

Addressing Nonresponse Bias in Household Surveys using Linked Administrative Data

Jonathan Rothbaum

Adam Bee

December 28, 2021

Abstract

Nonresponse has been increasing in household surveys. To address nonresponse bias, we use address-linked administrative data to identify individuals in respondent *and* nonrespondent households. We link them to income, demographic, and socioeconomic information from administrative data, prior surveys, and the decennial census. We use entropy balancing to adjust survey weights to match a high-dimensional vector of moment constraints. In the 2020 CPS ASEC, nonresponse biased income estimates up by 2 to 3 percent and poverty estimates down by about 0.45 percentage points. In other years, we do not find evidence of bias in income or poverty. We produce public-use weights that address this nonresponse bias while protecting respondent confidentiality.

Keywords: Nonresponse Bias; Income Measurement; Household Surveys

JEL Codes: C83, I3

Jonathan Rothbaum and Adam Bee are economists in the Social, Economic, and Housing Statistics Division at the U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233. This report is released to inform interested parties of ongoing research and to encourage discussion. Any views expressed on statistical, methodological, technical, or operational issues are those of the author and not necessarily those of the U.S. Census Bureau. The U.S. Census Bureau reviewed this data product for unauthorized disclosure of confidential information and approved the disclosure avoidance practices applied to this release. CBDRB-FY20-380, CBDRB-FY20-414, CBDRB-FY21-POP001-0060, and CBDRB-FY21-POP001-0238. The public-

1 Introduction

Nonresponse in household surveys has been increasing for decades, both in the United States (Williams and Brick, 2018) and around the world (Luiten, Hox and de Leeuw, 2020). If nonresponse is nonrandom, higher nonresponse may result in increased nonresponse bias. Over the same period, additional data, including administrative records, have become more available. Administrative data can help us both evaluate whether nonresponse is random and, if not, correct for nonresponse bias.

In this paper, we apply an improved method for survey weighting, entropy balancing (Hainmueller, 2012), which allows us to efficiently reweight respondents to a high-dimensional vector of moment conditions. We also incorporate additional data into our reweighting procedure. The additional information includes administrative data on income from the Internal Revenue Service (IRS) as well as linked responses from the decennial census and the American Community Survey (ACS) and administrative records from the Social Security Administration (SSA) on the race, ethnicity, gender, citizenship, and nativity of household residents. Crucially, the linked information is available for both respondent and nonrespondent households, which allows us to estimate the distribution of characteristics in the linked data for the full target population. We use entropy balancing to produce weights for public use that incorporate the value of administrative data in addressing nonresponse bias while protecting the confidentiality of the households and individuals whose administrative data was used.

With the linked data and improved weighting methods, we can better characterize and address nonresponse in surveys. We apply these methods to the Current Population Survey Annual Social and Economic Supplement (CPS ASEC).¹ We find limited evidence of non-

use weights are approved for release under approval CBDRB-FY21-126. We would like to thank David Hornick for help understanding CPS ASEC sampling and weighting and participants of the Census Research Seminar for their helpful suggestions and comments. Earlier drafts of this paper were titled “Coronavirus Infects Surveys, Too: Survey Nonresponse Bias and the Coronavirus Pandemic.”

¹The CPS is jointly sponsored the Census Bureau and the Bureau of Labor Statistics (BLS) and fielded monthly by the Census Bureau in order to track the nation’s labor force statistics, including the unemployment rate. Each year between February and April, the Census Bureau administers the ASEC by telephone

response bias in income from 2017 to 2019, but strong evidence of nonrandom nonresponse during the pandemic in 2020. In the 2020 and 2021 CPS ASECs (for income in 2019 and 2020), higher income households were considerably more likely to respond to the CPS ASEC, biasing income statistics up. With our adjusted weights, we estimate that the survey overstated household income across the distribution, including by 2 to 3 percent at the median. Likewise, estimates of official poverty were biased down by about 0.45 percentage points in both years, for an additional 1.5 million individuals in poverty compared to the official estimates.

Nonresponse bias has concerned survey sponsors throughout the development of scientific household surveys, so the literature on nonresponse bias is extensive and varied. Groves and Peytcheva (2008) survey 59 nonresponse analyses across a variety of research designs. Their meta-analysis comprises comparisons using survey frame variables, comparing responses to an earlier screener interview or other waves of the same survey, comparisons by the respondent's reported willingness to respond to a later interview, comparing respondents recruited from varying levels of field effort (e.g., rounds of follow-up or varying incentives), as well as the method we use: individually linking data from auxiliary records to sample units. They find that nonresponse bias is only weakly correlated to a given survey's response rate, and that the bias can vary widely across various estimates from the same survey.

Many analysts have also previously measured nonresponse bias in the CPS specifically. Groves and Couper (2012) match CPS sampled households to their responses in the 1990 decennial census, finding differences by demographic characteristics. John Dixon, working at BLS, has written a series of CPS nonresponse analyses. For example his 2007 paper, matching the 2006 Basic CPS to the 2000 decennial census, finds slightly less biased unemployment rates during the summer months. Research at the World Bank (e.g., Korinek, Mistiaen and Ravallion, 2006, 2007; Hlasny and Verme, 2018; Hlasny, 2020) developed an iterated method

and in-person interviews, with the majority of data collected each March. This supplemental questionnaire asks respondents about their income, health insurance status, etc. for the prior calendar year. The resulting data are heavily used in policy and academic research, and they are the source of the nation's official income and poverty estimates.

to correct for nonresponse bias based on the observed relationship of income to nonresponse across geographic areas. Heffetz and Reeves (2019) use difficult-to-reach respondents as proxies for nonrespondents.

The methods we employ in this paper follow most directly from a line of nonresponse papers developed at the U.S. Census Bureau. Extending Sabelhaus et al.'s (2015) linkage of Consumer Expenditure Survey and CPS ASEC samples to IRS ZIP-code-level income tables, Bee, Gathright and Meyer (2015) pioneered the method of linking nonrespondents of nationally representative surveys to administrative records via the Master Address File. Linking IRS Form 1040 records to the 2011 CPS ASEC, they find little selection into response across much of the unconditional income distribution, but uncover some selection on other demographic characteristics like marital status and number of children in the household. Eggleston and Westra (2020) extend the address-linking method to estimate new weights that correct for differential nonresponse. Reweighting the Wave 1 2014 SIPP respondents, they find similarly negligible biases across the income distribution.

Various applications of this method have found varying patterns of nonresponse bias. Brummet et al. (2018) apply this method to the Consumer Expenditure Survey, finding that high-income households are less likely to respond. Mattingly et al. (2016) apply the method to the Wave 1 2008 Survey of Income and Program Participation (SIPP), finding no evidence of nonresponse bias. However, Eggleston (2021) uses a similar approach as we do, applied to the American Community Survey (ACS), and he finds that high-income households are more likely to respond. He also finds that differences between respondents and nonrespondents have been increasing over time.

Our method, in turn, extends the prior work along a number of dimensions. First, we link a wider set of auxiliary data, in particular using the universe of information returns to link addresses to people for both nonrespondents and respondents. Second, we link multiple survey years to track changes in nonresponse functions over time. Third, we use a more flexible reweighting mechanism, entropy balancing (Hainmueller, 2012), which allows us to

control for nonresponse bias in a large number of dimensions. Fourth, we create and publish weights for public use that incorporate the information from the protected administrative records. Rothbaum et al. (2021) extend this work to the ACS and find biases in income and poverty statistics for that survey both before and during the pandemic.

