting the survey reports x^s. We estimate this term from the linked data, using the survey weights w_i^S adjusted for incomplete linkage. 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 as the total of x_i^A from the administrative microdata divided by the population size. However, the administrative data may contain records the survey does not intend to cover (e.g., the CPS does not cover the non-resident population) as well as records that the survey intends to cover but cannot (such as records without a PIK). 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:

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

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 ASEC non-response error, $\hat{\varepsilon}_{ANR}$, which is the difference between the statistic of interest calculated using whole-imputed values (x_i^S) and actual values (x_i^A) for ASEC non-respondents. Item non-response error, $\hat{\varepsilon}_{INR}$, is the difference between the statistic of interest calculated using item-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 further decompose each of the three components estimated from unit-level observations – ASEC 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 ($\mathcal{CC} = \{x_i^S > 0 \ \& \ x_i^A > 0\}$). As a result, for 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 and measurement error similarly, with $r_i^U r_i^I$ changed to $1 - r_i^U$ for ASEC non-response error and to $r_i^U (1 - r_i^I)$ for item non-response error.

Note that for decomposing a mean, the terms for false positives and false negatives at least partially offset each other. Beyond understanding how much bias there is in a statistic, which our decomposition captures, it is important to also measure noise to better understand overall data accuracy and bias in multivariate analyses (as discussed in the introduction). For each of the three components we further decompose, we first estimate total absolute error in amounts, denoted by $\hat{\varepsilon}_{ANR}^{|Amount|}$ for ASEC non-response error, by replacing $x_i^S - x_i^A$ in the definition of $\hat{\varepsilon}_{ANR}^{Amount}$ 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 for absolute error in amounts to false positive error and subtracting false negative error (which is negative by definition). That is, for ASEC non-response error, we estimate our absolute error components as $|\hat{\varepsilon}|_{ANR} = \hat{\varepsilon}_{ANR}^{FP} - \hat{\varepsilon}_{ANR}^{FN} + \hat{\varepsilon}_{ANR}^{|Amount|} = |\hat{\varepsilon}_{ANR}^{FP}| + |\hat{\varepsilon}_{ANR}^{FN}| + \hat{\varepsilon}_{ANR}^{|Amount|}$. Absolute error components for item non-response error and measurement error are defined analogously. Total absolute survey error is then the sum of the absolute values of the four error components:[6]

[6] Our measure of absolute error thereby is a measure of average absolute prediction error, i.e. how well the survey values "predict" the administrative variable at the individual level. Alternatively, one could decompose mean absolute/squared error or the variance of the survey estimate, which would take into account factors such as changes

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

3. Data

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 contains a battery of detailed questions on income received during the previous year along with extensive demographic information (see U.S. Census Bureau 2025a for more detail). In this paper, we focus on the CPS ASEC from survey years 1997 and 2018, which correspond to reference years 1996 and 2017, respectively. The CPS ASEC is intended to be representative of the U.S. civilian non-institutional population and is the official source of income and poverty statistics in the U.S., making it one of the most important economic surveys in the country. The ASEC itself is a supplement to the Basic Monthly CPS, which samples approximately 60,000 households every month and is the official source of labor force statistics in the U.S. (see U.S. Census Bureau 2025b for more detail). As a result, CPS ASEC observations are a subset of those who respond to the Basic Monthly CPS.

There are two types of non-respondents to the CPS ASEC for whom the Census Bureau must impute income information: ASEC non-respondents (also called "full supplement" imputations) and item non-respondents.[7] ASEC non-respondents consist of persons or households who do not respond at all to the CPS ASEC but did respond to the Basic Monthly CPS. Consequently, the entire ASEC record is imputed from a donor record via a hot deck procedure that sorts observations into eight cells based on observed characteristics from the Basic Monthly CPS (U.S. Census Bureau 2002; Hokayem, Raghunathan, and Rothbaum 2015). Item non-respondents are persons or households who complete the ASEC supplement but fail to answer specific questions (e.g., recipiency or dollar amount for a given income source).[8] Imputation again

in the sample size and means. Such measures are more complicated than ours and still allow errors to cancel in the aggregate, so we believe that our measure is more informative about noise and likely biases.

[7] 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.

[8] In our paper, we classify an income source as item imputed if only its receipt is imputed. Note that if receipt is imputed, then amounts conditional on receipt are naturally imputed as well. However, in some cases, respondents report receipt but have their amounts imputed—in this case we do not classify that income source as imputed. In

follows a hot deck approach, but the details of implementation vary across variables. For example, the hot deck for SNAP classifies households into 648 cells that are specific to the SNAP income source. In contrast, some items are imputed jointly – for instance, Public Assistance is imputed jointly with other missing cash income components from a single donor. For more details, see U.S. Census Bureau (2021) and Celhay, Meyer, and Mittag (2021).

A key aspect of the imputation process is that measurement error among reporters will naturally propagate to item imputations because these imputes are based on reports. Similarly, these errors will further propagate to whole imputations that are 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, if not more noisy, than reports.

Table 1 lists the ten income sources for which we examine survey error in the CPS in this paper. Our unit of analysis varies by income source and corresponds to the level at which an income source is reported in the CPS. For seven of the ten income sources we examine – retirement income, OASDI, OASI, DI, SSI, UI, and VA disability compensation – our unit of analysis consists of individuals age 15+ in the CPS. For another income source (Medicaid), our unit of analysis consists of individuals of all ages. For two more income sources (SNAP and housing assistance), our unit of analysis consists of households. For receipt, the CPS contains separate questions for each income source.[9] The survey also asks respondents to provide amounts corresponding to each income source, with the exception of Medicaid and housing assistance.[10] All income sources are available in the CPS ASEC for both 2017 and 1996, with the exception of OASI and DI (which are not available in 1996 since the survey question about reason for receipt 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 Sources

addition, when an income source is defined as the sum of components (e.g., for retirement income), we call the overall income source imputed if any component is imputed.

[9] For retirement income, we define it as the sum of several components – i.e., private pensions, other private retirement distributions, private survivors' income, and private disability income. In addition, we are able to classify OASDI into OASI and DI based on an additional survey question that asks about the reason for receiving Social Security.

[10] For housing assistance, the Census Bureau imputes a benefit amount that is capped at the housing portion of the Supplemental Poverty Measure (SPM) threshold. We use this amount given that it is what is reported on public-use SPM files and likely what researchers rely on, but we acknowledge that it potentially understates true amounts (Renwick 2017).

We bring in a variety of administrative sources to provide measures of "truth" against which to evaluate survey accuracy.[11] A key challenge here 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 typically regarded as the gold standard – often government benefits for which we have records from the agencies that administer them, meaning the data are accurate and nearly complete. In contrast, we do not focus on income sources like earnings, business income, and asset income where there may be greater divergence between survey and administrative income concepts. Still, there are several income sources we analyze for which the administrative data cover a narrower concept of benefits than survey (namely SSI and housing assistance, as we discuss more below). For these sources, any negative (positive) survey errors are likely to be understated (overstated).

Table 1 lists the various administrative sources from which we draw. Data on retirement distributions come from IRS Form 1099-R, which covers gross payments from employer-sponsored plans and Individual Retirement Account (IRA) withdrawals and excludes rollovers (see Bee and Mitchell 2017 for a discussion on the conceptual alignment between these records and the CPS variables). Social Security (OASDI) benefits come from the SSA's Payment History Update System (PHUS) file, and we use the SSA's Master Beneficiary Record (MBR) to distinguish between payments from Old-Age and Survivors Insurance (OASI) and from Disability Insurance (DI). Our preferred administrative measures of OASI and DI are gross amounts that include medical insurance premiums withheld from the net disbursement, although we show robustness checks only using the net disbursement. Supplemental Security Income (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 administrative benefits can be received by children below 15, even though these children cannot report SSI and OASDI claimed on behalf of children, we attempt to reallocate some of the linked administrative dollars received by children to the adults in the household.[12]

[11] See for example Celhay, Meyer, and Mittag (2021), Meyer and Mittag (2019, 2021a), and Mittag (2019) for discussions of a specific data source.

[12] We handle three cases: (1) If no adults (age 15+, PIKed) in the household report program 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

Unemployment Insurance (UI) records come from Box 1 on 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.[13] Administrative 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 are available from the Medicaid Statistical Information System (MSIS), which is a national database managed by the Centers for Medicare and Medicaid services. 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). We calculate a household's benefit amount as the difference between the gross rent and actual tenant payment. Note that these HUD data cover only federally-administered programs and thus miss programs administered by the USDA and by states and localities (which may be asked about in the CPS).

In sum, our administrative data cover the entire U.S. for all income sources examined, except for SNAP (for which we bring in records from only 5 states). Furthermore, while administrative3 records for all ten income sources are available in 2017, they are available for only six income sources in 1996. At the Census Bureau, linkable records are only available from 2000 for Medicaid, from 2004 for SNAP, from 2007 for VA disability, and from 2010 for UI. Taking the intersection of what is available in both survey and administrative data for 1996, our dynamic analyses examine four income sources: retirement income, OASDI, SSI, and housing assistance.

3.3. Linking Data Sources

We link the administrative microdata to the CPS using person identifiers created by the Person Identification Validation System (PVS) of the U.S. Census Bureau (Wagner and Layne 2014). This linkage is essential for measuring survey error at the individual level. The PVS uses

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.

[13] 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 a couple states during the years we examine and are thus likely to be trivial (see Meyer et al. 2023 for more details).

the 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 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, respectively. The increase in PIK rates over time corresponds to the CPS changing from an opt-in to an opt-out question for data linkage.

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. 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 for SNAP and housing assistance, for which we adjust at the household level.

3.4. Constructing Survey Targets

To estimate total survey error, we need an aggregate measure of what the survey intends to yield – i.e., the survey target. Given that we treat administrative records as truth for most income concepts we analyze, a natural starting point is to calculate the total number of recipients and dollars for our various income sources in the administrative microdata. This should approximate aggregate totals typically available from publicly available 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 (2015), Meyer, Mok and Sullivan (2015), Meyer and Mittag (2021) and Corinth, Meyer, and Wu (2022).

We first implement a series of adjustments for intentional coverage differences, subtracting payments and recipients from the administrative microdata that are not covered by the survey

household sample. 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 in the CPS ASEC. We do this by linking the administrative records 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.[14] 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).[15] To do this, we link administrative records for year *t* to GQ records in the American Community Survey (ACS) for the subsequent survey year *t* + 1 and calculate the weighted total of linked dollars in GQs.[16] Third, we subtract records linked to residents who appear in a different state when surveyed in year *t* + 1. Note that this is relevant for administrative income sources with incomplete geographies, which is only SNAP in our case. Finally, to account for records not covered by the linked data, we subtract administrative records that do not link to a PIK.

We apply this methodology to all income sources for which we have universe administrative data – i.e., all of our income sources except for those from the SSA (OASDI, OASI, DI, and SSI). For these SSA income sources, we only have administrative extracts for those who link to the CPS ASEC. As a result, we apply public-use analogs of the adjustments for decedents and GQs to publicly available administrative totals (see Appendix Section A.2 for more details).[17] We do not apply an adjustment for non-PIKing in the administrative data for these sources.

In addition to adjusting survey targets from the administrative data for coverage differences, we must also ensure that the survey targets align in terms of unit of analysis (individual or household) and reference period (calendar year). For the unit of analysis, most income sources in the administrative data are already at the relevant unit. The only exception is SNAP, where the administrative records are at the assistance unit (case) level, which differs slightly from that of a

[14] 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.

[15] While the CPS in principle includes non-institutional GQs, in practice the counts are very small. Moreover, starting in 2017 (one of the key years we analyze), the CPS stopped including college dormitories (one of the largest non-institutional GQ types) in its sample.