2 Data

To assess the likely severity of nonresponse bias and adjust for it, we would like information on both respondent and nonrespondent households. This has been difficult to achieve in the past, given the absence of information for nonrespondents. We use administrative data linked to the address of the surveyed housing unit, which therefore is available for all households, independent of response type.²

In Table 1, we summarize the data used. A diagram of this process is also shown in Figure 2. We start with the CPS ASEC household file to get sample frame information. From that file, we get information on household response type (respondent, Type A non-interview, and Type B and C non-interview) and the Master Address File ID (MAFID) for each housing unit in sample.³ The MAF is the comprehensive address database maintained by the Census Bureau for its survey operations. Housing units in the CPS ASEC are selected from the MAF. Administrative data sets with addresses are also linked to the MAF using

²The linking methods we exploit here were developed independently by Census Bureau researchers. Brummet (2014) describes the development and performance of the system used to link household records, via residential address fields, to the Master Address File (MAF), called the “MAF Match.” Wagner and Layne (2014) describe the Person Identification Validation System (PVS) used to assign individual PIK values for linkage. PIKs are assigned by a probabilistic matching algorithm that compares characteristics of records in administrative and survey data to characteristics of records in a reference file constructed from the Social Security Administration (SSA) Numerical Identification System (or Numident) as well as other federal administrative data. These characteristics may include Social Security Number (SSN), full name, date of birth, address, place of birth, and parents’ names depending on the information available in the data source. The PIK uniquely identifies a particular person and is consistent for that person over time. PIKs correspond one-to-one with a particular SSN. Consequently, the PIK allows us to link individuals across data sources. In administrative data with SSNs, that one-to-one mapping can be used to easily assign PIKs to individuals. See Wagner and Layne (2014) for more information on the assignment of PIKs to survey and administrative data.

³Type A non-interview housing units are nonrespondents. Type B non-interviews are vacant units. Type C non-interviews are non-residential addresses and are thus also ineligible for inclusion the survey.

probabilistic linking on the address string. As a result, the MAFID can be used to link addresses across data sets.

We use the MAFID to link survey households to the 1099 Information Return Master File (IRMF). This file contains data on information returns filed on behalf of individuals, including for Forms W-2, 1098, 1099-DIV, 1099-G, 1099-INT, 1099-MISC, 1099-R, 1099-S, and SSA-1099. There is no income information on this file, as it only includes flags indicating which forms were filed. The file contains address information, including the corresponding MAFIDs, which we use to link it to the sample frame information. It also contains Protected Identification Keys (PIKs) for the individuals that received the information returns.

These PIKs enable all further links to other administrative and survey information. The PIKs do not necessarily identify all residents of a given housing unit, just those that received information returns. However, this roster of individuals is available for responding and nonresponding housing units. It does not necessarily correspond to the set of individuals we observed or would have observed living in the housing unit in the CPS ASEC.

We use these PIKs to get income information from the W-2 Master File and the 1099-R Information Return Master File. The W-2 files include taxable wage and salary earnings and deferred compensation amounts for all W-2 covered jobs. The 1099-R files include income amounts from pension plans and withdrawals from defined-contribution retirement plans (such as 401(k)s) as well as income from survivor and disability pension plans, but excluding rollovers. For both files, the income covered matches the CPS ASEC reference period. We use only those forms posted to IRS databases by week 19 of the CPS ASEC calendar year, to match the data availability for 2020 during regular CPS ASEC production.⁴

Next, we link the PIKs to the 1040 Returns Master File from the prior calendar year. Due to the pandemic, the 2020 tax filing deadline was extended to July 15. We do not use 1040s filed in the calendar year of the CPS ASEC survey as we are concerned about non-random selection of households into early filing, particularly in 2020, which might affect comparisons

⁴Week 19 ended May 10, 2020, and May 12, 2019. W-2s are due to the IRS by January 31 each year. 1099-R filings are due to the IRS by March 31.

to prior years.⁵ Instead, for each CPS ASEC year, we use 1040s filed by the linked individuals in the prior calendar year for income from the year before the CPS ASEC reference period. For example, for the 2020 CPS ASEC, individuals report income for 2019 in the survey, but the linked 1040 filed in 2019 covers income from 2018. Although this income is not from the CPS ASEC reference period, it does provide information on the characteristics of responding and nonresponding households. For tax filers, the 1040 file contains information on adjusted gross income (AGI), wage and salary earnings, interest, dividends, gross rental income, and social security income. The 1040 also contains information on marital status (through joint filings) and PIKs for up to four dependents.

We also use the PIKs to link to several other sources of demographic and socioeconomic information. From the Social Security Administration’s (SSA) Numident file, we get information on each individual’s age, gender, and citizenship status.⁶ From the 2010 Decennial Census short form file, we get information on age, gender, race, and Hispanic origin. From the American Community Survey (ACS), we get information on an individual’s educational attainment if that individual was surveyed in any ACS from 2001 to 2018.

3 Weighting for Nonresponse

To correct for potential selection into response, we would like weights that condition on income and other characteristics available in the linked administrative, census, and survey data. The existing survey weights, however, condition only on the available demographic information in the survey. In this section, we first describe the existing weighting procedure for the CPS ASEC and then discuss our alternative weighting procedure, entropy balancing.

⁵Tax filing in 2020, for tax year 2019, may also have been affected by incentives around stimulus payments. For example, people who did not file a return for tax year 2018 had an incentive to file their tax year 2019 returns to receive a stimulus payment, even if they would not otherwise have been required to file.

⁶The Numident, or Numerical Identification System, contains information on all individuals that have ever filed for an SSN.

3.1 CPS ASEC Survey Weights

The CPS ASEC sample is a combination of several subsamples. The largest portion of the sample comes from the March Basic CPS. In 2019, 75 percent (71,000) of the approximately 95,000 housing units sampled for the ASEC came from the March Basic CPS sample. In addition, there is the Hispanic oversample. The Hispanic oversample draws households from the prior November's CPS sample, which were identified as Hispanic from a previous CPS response. The Hispanic oversample comprised 7 percent (6,600) of the housing units in the 2019 ASEC sample. Finally, the CPS ASEC includes additional households, primarily to improve the precision of state-level children's health insurance coverage estimates, called the SCHIP oversample.⁷ The SCHIP oversample has three components: 1) asking the ASEC Supplement questions of one-quarter of the February and April Basic CPS samples; 2) interviewing selected sample households from the preceding August, September, and October Basic CPS samples during the February-April period using the ASEC Supplement; and 3) increasing the monthly CPS sample in states with high sampling errors for uninsured children. The SCHIP oversample comprises 18 percent (17,000) of the housing units in the ASEC sample.

Each subsample is selected separately, and each household has a base weight defined by the probability of selection into that subsample. The final CPS ASEC person weights are estimated as follows:

1. Set the initial subsample base weight to account for the probability of selection into each sample group,
2. Make any needed special weighting adjustments (for selection into the main sample or each oversample),
3. Adjust for differential nonresponse of those inside and outside of Metropolitan Statistical Areas,

⁷SCHIP, for the State Children's Health Insurance Program.

4. Apply a two-stage coverage procedure (national-level and state-level coverage ratios) and a three-step iterative raking procedure to match to external estimates of state population totals by age and sex; to race population totals by age and sex; and to Hispanic origin population totals by age and sex. This raking also includes a step where the weights of spouses are equalized, with any necessary additional adjustments made to unmarried men and women to match the population totals after spousal equalization.⁸

The person weight for the “householder” is the supplement household weight.⁹

Step (4) in the weighting process simultaneously adjusts weights for differential nonresponse across age, sex, race, and Hispanic origin and accounts for oversampling of various demographic groups as part of the Hispanic and SCHIP oversamples.¹⁰ This step is not amenable to adjustment for differential nonresponse by many additional characteristics, such as various measures of income, education, citizenship, etc., that are used in this paper.¹¹

3.2 Entropy Balance Weights

To correct for nonrandom nonresponse we create weights using entropy balancing (Hainmueller, 2012) that condition on characteristics that are not observable in the survey. Entropy balancing is an application of exponential empirical calibration (Deville and Särndal, 1992), which has a long history of use in survey weighting. We use the unobservable in-

⁸For a more complete description, see the technical documentation at <https://www2.census.gov/programs-surveys/cps/methodology/CPS-Tech-Paper-77.pdf> and <https://www.census.gov/programs-surveys/cps/technical-documentation/methodology/weighting.html>.

⁹The householder is the person (or one of the people) in whose name the home is owned or rented. If a married couple owns the home jointly, either spouse may be listed as the householder, depending on who responded to the survey.

¹⁰The base weights account for the probability of selection into each sample group: the March Basic CPS sample, the Hispanic oversample, and the SCHIP oversample. Without differential nonresponse by demographic group, the adjustment in (4) will decrease the weight on Hispanic individuals in the March Basic CPS, for example, to adjust for the additional individuals present in the Hispanic oversample. However, if Hispanic individuals are also more or less likely than non-Hispanics to respond to the survey, the relative weights of the two groups in (4) will also change to control for the differential nonresponse.

¹¹The challenge is both in the higher dimensionality of the weighting adjustment in this paper and in the complicated nature of the current code.

formation (in the survey) from the linked administrative, census, and survey data, which is available for all linkable households, regardless of whether they responded. Entropy balancing estimates the set of weights that matches a specified set of moment constraints while keeping the final weights as close as possible to the initial weights.

More specifically, suppose we have n observations, where $i = 1, 2, \dots, n$ with base weights based on sampling probabilities of $q = \{q_1, q_2, \dots, q_n\}$. Entropy balancing estimates the set of weights $w = \{w_1, w_2, \dots, w_n\}$ that solve the following minimization problem:

$$\min_w \sum_{i=1}^n w_i \log\left(\frac{w_i}{q_i}\right) \quad (1)$$

subject to several sets of constraints. First, we have p moment conditions. Let $X = \{X_1, \dots, X_p\}$ be a matrix of observable characteristics. For characteristic j , the moment conditions are defined to match a vector of pre-specified constants \bar{c}_j , where:

$$\sum_{i=1}^n w_i c_j(X_{i,j}) = \bar{c}_j. \quad (2)$$

$c_j(\cdot)$ can be any arbitrary function.

Second, we have constraints on the weights themselves:

$$\begin{aligned} \sum_{i=1}^n w_i &= \bar{w} \\ w_i &\geq 0, i = 1, \dots, n \end{aligned} \quad (3)$$

which ensure that the weights sum to some pre-specified total weight \bar{w} , which can be the population count or 1. The value of \bar{w} does not affect the relative weights of each observation.

As such the weights can be adjusted to match pre-specified moments such as population means, variances, higher-order moments, moments of any transformed distribution of $X_{i,j}$, etc. In summary, entropy balancing adjusts the weights according to (1), subject to the

constraints in (2) and (3).¹²

Entropy balancing was developed as a method to balance covariates across treatment and control groups when estimating causal treatment effects in observational studies. Zhao and Percival (2017) show that, in that context, entropy balancing is equivalent to estimating a logistic model for the propensity score and a linear regression model for the outcome, conditional on the covariates used in the moment conditions. They find that entropy balancing is doubly robust—if at least one of the two models is correctly specified, the estimated population average treatment effect on the treated (PATT) is consistent. Using the notation of that literature, let γ be the PATT, Y be an outcome of interest where $Y(1)$ is the outcome if treated and $Y(0)$ is the outcome if untreated, then:

$$\gamma = E[Y(1)|T = 1] - E[Y(0)|T = 1]. \quad (4)$$

In the causal inference literature, the challenge is that $E[Y(0)|T = 1]$ is not observed. Under entropy balancing, given $\sum_{i=1}^n q_i = \bar{q}$, the PATT is estimated as:

$$\hat{\gamma}_{ebw} = \frac{1}{\bar{q}} \sum_{T_i=1} q_i Y_i - \frac{1}{\bar{w}} \sum_{T_i=0} w_i Y_i. \quad (5)$$

In the case of survey weights, the “treatment” is nonresponse, and the double robustness result applies. Entropy balancing reweights the sample so that the estimate of Y for the weighted respondents is equal to the estimate of Y for the population,¹³ or:

$$E[Y] = \frac{1}{\bar{w}} \sum_{i=1}^n w_i Y. \quad (6)$$

Entropy balancing has several other appealing features for this application. The first is

¹²In practice, as is not necessarily possible to satisfy all constraints simultaneously through weighting adjustment, the analyst sets a tolerance level for the moment constraints. The weighting algorithm adjusts the weights iteratively until all constraints are satisfied subject to the specified tolerance.

¹³Conditional on strong ignorability ($Y(0), Y(1) \perp T|X$) and overlap ($0 < P(T = 1|X) < 1$), from Rosenbaum and Rubin (1983), as well as the proper specification of the moment conditions required for the Zhao and Percival (2017) double robustness result.

flexibility. Inverse probability weighting (or any simple regression-based reweighting technique) is only amenable to matching characteristics of the distribution in the sample, but not external targets. Entropy balancing, on the other hand, will adjust the weights to match any properly specified target moment, whether that moment constraint was estimated on the sample or with external data. The second is statistical efficiency, which is achieved by keeping the final weights as close as possible to the initial probabilities of selection through the inclusion of w_i/q_i in (1). The third is computational efficiency—entropy balancing allows matching to a high-dimensional vector of moment constraints. In our application, we use state-level population controls that include estimates of the share of the population in 20 separate groups in each of the 50 states and the District of Columbia.¹⁴ That yields 1,020 separate target population moments before we add in the target moments from the linked data. Fourth, entropy balancing directly adjusts the weights to the moment conditions, like with raking but unlike single-index propensity score reweighting approaches (such as inverse probability weights). In propensity score approaches, the adjustment is made to the single index generally estimated from a regression. The resulting balance must be assessed to evaluate the success and quality of the propensity score model. In some cases, a misspecified propensity score model can make balance worse on a given set of dimensions. As entropy balancing directly targets those moments, balance is assured.

We would like to reweight the respondent sample so that its distribution of characteristics matches the target population from which the sample was drawn. However, some characteristics are not observable for all housing units with the available linked census, survey, and administrative data. For example, we do not observe any demographic information for housing units that are not linked to an information return in the IRMF file, as the IRMF provides the identifier needed (PIK) to link each individual to all other data sources. Therefore, we use a second source of data for our reweighting—external estimates of population by

¹⁴The 20 groups are 12 estimates from 3 age groups (0-17, 18-64, 65 and over) by demographic cells (Black, White, Hispanic, and female) as well as state-level estimates of the population in 8 age groups (0-5, 6-12, 13-17, 18-24, 25-34, 35-44, 45-54, 55-64, and 65 and over, where the total is 8 because one is excluded).

geography. For both the linked data and the external population estimates, we can specify a set of moment conditions, which are intended to capture the distribution of characteristics in the target population.

Our data have one additional complication, however—the target moments are at separate levels of aggregation. Estimates from the linked administrative, survey, and census data are at the housing unit level whereas the external state-level population moments are at the individual level. Entropy balancing is not amenable to matching moments at different levels of aggregation. Therefore, we proceed with a two-stage reweighting procedure, which we discuss below and summarize in Table 6. This is analogous to two-step calibration, as discussed in Estevao and Säärandal (2006).