[16] 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. Furthermore, 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.

[17] 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. See Appendix Section A.2 for more detail.

survey household. However, we are able to calculate the average number of SNAP cases per survey household in our linked data (1.07 for our 5 states in 2017), which we use to scale down the survey target to the level of a survey household. For the reference period, the tax-based income sources (retirement income and UI) are already at the calendar year level, and many other income sources are at the monthly level (which is straightforward to convert to calendar year). For the SSA income sources (OASDI, OASI, DI, and SSI), the publicly available recipient targets are monthly for December, but we can convert them to annual counts by adding total terminations at any point during the year (also published by SSA).

Table 2 shows the coverage adjustments for dollar amounts and recipients in the main year that we analyze (2017). In general, the largest adjustment is for GQs, who comprise at least 2% of observations for all income sources except for retirement income, UI, and VA disability. The GQ adjustment is particularly large for SNAP (6.8% of dollars and 7.2% of recipients), SSI (4.5% for dollars and 3.0% of recipients), and housing assistance (3.8% of dollars and 3.6% of recipients). The adjustment for decedents never exceeds 3%, though as expected it tends to be larger for income sources targeted to those who are elderly or disabled (for dollars: 2.3% for OASI, 2.0% for OASDI, 1.9% for retirement distributions, and 1.6% for VA disability). The adjustment for removing un-PIKed records comprises less than half of a percent for nearly all income sources for which we have universe administrative data, because one generally needs an SSN to receive a government benefit and there is a one-to-one match between SSN and PIK.[18] Finally, for SNAP we estimate that 1.2% of benefit dollars and 1.1% of recipients are surveyed in a different state from the one in which they received benefits per the administrative data.

4. Measures of Survey Error

We begin by establishing the need for our approach and presenting our first set of results. We first review commonly used indicators of survey accuracy, including unit non-response rates, item non-response rates, and measurement error rates. While these indicators raise clear concerns

[18] There are two prominent exceptions. 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.

about data quality, they are indirect measures that do not directly quantify bias in survey estimates. We then estimate TSE for dollars and recipients across our income sources. These TSE estimates are our first key results, as they provide a direct measure of bias and establish that the survey systematically understates both receipt and dollars, with this problem worsening over time for many programs. Finally, we explain why these TSE estimates, while valuable, still leave important questions unanswered. These questions require decomposing TSE into its components to address.

4.1. Indirect Indicators of Survey Error

A first set of indicators emphasized in the survey methodology literature are unit non-response rates, which capture situations where sampled units fail to provide survey information at all (or, in practice, where the survey must impute an entire record because the selected unit does not respond). In the CPS context, two types are particularly salient. Appendix Figure A.1 documents that the CPS ASEC whole-imputation rate (cases where a household responds to the Basic Monthly CPS but does not respond to the ASEC supplement) hovered around 10% between 1990 and 2009, then rose steadily to roughly 20% in recent years. But whole imputations are conditional on responding to the Basic Monthly CPS in the first place, and the Basic CPS itself has also experienced rising unit non-response rates – typically at 5-10% prior to 2010 and then increasing sharply to above 30% in 2022. Taken together, these two stages imply that by 2022, roughly 45% of the original Basic Monthly CPS sample did not provide responses to *any* CPS ASEC question (either because they never responded to the Basic Monthly CPS or because they responded to the Basic Monthly CPS but did not respond to the ASEC).

A second set of indicators focuses on item non-response, which capture cases where a respondent completes the survey but does not answer particular questions. Appendix Figure A.2 shows that item non-response rates have increased over time for nearly all income sources in the CPS ASEC. Unlike the sharp post-2010 rise in unit non-response measures, these item non-response increases appear more sustained over roughly the full 1990–2022 period. For example, item non-response rates quintupled for retirement income and OASDI. In levels, by 2022 item non-response rates were nearly 8% for Medicaid, 6% for retirement income, and 4% for SNAP (with most other sources around 2% or lower).

Beyond non-response, a growing literature documents substantial measurement error among respondents, often summarized using false negative and false positive rates. Appendix

Table A.1 compiles estimates from prior studies for various income sources, surveys, and settings. A consistent pattern is that false negative rates are often high. For example, focusing on CPS-based analyses, Meyer, Mittag, and Goerge (2022) find that 45% of SNAP recipients in Illinois and Maryland failed to report receipt in the CPS (survey years 2001–2005), while Celhay, Meyer, and Mittag (2021) find that 37% of SNAP recipients and 59% of TANF recipients in New York did not report receipt (survey years 2007–2012). While these estimates are striking, they also tend to come from case studies in a small number of states, years, and programs, leaving open how these patterns generalize across sources and over longer horizons.

These conventional metrics – including non-response and misclassification rates – raise clear concerns, but they do not by themselves tell us how the different error sources translate into bias in survey quantities such as total dollars received and numbers of recipients. They are indirect measures in that they quantify the extent of the problem (e.g., how many non-respondents) but not the distortion to survey estimates (e.g., the magnitude of the bias stemming from non-response). This is the fundamental gap that motivates our TSE approach.

4.2. Total Survey Error: A Direct Measure of Bias

To directly measure bias in survey estimates, we calculate TSE for dollars and recipients across our income sources. TSE is the difference between the weighted survey estimate and the survey target constructed from administrative data (after making public-data adjustments for intentional coverage differences, as described in Appendix Section A.2). Figures 1 and 2 summarize TSE over time in the CPS over time as a share of the survey target for eight income sources (dollars) and seven income sources (recipients) from 1984–2022.

In levels, TSE is almost always negative for both dollars and recipients (with SSI dollars an exception in recent years), indicating that survey totals are systematically biased downward. Yet, there is substantial heterogeneity across income sources. For dollars, TSE in recent years ranges from less than 20% understatement for SSA programs (OASDI, OASI, DI, SSI) to 30-50% understatement for UI, VA disability, and retirement distributions and over 60% understatement for SNAP. For recipients, TSE shows a more compressed range, with the greatest understatement of around 40–50% for SNAP and Medicaid.

Trends over time also generally point toward worsening TSE, but again with heterogeneity 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. Among recipients, TSE has worsened even for SSI, despite SSI dollars showing reduced understatement in recent years. This pattern is potentially consistent with increasing confusion between SSI and the larger OASDI program, a hypothesis we can examine directly in the decomposition. Overall, these TSE estimates establish that bias is substantial and generally increasing. For consumers of survey statistics, this finding suggests that estimates of receipt rates and income of low-income households should be interpreted with caution and potentially corrected upward.

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 instance, for SSI, VA disability, and Medicaid, much of worsening TSE for recipients occurs before the mid- to late-2000s – well before the sharp rise in unit non-response and ASEC non-response rates after 2010. Conversely, item non-response rose dramatically over time for OASDI and retirement income, yet these sources see comparatively small decreases in net reporting rates. More fundamentally, the indirect measures track the prevalence of different error sources but not their severity. Both of these aspects could matter – for example, TSE could increase either because more people are nonrespondents or because non-respondents are becoming more biased over time. Only a decomposition that measures TSE attributable to each source can separate various channels and reveal the mechanisms driving changes in survey accuracy. The remainder of the paper decomposes TSE into comparable components to rigorously assess the roles of different error sources in explaining levels and trends in bias.

5. Decomposing Total Survey Error at a Point in Time

We begin by decomposing total survey error (TSE) for each of our ten income sources in a single recent year (2017). The goal is to move from overall net bias to a unified breakdown into the four components defined in Section 2 – coverage error, ASEC non-response error, item non-response error, and measurement error – and then further unpack each component into its sub-errors. We then complement net decompositions with absolute error measures that capture the extent to which surveys are noisy even when different errors cancel out on average.

5.1. Decomposition of Net Survey Error

Figure 3 summarizes the main net decomposition results for 2017. It displays, for each income source, the contribution of each error component to total net survey error, expressed as a percentage of the survey targets. Each income source has two bars (one for dollars and one for recipients**)** except Medicaid, where we focus on recipients only. At the bottom of each bar, we report the TSE in parentheses for each variable and income source.

We start by walking through the results for a single income source (retirement distributions, in the top left of Figure 3). Total net TSE is -44% for dollars and -46% for recipients, indicating substantial understatement. The dominant contributor is measurement error among respondents, for which there is a 38% understatement for dollars and 36% understatement for recipients – accounting for roughly 79-87% of the total net understatement. Non-response error components also generate negative bias, but they are smaller in magnitude (8-9% for ASEC non-response and 3-4% for item non-response). Part of the reason for the smaller role of non-response error is that non-respondents are simply a minority of the sample; about 19% are ASEC non-respondents and, for retirement income in 2017, another 4% are item non-respondents. Yet, as we show below, another reason is that non-respondent errors are often noisier and can partially offset in net terms. Finally, coverage error (the residual category) is slightly positive for retirement income (+6% for dollars and +2% for recipients), consistent with mild over-coverage of the retirement-income population (e.g., through over-sampling or over-weighting).

Turning to patterns across all ten income sources, it is useful to first set aside SSI and housing assistance (which are outliers for reasons discussed below) and focus on the other eight sources. For these eight, net TSE is consistently negative for both dollars and recipients, and tends to be at least as large for recipients as for dollars. However, outside of the SSA programs (OASDI, OASI, and DI), net TSE for dollars and recipients is often quite similar in magnitude. This provides an initial suggestion that the extensive-margin understatement of receipt is likely the major driver of dollar TSE, as opposed to the intensive-margin misreporting of amounts conditional on receipt. Across these eight income sources**,** measurement error among respondents is the dominant component and is always negative. For dollars, measurement error accounts for the majority of TSE for all sources and exceeds two-thirds of TSE for five of the seven non-Medicaid sources. For recipients, it again constitutes the majority of TSE for six of the eight sources and exceeds two-thirds for four sources. This is notable given that the literature has often under-emphasized

measurement error among respondents and placed more weight on non-response error. Indeed, non-response errors (ASEC and item) consistently generate negative bias but are smaller in magnitude than measurement error. ASEC non-response typically contributes more than item non-response; in 2017, it is never more than a third of TSE, while item non-response is never more than about 10% of TSE. Finally, coverage error is generally modest: retirement-income recipients appear slightly over-covered, while other program recipients tend to be slightly under-covered. For DI and SNAP, under-coverage is more pronounced for recipients than for dollars, consistent with the survey frame missing lower-dollar recipients (e.g., for SNAP, potentially households with more churn and instability).

In contrast to the other eight sources, SSI and housing assistance behave differently in net terms. SSI is the only source with positive net TSE for dollars (a small overstatement of +2%) but still exhibits a large negative TSE for recipients (-25%). A natural hypothesis, which we investigate in the next subsection, is confusion between SSI and OASDI in survey reporting (especially given that both programs are administered by the SSA). If some true OASDI recipients misclassify themselves as SSI recipients in the survey, and OASDI benefits are larger on average, then that can generate the patterns we see. SSI also displays a large apparent coverage-error component (−17%), but note that this is mechanically inflated by a unit-of-analysis mismatch.[19] For housing assistance, net TSE is +12% for recipients but -56% for dollars. The positive recipient TSE is likely due to the fact that our administrative data cover only HUD-administered programs and thus misses millions of non-HUD units. As a result, the positive bias may not necessarily reflect error but instead recipients missed by our administrative data. At the same time, the large negative dollar TSE reflects the survey's use of a capped SPM housing subsidy amount (imputed by the Census Bureau for its SPM poverty measure) rather than an elicited benefit amount.