In the first stage, we adjust the household base weights for nonresponse, controlling to moments estimated from the linked administrative, census, and survey data. The target distribution is estimated using the nonvacant housing units in the March Basic CPS Sample, which includes both respondent and nonrespondent housing units. Given the known probability of inclusion in the sample (using the base weights), these moments are estimates of the underlying population moments for each of the included characteristics. The moments include housing-unit-level summary statistics on race, Hispanic origin, age, marital status, income, sources of income (through information return dummies), citizenship, and nativity.

Entropy balancing adjusts the housing unit weights so that the weighted estimates from respondent units matches the moments estimated from all nonvacant households. Let us designate the housing-unit moment constraint variables as $X_{i,j}^L$, where L indicates linked data. Let w_i^1 be the output weights of the first-stage reweighting. Given n respondent households, and a set of nonvacant (occupied) households O , where $i = 1, \dots, n_O$ with survey base weights q_i , the moment conditions are of the form:

$$\sum_{i=1}^n w_i^1 c_j(X_{i,j}^L) = \sum_{i=1}^{n_O} q_i^1 c_j(X_{i,j}^L). \quad (7)$$

With these moment conditions, we estimate w_i^1 for each household using entropy balancing.

In the second stage, we would like to create weights (denoted $w_{p,i}i^2$) for each individual m and household i , where $m = 1, \dots, M$, that adjust to external population controls while maintaining the household weighting adjustment from the first stage. We do so by simultaneously matching to three sets of target moments (2A-C in in Table 6):

A Preserve the distribution of housing unit characteristics

B Spousal equivalence

C External population targets

In the first set of constraints, we calculate per-person weighted moments from the stage-1 weights. Given the number of people in household i , n_i^{HH} , we define the moment condition, with the balance constraints defined using the stage-1 weights:

$$\sum_{m=1}^M w_{m,i}^2 \frac{1}{n_i^{HH}} c_j(X_{i,j}^L) = \sum_{i=1}^n w_i^1 c_j(X_{i,j}^L) \quad (8)$$

This ensures that if we take the average weight of household members in household i (HH_i) as $\bar{w}_i^2 = \frac{1}{n_i^{HH}} \sum_{p \in HH_i} w_{m,i}^2$ will satisfy the following condition:

$$\sum_{i=1}^n \bar{w}_i^2 c_j(X_{i,j}^L) = \sum_{i=1}^n w_i^1 c_j(X_{i,j}^L) \quad (9)$$

This does not require that \bar{w}_i^2 is equal to w_i^1 for any household i , just that the specified constraints from stage one hold in the final entropy-balance weights, when the final weights are averaged across all household members. This procedure of dividing the household moments equally among the family members helps ensure that each person contributes to satisfying the moments from linked decennial census, administrative, and third-party data, which should reduce the variability of weights among household members. This can be particularly important for person-level statistics, such as poverty.

For the second set of moments in the second-stage reweighting (2.B. in Table 6), we ap-

proximate the spousal equalization that is part of existing CPS ASEC weights. We include this set of conditions because family-level statistics should also be invariant to which spouse’s weight is used as the family weight. This is the family-level equivalent of balancing householder and householder-partner weights in (8). Let $S = 0, 1, 2$, where $S = 0$ if an individual is unmarried, 1 if the individual is the first spouse or cohabiting partner on the file, and 2 if the individual is the second spouse or partner on the file. Given an indicator function $I(\cdot)$, the spousal equivalence moment condition for a given characteristic in the linked data is:

$$\sum_{i=m}^M [I(S = 1)w_{i,m}^2 c_j(X_{i,k}^L) - I(S = 2)w_{i,m}^2 c_j(X_{i,k}^L)] = 0. \quad (10)$$

This does not require that each individual’s weight be equal to their partner’s, as that would require a separate moment condition for each couple. Instead it requires that the characteristics of the households of partners in the linked data be balanced.

The third set of moment conditions (2.C. in Table 6) reweight the individual observations to match the age by race/Hispanic-origin/gender cells for each state and the District of Columbia, as noted above.¹⁵ These conditions have the simple form of equation (2).

With these three sets of conditions, we reweight the March Basic CPS sample to simultaneously match the household-level linked administrative data and the individual-level state population targets. For each individual, the initial weights (q_i) for the stage 2 reweighting are the households weights from the stage 1 reweighting (w_i^1), so that the minimization from (1) becomes:

$$\min_{w^2} \sum_{i=1}^n w_i^2 \log\left(\frac{w_i^2}{w_i^1}\right). \quad (11)$$

However, for the full CPS ASEC sample, there is an additional complication. The full sample includes groups that were oversampled based on characteristics reported in earlier survey responses, including Hispanic origin and the presence of children. Therefore, in

¹⁵External estimates of population by geography can be found at <https://www.census.gov/data/tables/time-series/demo/popest/2010s-state-detail.html> (accessed 1/15/21). For this paper, because the existing CPS ASEC weights already incorporated these population totals, we estimated target moments directly from the existing survey weights.

the full sample, the weights for these oversampled individuals and households need to be adjusted to reflect their prevalence in the population. To do this, we add a fourth set of moment conditions (2.D. in Table 6). We create these conditions from the entropy-balance weighted March Basic sample, because that sample is a stratified random sample that is not affected by oversampling based on observable characteristics from prior survey responses. Let $w_{i,m}^{2, March}$ be the second-stage weights from the March Basic Sample, $w_{i,m}^{2, Full}$ be the second-stage weights from the full CPS ASEC sample, and M_{Full} and M_{March} be the number of individuals in the full and March Basic CPS samples. This fourth set of conditions has the form:

$$\sum_{m=1}^{m_{Full}} w_{i,m}^{2, Full} H_i c_j(X_{i,k}) = \sum_{m=1}^{m_{March}} w_{i,m}^{2, March} H_i c_j(X_{i,k}). \quad (12)$$

This fourth set of moments includes information on race, Hispanic origin, income (from the linked administrative data), and the number of adults and children in the household. Without this set of conditions, estimates of the number of households by type (especially for oversampled groups) differ between the full and March Basic CPS ASEC samples. Additionally, without these constraints, observables-based oversampling in the full CPS ASEC biases estimates for oversampled subgroups relative to estimates from the March Basic sample. Although we focus on the estimates from the full CPS ASEC sample in this paper, we present the results from the Basic March sample as well, because it is a stratified random sample with no oversampling based on observable characteristics from earlier survey responses.

We call the final weights using this procedure the entropy balance weights (EBW). For valid inference, we repeat the above two-stage reweighting procedure 160 additional times using the baseline successive difference replicate factors created during the sampling process, which are available for all households regardless of response status. These replicate factors account for the sampling design of the monthly Basic CPS and CPS ASEC. Also, the first-stage target moments from the March Basic CPS sample are estimates and thus subject to sampling error. By repeating the procedure with the base weights and replicate factors, the

variation in the final weights across the replicates will reflect this uncertainty as well.¹⁶ All standard errors reported using EBW are calculated with these 160 replicate-factor EBW.

4 The Coronavirus Pandemic and Nonresponse in the 2020 and 2021 CPS ASEC

We apply this weighting technique to evaluate nonresponse bias during the Coronavirus pandemic. The pandemic has had wide-ranging impacts on the lives and well-being of individuals and households. Surveys of those individuals and households are an important input into understanding those impacts. However, survey operations themselves have also been affected by the pandemic, which may affect the quality of the data we use to evaluate these impacts.