5.2. Extensive vs. Intensive Margin Errors

Given that we find net TSE to be almost always negative, we now examine – at least for dollar amounts – whether that bias arises primarily from (1) the understatement of receipt on the

[19] Specifically, children can receive SSI in administrative data while only those aged 15+ in the survey can report it. In cases when we cannot assign a linked child recipient to a survey adult, we classify these dollars and recipients into the coverage error residual. See Section 3 for more details. Based on survey-linked aggregates for SSI disclosed in prior work, our sense is that actual coverage error is approximately half what is reported here.

extensive margin versus (2) the understatement of amounts conditional on being a true recipient on the intensive margin.[20]

We start by calculating false negative rates, defined as the share of true recipients (based on administrative records) who fail to report receipt in the survey. This is therefore a measure of underreporting on the extensive margin. In the full sample, Table 3 shows that false negative rates exceed 39% for eight of the ten income sources. OASDI and OASI are the two exceptions, with false negative rates around 17-18%. For four sources (retirement income, DI, UI, and VA disability), more than half of true recipients fail to report receipt. Breaking false negatives out by respondent type reveals considerable heterogeneity. Respondents have lower false negative rates than the overall average but still exceed 36% for seven sources and exceed 50% for retirement income and UI. Item non-respondents have even higher false negative rates (consistent with the idea that imputations inherit bias from reports): among the seven sources that can be disclosed, false negatives exceed 64% for four (retirement income, DI, SSI, housing assistance). ASEC non-respondents, which are imputed based on even less information than item non-respondents, have staggeringly high false negative rates – above 68% for eight of ten sources and above 81% for five sources (DI, SSI, UI, VA disability, housing assistance).

Table 3 also reports false positive rates, defined as the share of true non-recipients who report receipt and thus an extensive-margin measure of overreporting. These rates range from 0.2% (VA disability) to 5.9% (Medicaid) in the overall sample. False positives are again lowest among respondents and higher among non-respondents. Notably, item non-respondents have the highest false positive rates (and are approximately 25% for Medicaid). In sum, these patterns suggest that non-respondents tend to have noisier information than respondents. They are more likely to be recorded as not receiving incomes when they actually do, and as receiving incomes when they actually do not. To translate these extensive-margin rates into contributions to TSE, we then combine false negatives and false positives with errors in reported (or imputed) amounts among those who correctly report receipt. Figure 4 presents this breakdown for dollars, with Tables 4 and 5 providing the underlying numbers for dollars and recipients.

For ASEC non-respondents and item non-respondents, net bias is almost entirely an extensive-margin phenomenon. Errors in amounts conditional on correctly reporting receipt are

[20] Note that the extensive- vs. intensive-margin distinction is relevant only for dollar error, since recipient error is only at the extensive margin.

trivial, which is not surprising given that the vast majority of true recipients among non-respondents are often imputed to not get any receipt. The negative contribution from false negatives exceeds the positive contribution from false positives for essentially all sources (SSI among ASEC 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 ASEC non-respondents, the bias from OASI false negatives (-4.3%) only slightly outweighs that from OASI false positives (+3.3%). But for many other sources (e.g., retirement income, UI, SNAP, and housing assistance), the bias from false negatives is more than twice as large as the bias from false positives.

Among respondents. who make up most of the sample and thus account for much of total TSE, false negatives again explain the bulk of net bias. For seven of nine dollar outcomes (excluding Medicaid), false negatives contribute more than 80% of net measurement error. 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 non-trivial for retirement income and SNAP, but still represent a minority of total error. SSI and housing assistance again stand apart from other income sources in a number of ways. For SSI, the bias from respondent false positives (+36%) exceed that from respondent false negatives (-23%), consistent with confusion between SSI and OASDI. In fact, we find that 53% of SSI false positives are also OASDI false negatives, suggesting that OASDI recipients (likely DI recipients) are often reporting their benefits as SSI. For housing assistance, the dominant contributor to respondent error is on the intensive margin in terms of amounts, which is consistent with the survey using a capped imputed housing subsidy measure that far understates actual benefit amounts.

Overall, our key takeaway is that extensive-margin underreporting (false negatives) is the dominant source of bias, and this pattern is remarkably consistent across income sources and respondent types.

5.3. Decomposition of Absolute Survey Error

So far we have focused on net bias. But because different error components can have opposite signs, net measures can conceal offsetting errors. We therefore also decompose absolute

TSE, which treats all component errors as positive and captures the overall "noise" of survey outcomes.

Figure 5 provides the absolute-error analog to Figure 3 for 2017, with the absolute error magnitudes all being positive by design. Table 4 also shows the absolute error magnitudes for the intensive- and extensive-margin subcomponents of each error component. In contrast to the patterns for net TSE, absolute TSE is now always larger for dollars than for recipients, reflecting the additional intensive-margin (amounts) error channel that applies to dollars but not to recipient counts. The income source for which net and absolute TSE are most strikingly different is SSI. 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 are obscured by net measures. More generally, SSI and housing assistance – which looked anomalous in net terms – look much more similar to other sources when it comes to absolute error.

Measurement error remains the largest contributor to absolute error for dollars for all income sources, and for recipients for seven of ten sources. However, measurement error tends to account for a smaller share of total error in absolute terms than in net terms, precisely because non-respondent errors are noisier and partially cancel in net measures but are fully counted in absolute measures.[21] ASEC non-respondents contribute roughly 18–40% of absolute error and item non-respondents contribute 3–17%, much larger roles than they play in net error. However, this is not to say that every component of measurement error is de-emphasized when analyzing TSE in absolute terms, as intensive-margin errors in amounts sometimes become more salient. For example, among OASDI respondents, net error in amounts is essentially zero (−0.1%), but the absolute error in amounts is 16.0%, large enough to dominate the combined extensive-margin bias from false positives and false negatives. This implies a substantial degree of (approximately mean-zero) noise in reported OASDI amounts conditional on true receipt.

6. Decomposing Total Survey Error Over Time

Having established the decomposition of both net and absolute TSE at a point in time, we next apply the framework to track how TSE and its components change over time. We focus on

[21] For example, for retirement income measurement error is around 80% or more of net TSE in net terms, while it is closer to two-thirds of absolute TSE. And for Medicaid, measurement error is around three-quarters of net TSE but less than half of absolute TSE.

four income sources (retirement income, OASDI, SSI, and housing assistance) for which survey and administrative data are consistently available in both 1996 and 2017.

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

To start, Figure 6 decomposes net TSE for these four sources in 1996 and 2017 and separately for dollars and recipients. From the descriptive TSE patterns in Figures 1 and 2, these sources are among the least problematic in terms of changes in net error over time. And indeed, there is a fair amount of heterogeneity in the direction of change over time: net TSE improves slightly for retirement income, worsens somewhat for OASDI, moves from negative to slightly positive for SSI dollars while worsening for SSI recipients, and becomes slightly more negative for housing assistance dollars and less positive for housing assistance recipients. Much of this ambiguity reflects the fact that measurement error among respondents – still the largest component in most cases – does not change monotonically in net terms. Coverage error also exhibits heterogeneity across income sources but in general tends to be smaller than measurement error. In contrast, net errors attributable to ASEC and item non-respondents grow considerably over time (often more than doubling), even if they remain smaller than measurement error in level. This raises a key question: when net error improves or stays flat, does that reflect genuine improvements in accuracy or simply greater cancellation of offsetting errors?

Figure 7 answers this by examining trends in absolute error, which no longer allows offsetting components to cancel. Across all four income sources, absolute error increases sharply from 1996 to 2017.[22] This implies that even when net bias changes little (or even improves), survey accuracy can still be deteriorating in that the survey may become noisier. SSI again illustrates the point starkly: net TSE for SSI dollars moves from a 13% understatement in 1996 to a 2% overstatement in 2017, yet absolute TSE rises from 79% to 124%. 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.

[22] Specifically, for measurement error among respondents (which often evolved ambiguously for net error), the evidence is stronger here that it has increased in absolute terms for both OASDI and SSI while it has only slightly decreased for retirement distributions and housing assistance. But even for the latter two income sources, the decreases in absolute error among respondents are not sufficient to offset the larger increases in absolute error coming from imputations – such that overall absolute error rises over time.

6.2. Extensive vs. Intensive Margin Errors Over Time

Overall, we have documented increases in absolute TSE for each of the four income sources we examine over time, with the increases most pronounced for ASEC and item non-respondents. Table 6 (Panel A) corroborates this pattern on the extensive margin. In particular, overall false negative rates increase for all four income sources between 1996 and 2017, nearly doubling for OASDI and rising by more than 50% for SSI. Over the same period, overall false positive rates rise by roughly 50–80% for three of the four income sources; the only exception is housing assistance, where false positives fall slightly (though this could be attributable to changes in the completeness of the administrative data over time).

What explains these increases in extensive-margin error over time? One clear force is compositional. As Appendix Figures A.1 and A.2 show, non-response has increased substantially, and because non-respondents are markedly noisier at baseline, rising non-response mechanically shifts the survey sample toward observations that are more error-prone. But there also remains the possibility of a second layer of deterioration – namely, that errors are also rising within respondent type, such that respondents and non-respondents are becoming less accurate over time. Figure 7 already provides suggestive evidence towards this. For several income sources, most notably OASDI and SSI, the bias from measurement error increases over time even though respondents are a declining share of the overall sample. In this subsection, we can test this hypothesis directly. Table 6 also reports how false negative and false positive rates evolve over time separately for respondents, ASEC non-respondents, and item non-respondents. Among respondents, false negative rates increase for OASDI (from 9% to 15%) and SSI (from 28% to 41%), while remaining essentially flat for retirement income (53–54%) and falling modestly for housing assistance (from 30% to 27%). False positive rates among respondents also rise for three of the four income sources (again with housing assistance as the exception). The increase in false positives for retirement income is especially noteworthy because the Census Bureau implemented a redesign of the CPS ASEC in the mid-2010s precisely to elicit more complete reporting of retirement income. At a superficial level, Figure 6 might suggest the redesign was successful, since net underreporting of both dollars and recipients for retirement income declines. But the increase in false positives implies that at least some of the additional reported receipt may be coming from the wrong people. Taken together, these patterns indicate meaningful evidence of rising noise among respondents,

and importantly, that this noise grows on both sides of the extensive margin (more missed true receipt and more spurious receipt), even when net bias changes only modestly.

We see parallel evidence of increasing noise among non-respondents. For ASEC non-respondents, false positive rates exhibit relatively little change over time, but false negative rates increase for three of the four income sources (all except retirement income). For item non-respondents, the pattern is closer to the reverse: false negative rates are relatively stable over time (with the important exception of housing assistance, where they rise from 40% to 68%), while false positive rates increase for three of the four income sources (all except SSI). Thus, even conditional on non-response status, extensive-margin errors tend to worsen over time in several key cases, reinforcing the interpretation that rising non-response rates are not the sole driver of deteriorating accuracy.

Tables 7 and 8 translate these extensive-margin changes into contributions to changes in overall net and absolute TSE for dollars and recipients, respectively. But our evidence of rising noise over time is not confined to the extensive margin. Intensive-margin errors in amounts also evolve meaningfully, particularly for respondents (Panel D of Table 7), where amount misreporting among true recipients who report receipt is most relevant. This evolution is largely missed if one focuses only on *net* amount error, since the bias in reported amounts tends to decline over time for all four income sources. In contrast, when we examine *absolute* amount error, we find it increases over time for both retirement income and OASDI, while falling for SSI and housing assistance. Overall, across the margins we can measure – extensive versus intensive, and net versus absolute – there is consistent evidence that survey outcomes have become noisier over time in important settings, even when net reporting rates appear stable or improve.

7. Conclusion

In this paper, we extend and implement a framework to measure survey error based on an empirical TSE decomposition. By measuring error in terms of its consequences, i.e. in terms of bias, our framework provides a general measure of survey error that is of direct relevance and comparable over time and across sources. We apply our approach at scale – to two outcomes each for ten income sources over many years – demonstrating its potential to continuously and comprehensively monitor survey accuracy at a low cost. Our results reveal important insights

about survey error in the CPS ASEC, one of the most important household surveys for both research and official statistics.

Key findings include that the survey misses a substantial fraction of both dollars and recipients for nearly all income sources we examine. This understatement is driven by misreporting among respondents (measurement error), which is by far the largest error component for nearly all income sources. Drilling deeper, we show that failure to report any receipt by true recipients (false negatives) constitutes the main source of measurement error. Both ASEC and item non-response error also cause consistently negative bias in dollars and recipients, but the net bias is much smaller especially for item non-response error. However, our analyses of net error show that both types of non-response add substantial noise to the data. In terms of dynamics, net error has tended to evolve ambiguously for the income sources we examine, but this ambiguous trend an increase in absolute error for all sources due to an increase in both false positive and false negative rates. Measurement error has tended to remain the dominant source of bias for income sources over time, but increased for some income sources and decreased for others both in terms of net and absolute error. While overall trends differ between income sources and across net and absolute error, decomposing error dynamics shows an unambiguous (and perhaps most dramatic) increase in net and absolute ASEC and item non-response error. These error components have increased both because the respective non-response rates have increased and because error among imputations has grown over time. Thus, both non-response problems have not only become more prevalent, but also more severe. On the positive side, without ASEC non-response error survey coverage of the recipients of our ten income sources appears to be good.

Our findings have several direct implications. For the general public and consumers of survey statistics, the severe understatement of program receipt and dollars received suggests caution when 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. The bias in dynamic analyses is less clear, but they could be de-biased using estimates of the error dynamics. However, the rise in absolute error we document likely translates into more complicated biases in multivariate estimates, such as take-up models. Overall, when working with the error-ridden data, survey users should consider correcting their estimates upwards or imputing additional receipt in the survey data (see e.g. Mittag 2019). Our finding that the imputed ASEC

non-respondents add substantial noise to the data suggests that survey users may be better off excluding them and re-weighting the data, especially in multivariate models.

For survey producers, our finding that extensive margin underreporting causes surveys to understate both receipt and dollars suggests that improving reporting by true recipients is the most promising margin to improve survey accuracy. Probing more among those who are likely recipients, e.g. those who report receipt of other programs or live in areas of high program receipt, may help to improve this dominant problem of underreporting. Survey producers could monitor both net and absolute error to ensure that such survey design changes do not lead to more overreporting, but indeed increase reporting by true recipients. Reducing underreporting is an especially promising strategy to improve survey accuracy, since it would also improve the imputations to which the net understatement extends because they are based on reports. Therefore, reducing underreporting would also reduce a second important source of bias – non-response error.

More generally, our analyses and their practical implications highlight the need for a more universal and comparable system of measuring survey accuracy as well as the potential of our approach to address this problem. Beyond the specific findings of this paper, our framework could help survey producers continuously monitor and evaluate survey data. Doing so would allow them to test effects of survey design changes ex ante or to evaluate them ex post. Being able to estimate both net and absolute error is crucial in this context, since it allows them to distinguish between changes that reduce bias by adding a new, offsetting source of bias and changes that actually improve survey data. A measure of survey error that is comparable across time is thus essential for survey producers to improve survey accuracy efficiently. Such efforts are further supported by the fact that our measure is comparable across components, which allows survey producers to quantify and navigate the potential trade-off between increasing response rates and reducing response accuracy. Contrary to improvements of reporting, which has positive spillovers on item non-response error, increasing response rates may have negative side effects. The marginal respondents such measures add have been argued (e.g. Groves and Peytcheva 2008, Peytcheva 2012) to only modestly improve survey representativeness, which is in line with our finding of good coverage, but to increase reporting error, because marginal respondents may report inaccurately. Our approach could be used to evaluate this trade-off empirically by quantifying the effect of efforts to increase response rates on non-response error (rather than just rates) as well as whether and how much the measures increase response error. Survey users could use our approach to evaluate the

trade-offs between excluding non-respondents and using them with their imputed values and to evaluate which corrections for the various error sources improve the accuracy of their estimates.

There are various extensions of our approach that could help both survey users and producers adapt it to their setting. 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 re-weighting improves the problem. For survey users interested in multivariate analyses, it may be useful to amend our measure of noise with a measure of the degree to which survey error is related to their covariates. Likely most importantly, our approach can be applied in situations where it is more difficult to obtain the measures of truth that our linked data provide. Our analyses of housing assistance show that our approach can still deliver useful insights in cases when the administrative data are incomplete. The approach can also be applied to cases where the administrative data contains error. Net error components can still be estimated as long as this error cancels in the required aggregates. Even if it does not cancel in the aggregates, the administrative data may often be sufficiently more accurate than the survey data such that it provides a good approximation for the error components. In other cases, such as wages or self-employment income, both data sources likely contain substantial error and our approach could still help to measure what is missing from each data source.

Yet our results in Section 4 show that even in the absence of any linked data, our approach can be used to estimate total survey error and compare it over time. And future research could further explore whether and when common measures such as response rates track our measures of bias well. Understanding this relation could help survey users and producers to use these measures as proxies to gauge bias and its components when our measures are not (yet) available.

In summary, surveys are used to measure a wide range of objects beyond the income sources we analyze here. These income sources are an ideal first application of our approach, because accurate linked data simplifies creating the measures of truth our approach requires. Survey error has also been shown to be severe for these variables. However, we do not know how the extent and sources of survey error in income sources compare to other variables commonly measured with survey data, such as employment rates, predicted election outcomes or opinions. To understand how reliable studies and statistics that use these variables are and whether survey error is better or worse than for income, we need to apply a systematic measurement approach such as ours to these cases as well. The crucial step that makes our approach feasible is the

construction of measures of truth for key survey quantities, such as the survey target. While these measures of truth are more difficult to construct in the absence of linked data, they can be obtained or approximated in other ways. In that regard, our approach can be applied beyond the ideal case we study here and provide the much needed measures of survey accuracy more generally. In conclusion, surveys are one of our central instruments of measurement, but the results in this paper clearly show that we need to measure and calibrate the accuracy of this instrument better. The tools we develop and apply in this paper provide a series of methods for doing so.

References

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

Bee, C. Adam, Bruce D. Meyer, and Graton Gathright. 2015. "Unit Non-Response Bias in the Current Population Survey." Paper presented at the Allied Social Science Associations Annual Meeting, Boston, MA, January 3-5.

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): 2143-85.

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

Celhay, Pablo A., Bruce D. Meyer, and Nikolas Mittag. 2021. "Errors in Reporting and Imputation of Government Benefits and Their Implications." NBER Working Paper No. 29184. Cambridge, MA: National Bureau of Economic Research.

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.

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.

Greene, William H. 2018. *Econometric Analysis*. 8th ed. Upper Saddle River, NJ: Pearson.

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.

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." Paper presented at the Society of Government Economists Session at the Allied Social Science Associations Annual Meeting, Boston, MA, January 3-5.

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

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

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." Paper presented at the Federal Committee on Statistical Methodology Research Conference, Arlington, VA, November 5-7.

Mulry, Mary H., and Bruce D. Spencer. 2001. "Accuracy, Coverage, and Content Evaluation: A.C.E. Revision II Results." Report prepared for the Executive Steering Committee for A.C.E. Policy. Washington, DC: U.S. Census Bureau.

O'Hare, William, Peter Miskura, David Armstrong, and Jennifer Ortman. 2016. "Children Under Age 5 in the Census." In *The Undercount of Young Children in the U.S. Decennial Census*, edited by Barbara A. Bentley, Patricia J. Duff, and Enid J. Graber, 89-116. New York: Springer.

Peytcheva, Emilia. 2012. "Consequences of Survey Nonresponse." PhD diss., University of Michigan.

Renwick, Trudi. 2017. "Addressing Quality Issues with the Current Housing Unit Value Imputation Procedure in the CPS ASEC." SEHSD Working Paper 2017-22. Washington, DC: U.S. Census Bureau.

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, Housing and Household Economic Statistics Division.

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.

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.

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. 2021. "Imputation of Unreported Data Items." Current Population Survey, Technical Documentation. Washington, DC: U.S. Census Bureau. https://www.census.gov/programs-surveys/cps/technical-documentation/methodology/imputation-of-unreported-data-items.html.

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.

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**
Retirement Income	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 Dollars and Recipients that Should Be Covered by the Survey Sample, 2017

	Retirement Income	OASDI	OASI	DI	SSI	Housing Assistance	UI	VA Disability	SNAP	Medicaid
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Dollars ($ millions)										
Payments in Admin Microdata	1,203,000	941,554	798,728	142,826	54,509	42,420	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%	1.01%	0.23%	1.62%	0.63%	--
Estimated Payments to GQs	0.68%	2.35%	2.35%	2.35%	4.46%	3.77%	0.81%	1.52%	6.84%	--
Subtraction of Payments Not Covered by Linked Data (%)										
Unlinkable Admin Records	0.80%	--	--	--	--	0.16%	0.04%	0.05%	0.03%	--
Estimated Payments to Out-of-State Residents	--	--	--	--	--	--	--	--	1.22%	--
Total Subtraction	3.43%	4.39%	4.63%	3.10%	5.17%	4.94%	1.08%	3.18%	8.73%	--
Total Payments Covered	1,161,754	900,203	761,744	138,405	51,691	40,322	31,190	69,865	6,568	--
B. Recipients (millions)										
Recipients in Admin Microdata	53.27	66.09	54.36	11.73	9.00	5.40	7.54	4.78	3.44	89.84
Subtraction of Recipients Not Covered by Household Samples (%)										
Decedents	2.81%	1.95%	2.20%	0.71%	0.78%	1.83%	0.26%	1.44%	1.35%	1.25%
Estimated Recipients in GQs	0.95%	3.01%	3.01%	3.01%	3.01%	3.61%	0.91%	1.61%	7.21%	2.94%
Subtraction of Payments Not Covered by Linked Data (%)										
Unlinkable Admin Records	1.33%	--	--	--	--	0.18%	0.06%	0.04%	0.10%	2.31%
Estimated Residents of Other States	--	--	--	--	--	--	--	--	1.01%	--
Total Subtraction	5.09%	4.96%	5.21%	3.72%	3.79%	5.62%	1.23%	3.10%	9.67%	6.49%
Total Recipients Covered	50.56	62.82	51.53	11.30	8.66	5.10	7.45	4.63	3.11	84.01

Notes: This table shows aggregate dollars and recipients 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. False Negative and False Positive Rates for Income Sources, 2017

	Retirement Income	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 Whole Imputations	74.7	31.5	28.9	83.7	81.1	95.6	93.7	77.4	67.8	85.6
For Item Imputations	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 Whole Imputations	4.8	4.9	4.0	1.9	1.9	1.2	0.9	7.3	14.7	2.2
For Item Imputations	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 4. Total Survey Error for Dollar Amounts and Its Components (as a % of Survey Targets), 2017