In 2020 data collection faced extraordinary circumstances. On March 11, 2020, the World Health Organization announced that COVID-19 was a pandemic. Interviewing for CPS ASEC in March began on March 15. In order to protect the health and safety of Census Bureau staff and respondents, the survey suspended in-person interviewing and closed the two Computer-Assisted Telephone Interviewing (CATI) centers on March 20. Through April the Census Bureau continued to attempt all interviews by phone. For those whose first month in the survey was March or April, the Census Bureau used vendor-provided telephone numbers associated with the sample address to try to reach households.¹⁷

While the Census Bureau went to great lengths to complete interviews by telephone, the response rate for the Basic CPS was 73 percent in March 2020, about 10 percentage points lower than in preceding months and the same period in 2019.¹⁸ Figure 1 shows the

¹⁶At present, we do not include uncertainty in the external population targets, but we hope to explore how best to account for that uncertainty in the weights as well in future research.

¹⁷For a more complete description of data collection during the pandemic, see Berchick, Mykyta and Stern (2020).

¹⁸This paper focuses on response at the housing unit level, or unit nonresponse. In unit nonresponse, no response information is available from any individual in the household. Nonresponse is also possible at the item level. For item nonresponse, an individual responds to the survey but does not answer a particular

unweighted response rate of the Basic CPS from April 2010 to May 2021. The sharp decline in response in March and April 2020 is clearly visible.

Additionally, the BLS stated in their FAQ accompanying the April 3 release of the March Employment Situation, “Response rates for households normally more likely to be interviewed in person were particularly low. The response rate for households entering the sample for their first month was over 20 percentage points lower than in recent months, and the rate for those in the fifth month was over 10 percentage points lower.”¹⁹

By the 2021 CPS ASEC (February-March), in-person interviewing was again possible throughout the country and response rates had rebounded. However, they were still below the pre-pandemic trend.

The CPS ASEC response rate is complicated by the different months and samples that feed into the survey.²⁰ Further, it includes an adjustment factor to account for those who responded to the Basic survey but did not answer the supplement.²¹ The Census Bureau estimates that the combined supplement unweighted response rate was 61.1 percent in 2020, down from 67.6 percent in 2019.

In processing responses to the CPS ASEC (or any survey), the Census Bureau has methods in place to adjust for nonresponse, through survey weights. As discussed above, for the CPS ASEC this includes several stages of adjustment. One adjustment controls for differential response rates of housing units within and outside of Metropolitan Statistical Areas. Additional weighting adjustments control the CPS ASEC sample to independent population estimates by age, sex, race, and Hispanic origin at the national and state levels. These controls ensure that the weighted shares of groups in the CPS ASEC match closely to their

question. Because the CPS ASEC is a supplement to the Basic CPS, it is also possible for an individual to be a supplement nonrespondent. In that case, the individual answers the Basic CPS but does not provide enough information to questions in the ASEC supplement to be considered a respondent.

¹⁹<https://www.bls.gov/cps/employment-situation-covid19-faq-march-2020.pdf>. The Basic CPS uses a 4-8-4 design, where housing units are in sample for four months, called month-in-sample (MIS) 1-4, then out of sample for 8 months, and then back in sample for 4 months, MIS 5-8.

²⁰Additional housing units are added to the CPS ASEC sample to oversample Hispanics and households with children, as discussed later in the paper.

²¹These supplement nonrespondents are included in the ASEC sample, with their ASEC income imputed conditional on their responses to questions in monthly CPS.

independently estimated shares in the target population.

To assess nonresponse bias in the CPS ASEC, we link addresses selected for inclusion in the sample to various sources of administrative and prior survey and decennial census data. These data include administrative earnings and income as well as demographic information such as individual age, race, gender, citizenship, and education. Using this information we evaluate how households that do and do not respond to the survey differ over time.²²

Berchick, Mykyta and Stern (2020) also examine the 2020 CPS ASEC for evidence of nonresponse bias, with a particular focus on estimates of health insurance coverage. They examine changes in the characteristics of respondents over time and compare health insurance estimates from the CPS ASEC to estimates from other surveys.

Two papers assess nonresponse bias during the pandemic in the monthly CPS in 2020. Ward and Edwards (2020) show that the distributions of demographic and socioeconomic characteristics change as response rates decline in the early months of the pandemic. Heffetz and Reeves (2021) use survey design features and information on the number of contact attempts to estimate rotation-group bias and difficulty-to-reach bias. They find potential evidence of bias in estimates of the unemployment rate, but the direction and magnitude of the bias are uncertain.

5 Results

5.1 Evaluating Nonresponse using Linked Data

Table 2 shows the share of housing units that can be linked to each source of data used, either at the address/MAFID level for the 1099 IRMF or at the person/PIK level for the other files.

²²Households may not respond to a survey for a variety of reasons, such as inability to contact a household member, refusal to respond, or inability to respond (for example, due to language barriers). In 2020 in particular, one of those reasons could have been the inability of Census field representatives to reach a member of the household. Noninterview households may be a more accurate way to describe the households that could not be reached or refused the CPS interview. However, as nonresponse is the term used in the literature, we use that in this paper.

In non-pandemic years (2017 to 2019, in Columns (1)-(3)), respondents and nonrespondents differ slightly in the forms that can be linked to their addresses. Respondents are more likely to have any information return in the 1099 IRMF, less likely to have a W-2, more likely to have a 1099-R, more likely to have filed a 1040 (in the prior year), and more likely to have an individual that can be linked to a 2010 census or ACS respondent. However, the relationships are not statistically different over time as the year-to-year comparisons of respondents and nonrespondents show in Columns (6) and (7).²³

However, as shown in Column (8), the year-to-year change in the differences between respondents and nonrespondents is larger in 2020 for most linked data sets. Response in 2020 was increasingly associated with the presence of an information return (1099 IRMF), the presence of a W-2, filing a tax return (1040) in the prior year, and linkage to the 2010 census. In 2021, the linkage rate gaps between respondents and nonrespondents shrink slightly (Column (9)).

With the linked data, we can summarize the characteristics of responding and nonresponding housing units. Table 3 shows summary statistics on race, Hispanic origin, nativity, and education for linked housing units. Race and Hispanic origin come from the linked 2010 census. The value for a given household is set to one if at least one individual in the housing unit is in that race or Hispanic-origin group in the 2010 census and zero otherwise. Nativity information comes from the Numident, and again, the categories are set to one if a household member is in each group in the Numident and zero otherwise. Education information comes from the ACS, and a household is categorized by the reported educational attainment of the most educated linked individual. Housing units are only included in the sample for each summary statistic if at least one member is linked to the corresponding source data set.

In Columns (1)-(5), Table 3 compares the characteristics of respondents and nonrespondents in each year from 2017 to 2021.²⁴ In each year, respondents are less likely to be Black

²³All statistics in this section use the base weights that reflect the probability of selection into the sample and standard errors are calculated using the baseline replicate factors that account for the sample design.

²⁴For 2017, we use the CPS ASEC Research File, and for 2018, we use the CPS ASEC Bridge File. These files incorporate updates to the CPS ASEC processing system, implemented in 2019. By using these files,

and they are more likely to be White.²⁵ Columns (6)-(9) again show the change each year in the estimates shown in (1)-(5). The results show that response in 2020 was increasingly associated with being non-Hispanic, native born, and more educated. There were not many differences between 2020 and 2021 respondents and nonrespondents (Column (9)).

Using the linked data, we can also evaluate how household response correlates with administrative income. We test two measures of income: 1) the sum of all W-2 earnings at the address in the prior year (matching the survey reference year) and 2) the sum of adjusted gross income (AGI) for income one year before the reference period on tax returns filed by linked individuals at the address in the survey year.