	Retirement Income	OASDI	OASI	DI	SSI	UI	VA Disability	SNAP	Housing Assistance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Coverage Error	5.9	-0.4	0.0	-2.5	-17.0	-4.5	-2.0	-6.2	15.0
B. ASEC Non-Response Error									
False Positives	3.5	3.1	3.3	8.6	14.9	8.0	9.0	4.8	4.2
False Negatives	-11.0	-4.7	-4.3	-13.3	-11.9	-17.6	-17.6	-19.1	-16.8
Net Error due to Mis-Imputed Receipt Status	-7.5	-1.6	-1.0	-4.7	3.0	-9.6	-8.7	-14.3	-12.6
Absolute Error due to Mis-Imputed Receipt Status	14.5	7.7	7.5	21.9	26.8	25.7	26.6	23.9	21.0
Net Error in Imputed Amounts	-0.4	0.2	0.2	-0.2	0.6	0.3	-0.1	-1.0	-1.8
Absolute Error in Imputed Amounts	4.9	5.0	5.1	1.3	1.8	0.9	1.4	2.6	2.0
Total Net Error Due to ASEC Non-Response	-7.9	-1.4	-0.8	-5.0	3.5	-9.3	-8.8	-15.3	-14.4
Total Absolute Error Due to ASEC Non-Response	19.4	12.7	12.6	23.2	28.6	26.6	28.1	26.5	23.0
C. Item Non-Response Error									
False Positives	1.1	0.5	0.5	1.6	1.9	0.8	2.2	--	1.6
False Negatives	-4.3	-0.8	-0.7	-2.7	-3.0	-2.2	-3.6	--	-5.8
Net Error due to Mis-Imputed Receipt Status	-3.2	-0.3	-0.2	-1.1	-1.1	-1.4	-1.4	--	-4.3
Absolute Error due to Mis-Imputed Receipt Status	5.4	1.3	1.2	4.3	4.8	2.9	5.9	--	7.4
Net Error in Imputed Amounts	-0.5	0.0	-0.1	0.0	-0.1	--	--	--	-1.7
Absolute Error in Imputed Amounts	3.0	1.1	1.2	0.3	0.6	--	--	--	1.8
Total Net Error Due to Item Non-Response	-3.8	-0.4	-0.2	-1.1	-1.2	-1.4	-1.4	--	-6.0
Total Absolute Error Due to Item Non-Response	8.4	2.4	2.4	4.5	5.4	2.9	5.9	--	9.2
D. Measurement Error									
False Positives	4.1	2.4	3.7	17.5	36.1	7.4	1.7	2.5	12.7
False Negatives	-31.7	-9.7	-9.1	-35.5	-23.3	-32.8	-27.1	-21.6	-21.3
Net Error due to Mis-Reported Receipt Status	-27.6	-7.3	-5.4	-18.0	12.8	-25.4	-25.3	-19.0	-8.7
Absolute Error due to Mis-Reported Receipt Status	35.8	12.1	12.8	52.9	59.4	40.1	28.8	24.1	34.0
Net Error in Reported Amounts	-10.6	-0.1	0.0	-0.3	3.8	-1.8	-0.5	-6.3	-42.2
Absolute Error in Reported Amounts	26.7	16.0	15.9	10.2	13.7	18.7	20.9	14.2	45.1
Total Net Error Due to Measurement Error	-38.2	-7.4	-5.4	-18.3	16.6	-27.2	-25.9	-25.3	-50.9
Total Absolute Error Due to Measurement Error	62.5	28.1	28.7	63.1	73.0	58.9	49.7	38.3	79.2
Total Net Error	-44.0	-9.6	-6.4	-26.9	1.9	-42.3	-38.0	-46.9	-56.3
Total Absolute Error	96.2	43.6	43.7	93.3	124.0	92.9	85.6	71.1	126.3

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 (whole imputes, item imputes, 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. Corresponding absolute error rows report the same quantities using absolute values.

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

	Retirement Income	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	-5.5	-3.9	-12.9	-17.9	-4.5	-3.2	-25.1	-3.7	7.1
B. ASEC Non-Response Error										
False Positives	3.9	3.1	3.3	8.1	10.4	7.5	9.3	5.9	7.7	9.8
False Negatives	-12.9	-4.8	-4.4	-12.1	-11.4	-18.0	-17.3	-14.2	-12.5	-15.2
Total Net Error Due to ASEC Non-Response	-9.0	-1.7	-1.2	-4.0	-0.9	-10.4	-8.1	-8.3	-4.7	-5.5
Total Absolute Error Due to ASEC Non-Response	16.8	7.9	7.7	20.2	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	-0.7	-2.5	-2.7	-2.5	-2.8	--	-5.3	-5.1
Total Net Error Due to Item Non-Response	-2.8	-0.3	-0.2	-1.0	-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	2.9	4.1	18.6	21.4	6.2	2.0	2.6	3.5	32.9
False Negatives	-42.1	-11.4	-11.1	-33.8	-26.5	-40.1	-30.1	-20.5	-26.3	-21.9
Total Net Error Due to Measurement Error	-36.3	-8.5	-7.0	-15.2	-5.1	-33.8	-28.1	-17.9	-22.8	11.0
Total Absolute Error Due to Measurement Error	47.9	14.3	15.1	52.4	47.9	46.3	32.1	23.1	29.8	54.9
Total Net Error	-45.9	-16.0	-12.2	-33.1	-25.2	-50.6	-40.0	-51.3	-30.8	11.6
Total Absolute Error	72.3	29.1	28.0	89.4	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 (whole imputes, item imputes, 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 6. False Negative and False Positive Rates for Income Sources, 1996 and 2017

	Retirement Income		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 Whole Imputations	77.5	74.7	23.9	31.5	69.6	81.1	70.5	85.6
For Item Imputations	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 Whole Imputations	5.0	4.8	5.0	4.9	1.8	1.9	2.5	2.2
For Item Imputations	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.

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

	Retirement Income		OASDI		SSI		Housing Assistance	
	1996	2017	1996	2017	1996	2017	1996	2017
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Coverage Error	-4.4	5.9	-3.7	-0.4	-12.6	-17.0	15.1	15.0
B. ASEC Non-Response Error								
False Positives	1.6	3.5	1.3	3.1	4.8	14.9	1.4	4.2
False Negatives	-4.9	-11.0	-1.4	-4.7	-3.9	-11.9	-5.9	-16.8
Net Error due to Mis-Imputed Receipt Status	-3.3	-7.5	-0.1	-1.6	0.8	3.0	-4.6	-12.6
Absolute Error due to Mis-Imputed Receipt Status	6.5	14.5	2.7	7.7	8.7	26.8	7.3	21.0
Net Error in Imputed Amounts	-0.5	-0.4	-0.3	0.2	-0.1	0.6	-1.6	-1.8
Absolute Error in Imputed Amounts	1.6	4.9	2.2	5.0	1.2	1.8	1.7	2.0
Total Net Error Due to ASEC Non-Response	-3.8	-7.9	-0.4	-1.4	0.8	3.5	-6.2	-14.4
Total Absolute Error Due to ASEC Non-Response	8.1	19.4	4.9	12.7	9.9	28.6	9.0	23.0
C. Item Non-Response Error								
False Positives	0.3	1.1	0.2	0.5	1.3	1.9	0.7	1.6
False Negatives	-0.8	-4.3	-0.2	-0.8	-1.3	-3.0	-0.9	-5.8
Net Error due to Mis-Imputed Receipt Status	-0.5	-3.2	0.0	-0.3	-0.1	-1.1	-0.1	-4.3
Absolute Error due to Mis-Imputed Receipt Status	1.0	5.4	0.4	1.3	2.6	4.8	1.6	7.4
Net Error in Imputed Amounts	0.0	-0.5	0.0	0.0	--	-0.1	-1.3	-1.7
Absolute Error in Imputed Amounts	0.5	3.0	0.3	1.1	--	0.6	1.3	1.8
Total Net Error Due to Item Non-Response	-0.5	-3.8	0.0	-0.4	-0.1	-1.2	-1.4	-6.0
Total Absolute Error Due to Item Non-Response	1.5	8.4	0.7	2.4	2.6	5.4	2.9	9.2
D. Measurement Error								
False Positives	4.4	4.1	2.9	2.4	13.9	36.1	14.0	12.7
False Negatives	-39.3	-31.7	-5.9	-9.7	-21.0	-23.3	-28.8	-21.3
Net Error due to Mis-Reported Receipt Status	-34.9	-27.6	-3.0	-7.3	-7.1	12.8	-14.8	-8.7
Absolute Error due to Mis-Reported Receipt Status	43.6	35.8	8.8	12.1	34.9	59.4	42.8	34.0
Net Error in Reported Amounts	-13.5	-10.6	-0.5	-0.1	5.8	3.8	-48.3	-42.2
Absolute Error in Reported Amounts	22.2	26.7	13.8	16.0	18.5	13.7	52.0	45.1
Total Net Error Due to Measurement Error	-48.5	-38.2	-3.5	-7.4	-1.4	16.6	-63.0	-50.9
Total Absolute Error Due to Measurement Error	65.8	62.5	22.6	28.1	53.4	73.0	94.8	79.2
Total Net Error	-57.2	-44.0	-7.6	-9.6	-13.2	1.9	-55.5	-56.3
Total Absolute Error	79.8	96.2	32.0	43.6	78.5	124.0	121.7	126.3

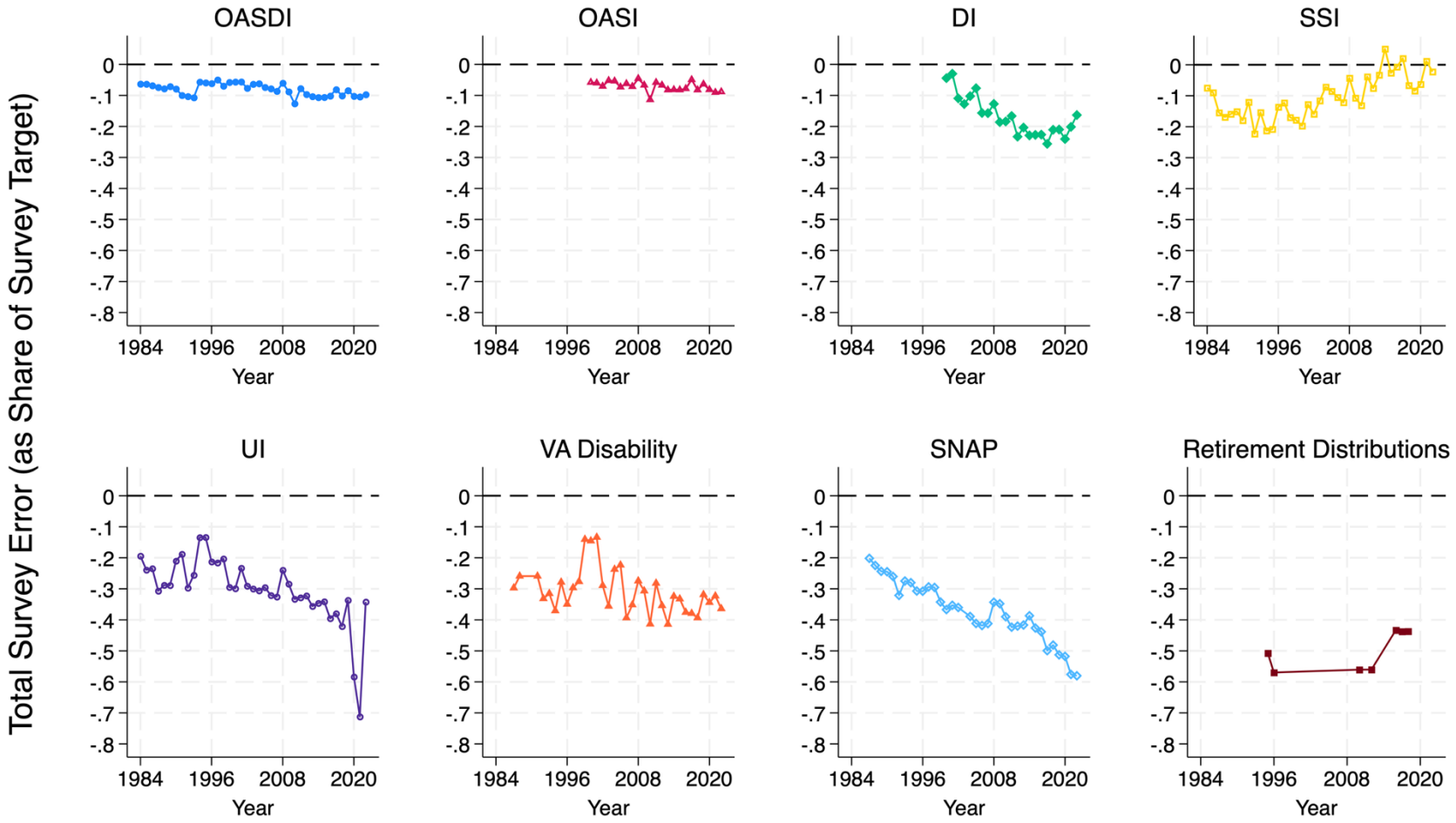
Notes: Cells marked "--" are suppressed to comply with Census Bureau disclosure avoidance rules. All reported values are expressed as shares of coverage-adjusted administrative aggregates. For each reporting group (whole imputes, item imputes, and respondents) we report false positives, false negatives, 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. Corresponding absolute error rows report the same quantities using absolute values.

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

	Retirement Income		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	-10.9	-5.5	-13.7	-17.9	3.4	7.1
B. ASEC Non-Response Error								
False Positives	2.1	3.9	1.5	3.1	4.3	10.4	4.5	9.8
False Negatives	-5.6	-12.9	-1.5	-4.8	-4.0	-11.4	-5.3	-15.2
Total Net Error Due to ASEC Non-Response	-3.5	-9.0	0.0	-1.7	0.3	-0.9	-0.8	-5.5
Total Absolute Error Due to ASEC Non-Response	7.7	16.8	3.0	7.9	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	-0.2	-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	3.3	2.9	13.1	21.4	38.9	32.9
False Negatives	-49.1	-42.1	-7.4	-11.4	-22.4	-26.5	-27.8	-21.9
Total Net Error Due to Measurement Error	-42.9	-36.3	-4.0	-8.5	-9.3	-5.1	11.1	11.0
Total Absolute Error Due to Measurement Error	55.3	47.9	10.7	14.3	35.5	47.9	66.8	54.9
Total Net Error	-47.1	-45.9	-15.0	-16.0	-22.8	-25.2	15.7	11.6
Total Absolute Error	64.7	72.3	25.1	29.1	60.1	91.7	83.8	96.2

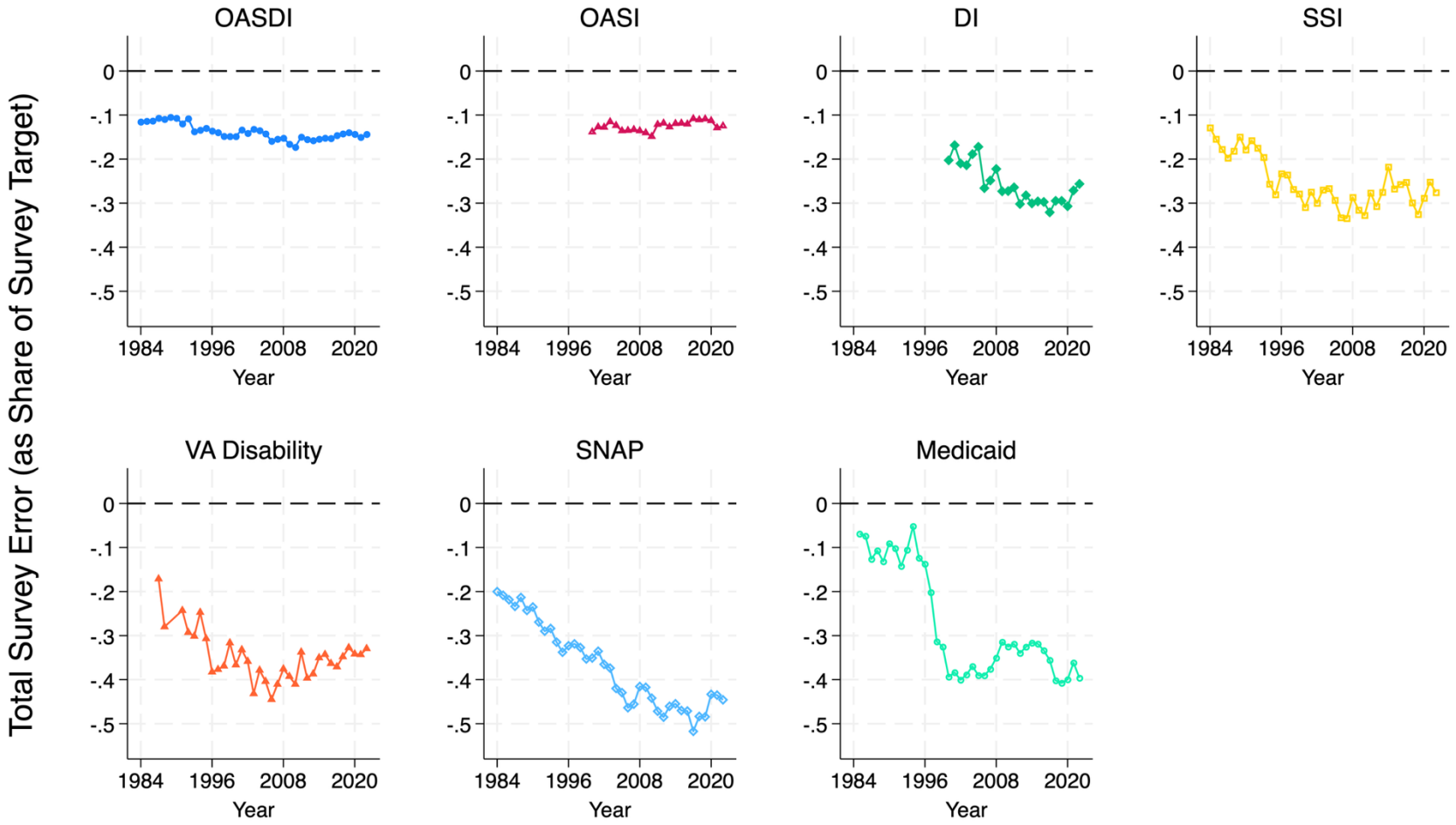
Notes: All reported values are expressed as shares of coverage-adjusted administrative aggregates. For each reporting group (whole imputes, item imputes, and respondents), we report false positives, false negatives, and their sum, defined as net error due to mis-imputed or mis-reported receipt status,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

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



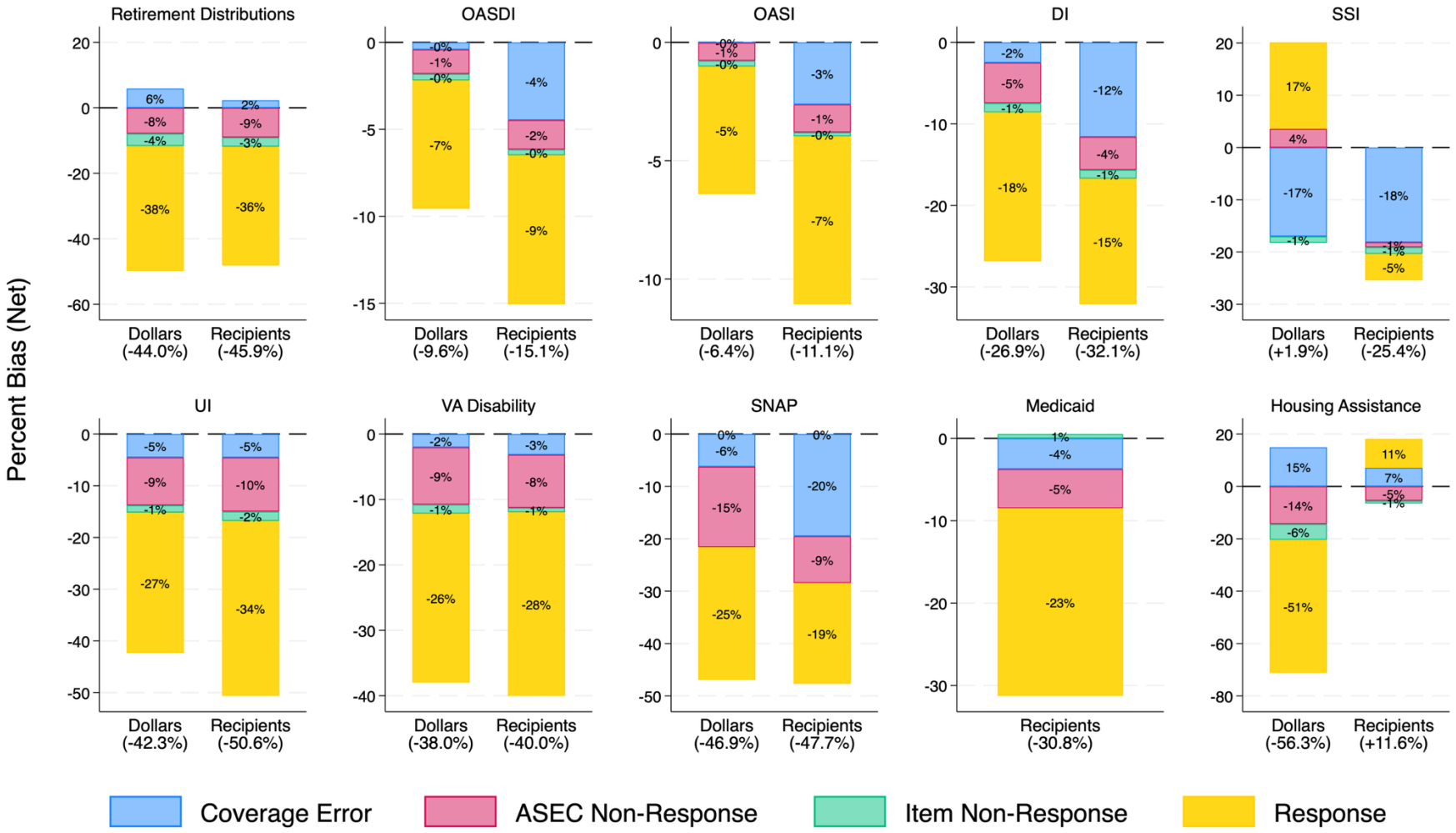
Notes: These figures show total net survey error for annual dollars in the CPS ASEC over time. Net TSE here is calculated as the difference between the weighted survey estimate (based on the public-use CPS ASEC) and the survey target constructed from publicly available administrative aggregates (after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