In Table 4, we compare the mean and various percentiles (10th, 25th, median, 75th, and 90th) of income for respondents and nonrespondents over time, with the results shown in Figure 3 as well. The annual estimates from 2017 to 2021 are shown in Columns (1)-(5). While there are differences between respondents and nonrespondents from 2017 to 2019, most comparisons of W-2 and AGI income statistics are not statistically different. However in 2020 and 2021, respondents have higher income than nonrespondents at nearly every percentile in the table. The difference-in-difference comparisons in Columns (6)-(9) also highlight how unique selection into response on income was in 2020. For every statistic except mean AGI, respondents had higher incomes relative to nonrespondents in 2020 than in 2019, whereas the same was not true for most other year-to-year comparisons of respondents and nonrespondents.

However, it is possible that income is highly correlated with observable characteristics, such as age, which are controlled for in the current weighting system. The state-level race, Hispanic origin, age, and gender information could in principle fully adjust the weights to account for selection into response by income. To test whether this is likely, we regress survey response on administrative income (in various income bins) with and without conditioning

we are not comparing across a break in series. See Semega et al. (2019) for more information on the updated processing system.

²⁵They are also less likely to be high-school graduates and more likely to be college graduates in three of the four years.

on the other demographic and socioeconomic information available in the linked data. In the controls, we include information from linked individuals on race, age, Hispanic origin, education, citizenship status, dummies for each linked administrative data source, state fixed effects, and the number of linked household members.

The results are shown in Table 5, Figure 4 (no controls), and Figure 5 (full controls) for W-2 earnings.²⁶ With or without controls, response in 2020 and 2021 was more strongly associated with income than prior years, whether income was measured as W-2 earnings or prior-year 1040 AGI.²⁷

From 2017 to 2019, we do not see strong evidence of nonresponse bias due to differential nonresponse by low- and high-income households. This is consistent with the results in Bee, Gathright and Meyer (2015), which does not find strong evidence of nonresponse bias using 1040 data in the 2011 CPS ASEC.

However, income is strongly associated with nonresponse in the 2020 and 2021 CPS ASEC. High-income households, as measured by their W-2 earnings or 1040 AGI in the prior year, are more likely to respond than low-income households. Conditioning on observable demographic and socioeconomic data did not eliminate this variation in nonresponse by income.

Differential nonresponse has the potential to bias many estimates generated from CPS and CPS ASEC data. The pattern of nonresponse during the pandemic could bias income up and poverty down, with additional effects on other correlated statistics such as health insurance coverage, education, etc.

²⁶For AGI in the prior year, the results are available in Figure A1 (no controls), and Figure A2 (full controls), with the values shown in Table A1.

²⁷We also conducted robustness checks to test whether was primarily due to respondents in their 1st and 5th months in the sample, where face-to-face interviews are more often required. We found selection in income for both groups in 2020 when we divided the sample into: 1) months-in-sample 1 and 5, and 2) months-in-sample 2-4 and 6-8, shown in Tables A2 and A3.

5.2 Entropy Balance Weights

5.2.1 Summary Statistics

To evaluate our weighting procedure, we compare the survey estimates to both sets of EBW: 1) the full CPS ASEC sample (denoted Full EBW or EBW in the tables and figures) and 2) the March Basic CPS ASEC sample (denoted March EBW in the tables and figures). In the text we will primarily focus on the Full EBW comparisons.

Table 7 compares summary statistics between the full sample of respondents and nonrespondent households to the respondents only using the unadjusted base weights. Columns (1)-(5) use the March base weights, which reflect the probability of selection into the sample for each housing unit. These estimates are the target distribution for the first-stage entropy balance adjustment. As expected, without adjusting for oversampling or selection into response, there are important differences in the samples. For example, from Columns (11)-(15), March Basic CPS respondents select into response by age, education, and race. The estimates for the CPS ASEC sample in Columns (6)-(10) reflect both nonrandom nonresponse and the characteristics of oversampled households.

Table 8 shows these same comparisons after the EBW nonresponse adjustment. By construction, we no longer see many meaningful or statistically significant differences between the EBW-based estimates and the baseline estimates from nonvacant units.²⁸

Table 9 summarizes various demographic and socioeconomic characteristics using the different weights at the person level. For the external population targets of the EBW adjustment (such as for Blacks, Whites, and Hispanics), the point estimates of the differences between the differences round to 0. However, there are differences in the estimates, especially for 2020 and 2021. For example, the EBW weights estimate lower levels of education during the pandemic than the survey weights. EBW weights also estimate different shares of native and foreign-born citizens than the survey in some years.

²⁸Even for characteristics that are targets for the entropy balance procedure, there can be differences in the estimates as not all moment conditions can be matched exactly, especially with a large number of moment constraints. However, the magnitudes of the statistically significant differences are small in all cases.

5.2.2 Income and Poverty Estimates

Using the alternative weights, we estimate various statistics of income and poverty to assess the bias from selection into response, for survey years 2017 to 2021 (and reference years 2016 to 2020).

Note that we continue to refer to the survey years in the text, tables, and figures to keep the year references consistent across table and more clearly identify the 2020 and 2021 CPS ASECs as the ones affected by the pandemic. However, keep in mind that the reference period is the prior year in the CPS ASEC. Therefore, for example, when we discuss statistics for the 2021 CPS ASEC, we are discussing income earned or received in 2020.

In Table 10, we estimate household income at five-percent intervals from the 5th to 95th percentiles, using linear interpolation. In Table 11 and Figure 6, Panel A, we show comparisons between the estimates using the survey weights and alternative weights. There are no statistically significant differences between the full EBW and survey estimates from 2017 to 2019 and only a handful for the March EBW compared to the survey. However, in 2020 and 2021 using the full EBW, we estimate much lower income across the distribution than with survey weights. In 2020 For the 25th, 50th, and 75th percentiles, the respective full EBW estimates are 3.2 percent, 3.0 percent, and 2.2 percent lower than the survey.²⁹

Table 12 and Figure 6, Panel B show estimates of year-to-year growth in real household income using each weight. For 2018 and 2019, year-to-year changes track very closely to the estimates using alternative weights, with no statistically significant differences in the year-to-year growth. However, there is a level difference in the estimates from the 2020 ASEC, with the EBW estimating substantially lower growth in income. Because there is a level difference for both 2020 and 2021, the year-to-year estimates using the survey weights and EBW are not statistically different from 2020 to 2021.

In the 2020 CPS ASEC, real median household income increased 6.8 percent using the survey weights, compared to 3.8 percent with the full EBW. This would change the year-to-

²⁹The three estimates (3.2, 3.0, and 2.2 percent) are not statistically different from each other.

year increase estimated from the 2020 CPS ASEC from the largest point estimate increase in the series (going back to 1967) to the 91st percentile of year-to-year changes. The adjusted estimates would indicate that 2019 (from the 2020 CPS ASEC) was still a very good year for income, even if it did not necessarily have the most year-to-year growth in the historical income series.

Poverty estimates are shown in Table 13. For both the 2020 and 2021 CPS ASEC, the EBW estimates of poverty were 0.44 and 0.46 percentage points higher respectively than the official survey-weighted estimate. The point estimates for poverty with the EBW were higher, but not statistically different, prior to the pandemic.

Estimates for the Supplemental Poverty Measure (SPM) are also shown in Table 13.³⁰ There are no statistically significant differences in levels or changes in the SPM estimates using the survey weights or EBW.

6 Public-Use Weights

Entropy balancing is also very amenable to the release of weights for public use. To release weights based on administrative data, we would like the weights to replicate important estimates while protecting the privacy of respondents.