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



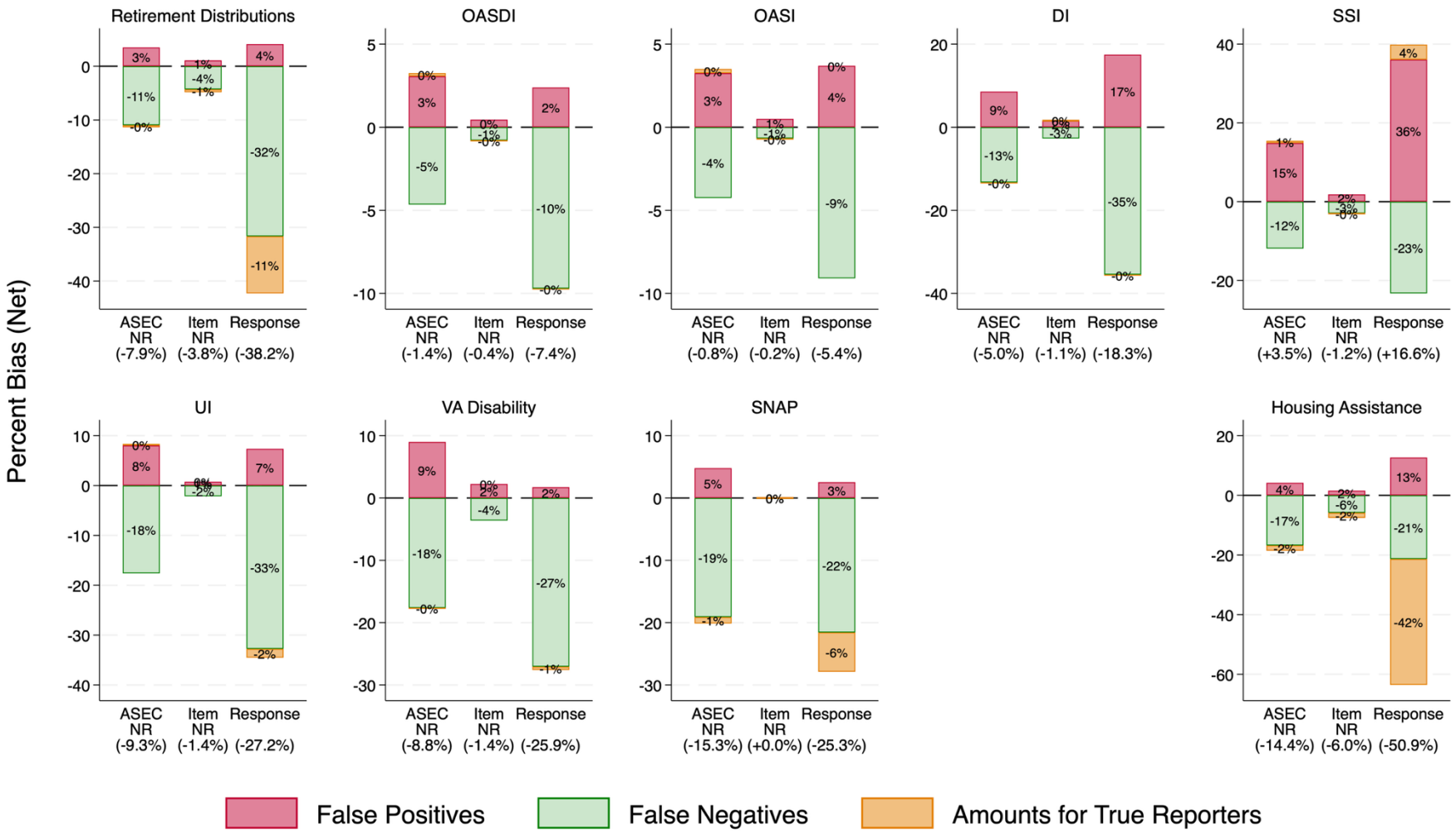
Notes: These figures show total net survey error for annual recipients in the CPS ASEC over time. Net TSE here is calculated as the difference between the weighted survey estimate (based on the public-use CPS ASEC) and the survey target constructed from publicly available administrative aggregates (after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

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



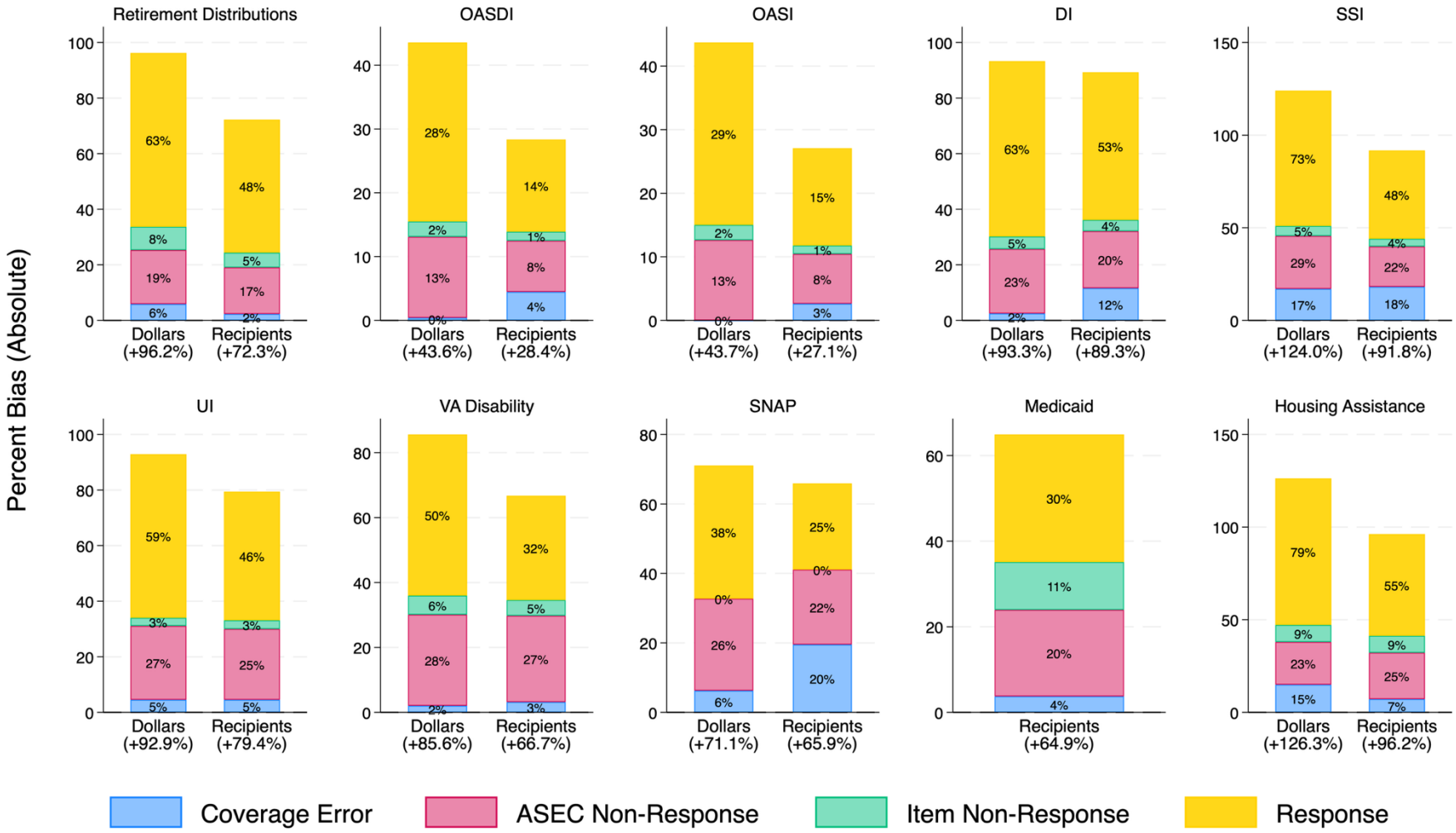
Notes: These figures decompose total net survey error for dollars and recipients into separate components for ASEC non-response, item non-response, measurement error among respondents, and coverage error (residual category) in 2017. Errors are reported as a share of the survey target. We examine both dollars and recipients 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



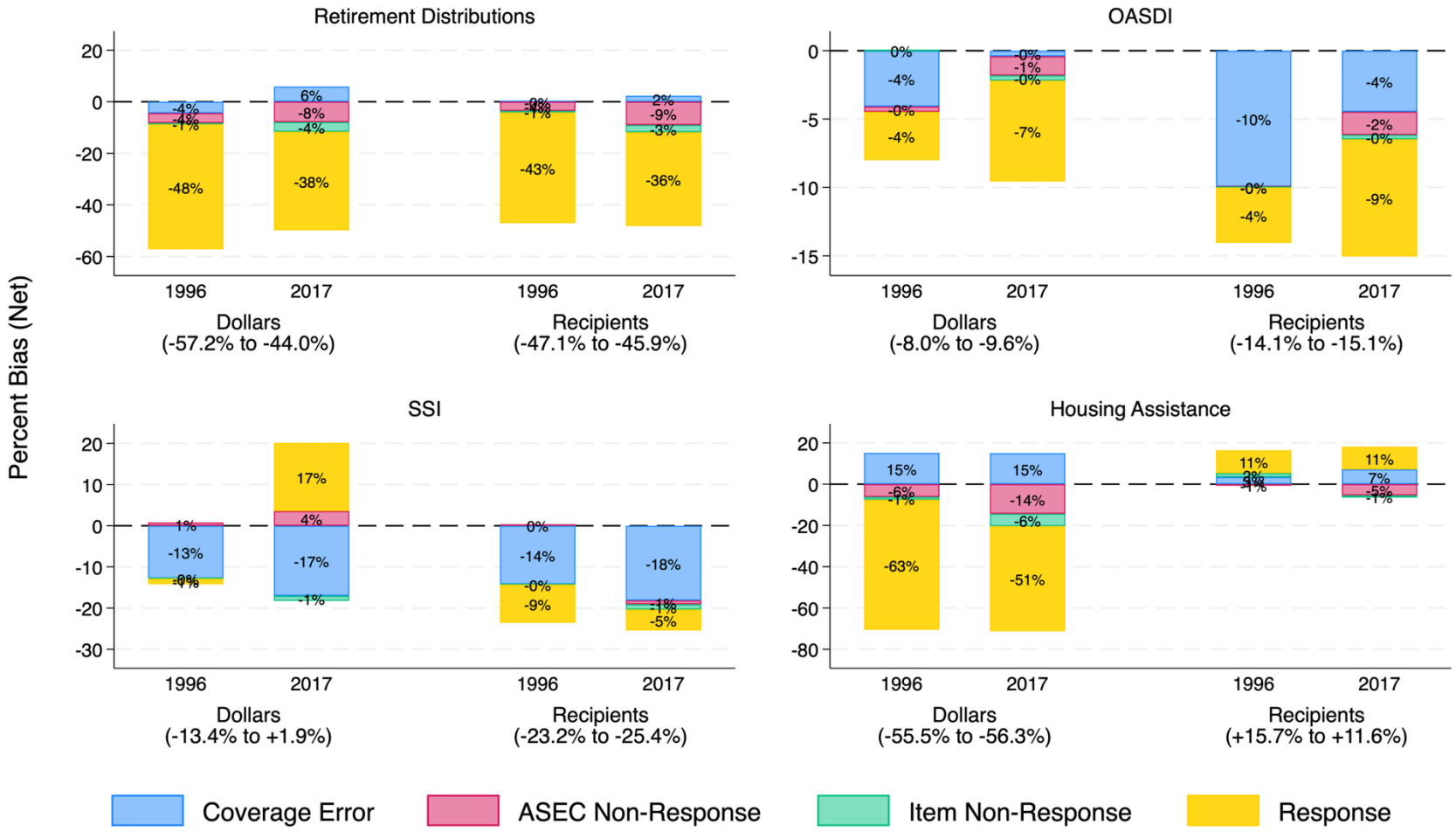
Notes: These figures decompose total net survey error for dollars (separately by ASEC non-response, item non-response, and measurement error among respondents) into extensive-margin components (false positives and false negatives) and intensive-margin errors in amounts among true reporting recipients 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 Dollars and Recipients, 2017



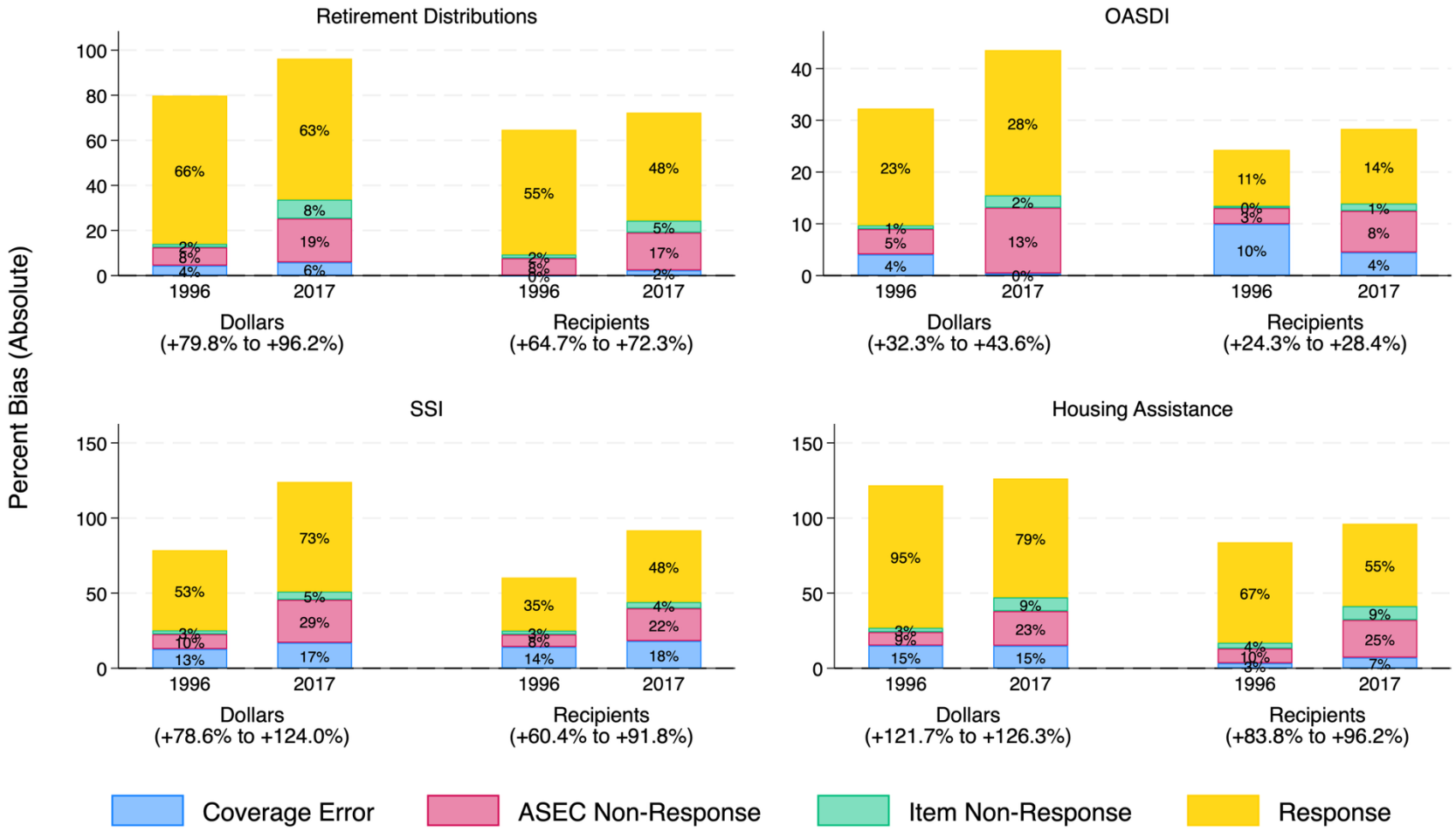
Notes: These figures decompose total absolute survey error for dollars and recipients into separate components for ASEC non-response, item non-response, measurement error among respondents, and coverage error (residual category) in 2017. Errors are reported as a share of the survey target. We examine both dollars and recipients 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 Dollars and Recipients, 1996 and 2017



Notes: These figures decompose total net survey error for dollars and recipients into separate components for ASEC 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 Dollars and Recipients, 1996 and 2017



Notes: These figures decompose total absolute survey error for dollars and recipients into separate components for ASEC 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.

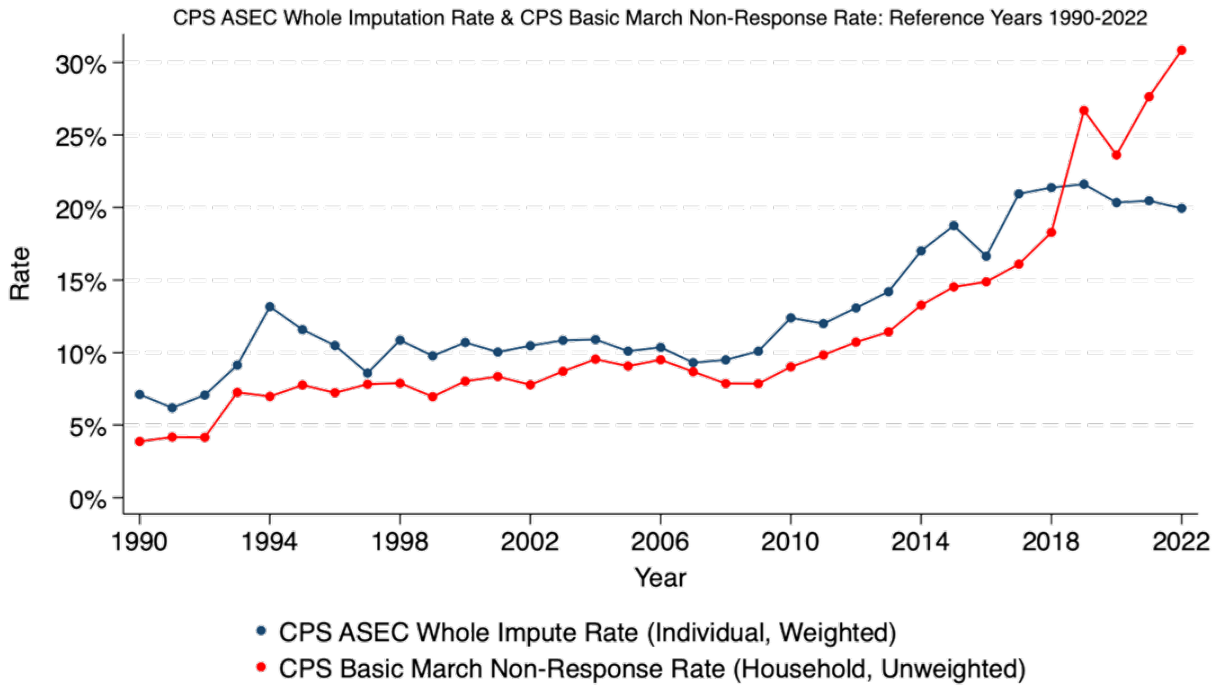
Appendix Tables and Figures

Table A.1. Misreporting Rates Among Respondents from Prior Studies

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%

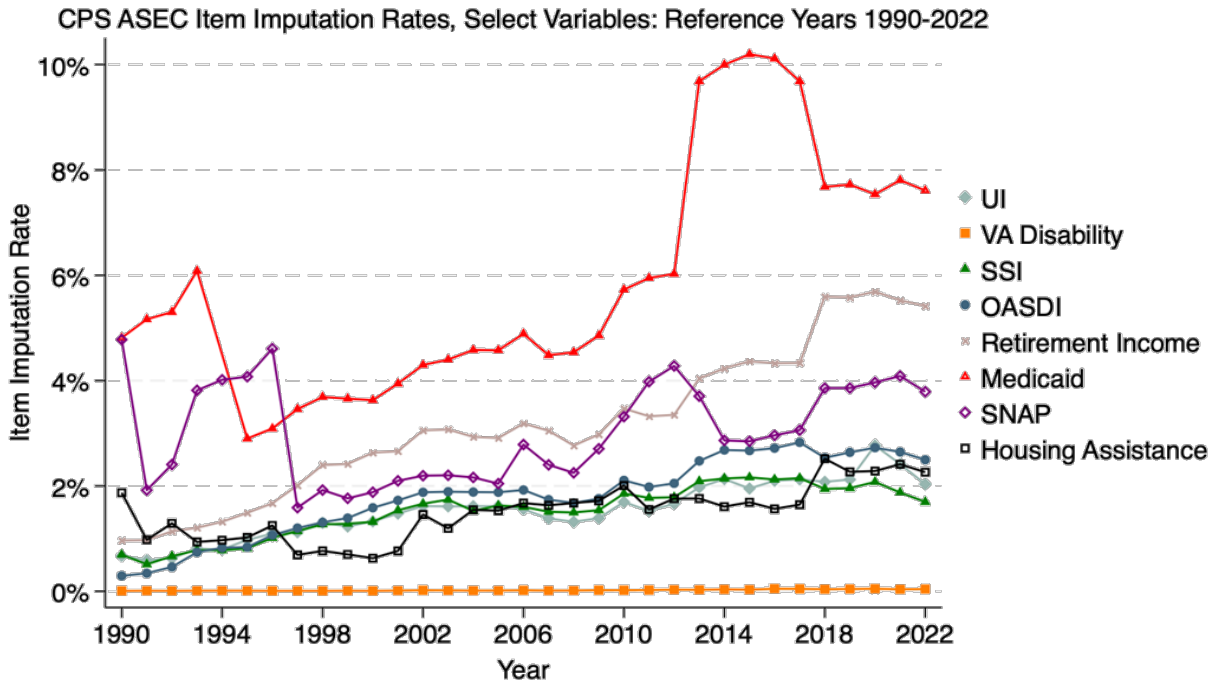
Notes: This table contains false negative and false positive rates for various programs, surveys, samples, and years calculated based on linked survey-administrative data from prior studies.

Figure A.1. 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 ASEC 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.2. 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.[24] 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.

Retirement Income

We define retirement income 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 retirement income recipient if they answer "yes" to any receipt question. Total retirement income equals the sum across all retirement-related amount variables.

CPS ASEC Years	Variable Names	Questionnaire
1985-2019	*ret_yn1/2*	Other than Social Security or VA benefits, did you receive any pension or retirement income in [year]?
	ret_val1/2	How much did you receive from this income source in [year]?
1985-2023	*sur_yn1/2*	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?
	sur_val1/2	How much did you receive from your survivor benefits in [year]?

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

1985-2023	*dis_yn1/2*	Did you receive any income in [year] as a result of your health problem other than Social Security Disability or VA Benefits?
	dis_val1/2	How much did you receive before deductions in these payments in [year]?
2019-2023	*pen_yn1/2*	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.
	pen_val1/2	How much did you receive in these pensions in [year]?
2019-2023	*dst_yn1/2* *dst_yn1/2_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?
	dst_val1/2 *dst_val1/2_yng*	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 Nutritional 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, then use HUD program rules to impute amounts based on the household income reported in the survey.

A.2. 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.[25]

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.

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

[25] 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.

tables report calendar-year totals of federally-administered SSI payments.[26] 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:

$$M_{1996} = \frac{\text{1996 Total Annual Recipients}}{\text{1996 December Monthly Recipients}} = 1.074$$

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

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,

[26] 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.

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-annual conversion factor.[27] These multipliers are applied to the public SNAP recipient aggregates to construct annual household-level recipient counts comparable to the CPS ASEC.

Retirement Distributions

Aggregate retirement dollars are constructed using disclosed 1099-R totals from restricted-use data, as no comprehensive publicly available administrative source exists that reports retirement income dollars consistent with CPS ASEC concepts. While some studies have used the National Income and Product Accounts as an administrative source of retirement income dollars, these data report retirement income in terms of accruals rather than distributions (as measured in the CPS ASEC). Retirement 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). Retirement dollars for 2012 are obtained from Bee & Mitchell (2017). One caveat is that, for many 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).

[27] 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. 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.

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.[28] 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 GQs 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 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.

[28] Additionally, beginning in survey year 2018, the CPS ASEC excludes college dormitories, the largest single source of non-institutional GQ populations.

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 retirement 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.