We achieve this by defining moment conditions from a set of covariates that is only available in the survey, $X_{i,j}^S$. We include target moments from survey-reported demographics, household and personal income, poverty, education, health insurance status, among other survey characteristics.³¹ We can then estimate public-use weights, w_i^{PU} , with initial weights equal to the sampling probability weights q_i , subject to the following constraints:

$$\sum_{i=1}^n w_i^{PU} c_j(X_{i,j}^S) = \sum_{i=1}^n w_i^2 c_j(X_{i,j}^S). \quad (13)$$

³⁰For more information about the Supplemental Poverty Measure, see Fox (2020).

³¹This approach is similar to the one taken in Mittag (2019). In both cases, summary statistics from linked data (For Mittag, regression coefficients and in this paper, means) are then used to adjust the unlinked survey data to address bias while protecting against the disclosure of the linked administrative data.

The constraints in Equation 13 ensure that important statistics match when estimated from the full EBW and the public-use EBW. However, because the public-use EBW only target the moments of characteristics available in survey responses, they help protect the linked information against disclosure. For example, if having high AGI or W-2 earnings predicts response after conditioning on survey responses, then having a lower weight than expected given the survey information in the full EBW suggests that an individual or household had higher than expected administrative income. With the public-use EBW, that would not necessarily be the case. The public-use weights reflect the expected response probability of people with the same survey characteristics (given the distribution of linked information for those people), not necessarily that individual or household’s administrative information.³²

Our public-use weights are estimated using the same two-stage procedure as discussed in section 3.2 and shown in Table 6. However, for the public-use weights, in both stages the moments are estimated from the full CPS ASEC sample using the full EBW. The first-stage public-use reweighting ensures that the included survey response moments at the household level match when estimated using the public-use EBW and the full EBW. The second stage of reweighting ensures that the person-level moments also match, while preserving the match at the household level as well.

For means and proportions (such as poverty rates or mean household income), the public-use EBW estimates will match the full EBW by construction. However, that is not the case for some statistics of interest, such as medians. Medians cannot be targeted as moment constraints in entropy balancing as medians are functions of the distribution, not of individual $X_{i,j}$ values. In Table A4, we show estimates of median household income for various subgroups using the survey weights, the full EBW, and the public-use EBW, for reference.

³²Public-use weights are available at <https://www.census.gov/data/datasets/time-series/demo/income-poverty/data-extracts.html>.

7 Conclusion

Survey response rates have been declining for decades. The Coronavirus pandemic also affected survey operations and, potentially, respondent behavior. As a result, response rates declined further in the CPS beginning in March 2020. They have rebounded since, but still are below the historical trend. We evaluated selection into nonresponse using administrative, survey, and decennial census data linked to respondent and nonrespondent addresses. We found that nonresponse varied by income during the pandemic in particular, with high-income households more likely to respond than low-income households. This relationship between income and nonresponse held even after controlling for other observable demographic and socioeconomic characteristics. We used entropy balancing to adjust the weights for selection into nonresponse in the CPS ASEC from 2017 to 2021. This adjustment had relatively small or no significant effect on income or poverty estimates from 2017 to 2019. However, estimates of income in 2020 and 2021 were adjusted downward by several percent. The nonresponse adjustment also increased estimates of poverty by nearly a half a percentage point, or about 1.5 million individuals. Finally, we applied entropy balancing to create public-use weights that protect the confidentiality of respondents, when it would be difficult to do so for weights estimated on the linked administrative data.

Furthermore, because the ASEC is a supplement to the monthly CPS, many of the households in our study are also the source of data for the monthly unemployment statistics. Therefore, this work also highlights an important potential source of bias (and variation in bias) over the course of the pandemic in the monthly data. Nonresponse to the monthly CPS increased from 17.7 percent in February to 27.0 and 30.1 percent in March and April respectively (the period of this study), to a peak of 35.3 percent in June before declining again to 19.7 percent by October (as shown in Figure 1). Future applications of this data and methods can help to untangle the interaction between changing economic conditions and the effects of those changes on the tools used to monitor economic activity.

While we did not see as large an impact of the adjustment on prior years, there are still

differences between the EBW estimates and the estimates using existing survey weights, such as by race, education, citizenship, and nativity in some years. We believe this approach has the potential to improve survey weights and reduce nonresponse bias in survey-based estimates beyond the CPS ASEC. Accordingly we have begun research to apply this method to several other national household surveys, such as the American Community Survey (Rothbaum et al., 2021), and as part of a larger ongoing project to combine the relative advantages of surveys and auxiliary records to create estimates of income that are less subject to survey misreporting and measurement error, as discussed in Bee and Rothbaum (2019).

References

- Bee, C. Adam, and Jonathan Rothbaum.** 2019. “The Administrative Income Statistics (AIS) Project: Research on the Use of Administrative Records to Improve Income and Resource Estimates.” *U.S. Census Bureau SEHSD Working Paper #2019-36*.
- Bee, C. Adam, Graton Gathright, and Bruce D. Meyer.** 2015. “Bias from Unit Non-response in the Measurement of Income in Household Surveys.” *Unpublished U.S. Census Bureau Working Paper*.
- Berchick, Edward R., Laryssa Mykyta, and Sharon M. Stern.** 2020. “The Influence of COVID-19-Related Data Collection Changes on Measuring Health Insurance Coverage in the 2020 CPS ASEC.” *U.S. Census Bureau SEHSD Working Paper #2020-13*.
- Brummet, Quentin.** 2014. “Comparison of Survey, Federal, and Commercial Address Data Quality.” *U.S. Census Bureau CARRA Working Paper #2014-06*.
- Brummet, Quentin, Denise Flanagan-Doyle, Joshua Mitchell, John Voorheis, Laura Erhard, and Brett McBride.** 2018. “What Can Administrative Tax Information Tell Us about Income Measurement in Household Surveys? Evidence from the Consumer Expenditure Surveys.” *Statistical Journal of the IAOS*, 34(4): 513–520.
- Deville, Jean-Claude, and Carl-Erik Särndal.** 1992. “Calibration estimators in survey sampling.” *Journal of the American statistical Association*, 87(418): 376–382.
- Dixon, John S.** 2007. “Nonresponse Bias Patterns in the Current Population Survey.” *Unpublished Paper, Bureau of Labor Statistics*.
- Eggleston, Jonathan.** 2021. “Comparing Respondents and Nonrespondents in the ACS: 2013-2018.” *Unpublished U.S. Census Bureau Working Paper*.
- Eggleston, Jonathan, and Ashley Westra.** 2020. “Incorporating Administrative Data in Survey Weights for the Survey of Income and Program Participation.” *U.S. Census Bureau SEHSD Working Paper #2020-07*.
- Estevao, Victor M, and Carl-Erik Särndal.** 2006. “Survey estimates by calibration on complex auxiliary information.” *International Statistical Review*, 74(2): 127–147.
- Fox, Liana.** 2020. “The Supplemental Poverty Measure: 2019.” *U.S. Census Bureau Current Population Reports*.
- 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.
- Groves, Robert M., and Mick P. Couper.** 2012. *Nonresponse in Household Interview Surveys*. John Wiley & Sons.
- Hainmueller, Jens.** 2012. “Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies.” *Political Analysis*, 25–46.

- Heffetz, Ori, and Daniel B. Reeves.** 2019. “Difficulty of Reaching Respondents and Nonresponse Bias: Evidence from Large Government Surveys.” *Review of Economics and Statistics*, 101(1): 176–191.
- Heffetz, Ori, and Daniel B. Reeves.** 2021. “Measuring Unemployment in Crisis: Effects of COVID-19 on Potential Biases in the CPS.”
- Hlasny, Vladimir.** 2020. “Unit Nonresponse Bias in Inequality Measurement: Worldwide Analysis Using Luxembourg Income Study database.” *Social Science Quarterly*.
- Hlasny, Vladimir, and Paolo Verme.** 2018. “The Impact of Top Incomes Biases on the Measurement of Inequality in the United States.”
- Korinek, Anton, Johan A. Mistiaen, and Martin Ravallion.** 2006. “Survey Nonresponse and the Distribution of Income.” *The Journal of Economic Inequality*, 4(1): 33–55.
- Korinek, Anton, Johan A. Mistiaen, and Martin Ravallion.** 2007. “An Econometric Method of Correcting for Unit Nonresponse Bias in Surveys.” *Journal of Econometrics*, 136(1): 213–235.
- Luiten, Annemieke, Joop Hox, and Edith de Leeuw.** 2020. “Survey Nonresponse Trends and Fieldwork Effort in the 21st Century: Results of an International Study across Countries and Surveys.” *Journal of Official Statistics*, 36(3): 469–487.
- Mattingly, Tracy, Jamie Choi, Tremika Finney, Rebecca Hoop, David Hornick, Danielle Nieman, Cynthia Rothhaas, Ashley Westra, and Michael White.** 2016. “Results of a Nonresponse Bias Analysis Using Survey of Income and Program Participation (SIPP) Addresses Matched to Internal Revenue Service (IRS) Data.” *Memorandum, U.S. Census Bureau*.
- Mittag, Nikolas.** 2019. “Correcting for Misreporting of Government Benefits.” *American Economic Journal: Economic Policy*, 11(2): 142–164.
- Rosenbaum, Paul R, and Donald B Rubin.** 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika*, 70(1): 41–55.
- Rothbaum, Jonathan, Jonathan Eggleston, Adam Bee, Mark Klee, and Brian Mendez-Smith.** 2021. “Addressing Nonresponse Bias in the American Community Survey During the Pandemic Using Administrative Data.” *U.S. Census Bureau SEHSD Working Paper #2021-24*.
- Sabelhaus, John, David Johnson, Stephen Ash, David Swanson Swanson, Thesia I. Garner, John Greenlees, and Steve Henderson.** 2015. “Is the Consumer Expenditure Survey Representative by Income?” *Improving the Measurement of Consumer Expenditures*, 74: 241.
- Semega, Jessica, Melissa Kollar, Emily A. Shrider, and John Creamer.** 2020. “Income and Poverty in the United States: 2019.” *U.S. Census Bureau Current Population Reports*.

- Semega, Jessica, Melissa Kollar, John Shrider, Creamer, and Abinash Mohanty.** 2019. "Income and Poverty in the United States: 2018." *U.S. Census Bureau Current Population Reports*.
- Wagner, Deborah, and Mary Layne.** 2014. "The Person Identification Validation System (PVS): Applying the Center for Administrative Records and Research and Applications' record linkage software." *U.S. Census Bureau CARRA Report Series #2014-01*.
- Ward, Jason M, and Kathryn A Edwards.** 2020. "Statistics in the Time of Coronavirus: COVID-19-related Nonresponse in the CPS Household Survey." *Rand Working Paper #WR-A842-1*.
- Williams, Douglas, and J. Michael Brick.** 2018. "Trends in US Face-To-Face Household Survey Nonresponse and Level of Effort." *Journal of Survey Statistics and Methodology*, 6(2): 186–211.
- Zhao, Qingyuan, and Daniel Percival.** 2017. "Entropy Balancing is Doubly Robust." *Journal of Causal Inference*, 5(1).

Table 1: Data Used in this Paper

Data Set	Link Variable	Description	Variables Added
CPS ASEC Household File		Sampling and geographic information for all households in the CPS ASEC sample, whether they responded or not	MAFID, housing unit survey identifiers, location, response type, other sampling information, and survey information for responding households
CPS ASEC Person File	Housing unit survey IDs	Survey information for responding individuals	
1099 Information Returns Master File	MAFID	Person-level file of information returns filed for each individual by week 19 of the survey year. Covers income earned during the CPS ASEC reference period. No income information is contained in this file.	PIK for individuals receiving returns, flags for forms: W-2, 1098, 1099-DIV, 1099-G, 1099-INT, 1099-MISC, 1099-R, 1099-S, and SSA-1099
W-2 Return Master File	PIK	Universe of job-level earnings filed through week 19 of the survey year. Covers income earned during the CPS ASEC reference period.	Taxable earnings, deferred compensation
1099-R Return Master File	PIK	Universe level information return covering defined-contribution and defined-benefit pension plan earnings, as well as other survivor and disability income. Includes returns filed through week 19. Covers income earned during the CPS ASEC reference period.	Income from pension plans, withdrawals from defined-contribution retirement plans (such as 401(k)s), income from survivor and disability pension plans
1040 Master File	PIK	Universe of 1040 filings filed in the prior calendar year for income earned the year before the CPS ASEC reference period.	Adjusted gross income, wage and salary income, interest income, dividend income, gross rental income for tax units that filed taxes in the year prior to the CPS ASEC
SSA Numident	PIK	SSA master file of individuals with Social Security Numbers	Age and citizenship status
Census 2010 Short Form	PIK		Race and age
American Community Survey	PIK	Pooled responses to all ACS files from 2001-2018	Education

Notes: This table shows the administrative and survey data sets that are linked to CPS ASEC respondents and nonrespondent households. The initial link is at the address level to the 1099 IRMF file of information returns. Each subsequent is conditional on the 1099 IRMF link at the housing unit level, and all subsequent links are at the person level, using PIKs. Because the tax filing deadline was delayed in 2020 until July 15, we do not use 1040s filed in 2020 due to concerns about non-random selection of households into early filing in 2020 that would make comparisons to prior years difficult.

Table 2: Linkage Rates for Various Data Sources to CPS ASEC Respondents and Nonrespondents

Households Linked To:	Year					Difference			
	2017 (1)	2018 (2)	2019 (3)	2020 (4)	2021 (5)	2018-2017 (6)	2019-2018 (7)	2020-2019 (8)	2021-2020 (9)
1099 IRMF									
Respondents	0.8242*** (0.002398)	0.8231*** (0.002194)	0.8128*** (0.00247)	0.8355*** (0.002483)	0.8345*** (0.002159)	-0.001084 (0.002215)	-0.01039*** (0.002427)	0.02272*** (0.00231)	-0.001017 (0.002263)
Nonrespondents	0.7874*** (0.004893)	0.7818*** (0.005006)	0.7663*** (0.004175)	0.753*** (0.00427)	0.7637*** (0.004241)	-0.00552 (0.006337)	-0.01557** (0.006122)	-0.01324** (0.005414)	0.01069** (0.005223)
Respondents - Nonrespondents	0.03687*** (0.004346)	0.04131*** (0.004687)	0.04649*** (0.004347)	0.08246*** (0.004233)	0.07075*** (0.00413)	0.004436 (0.006175)	0.005185 (0.00632)	0.03596*** (0.005657)	-0.01171** (0.005774)
W2									
Respondents	0.6498*** (0.002841)	0.6429*** (0.002458)	0.6338*** (0.002646)	0.6542*** (0.002746)	0.6445*** (0.002581)	-0.006874** (0.002852)	-0.00907*** (0.002542)	0.02037*** (0.002668)	-0.009741*** (0.002637)
Nonrespondents	0.6718*** (0.005939)	0.6571*** (0.005405)	0.643*** (0.004737)	0.6352*** (0.004768)	0.6372*** (0.004838)	-0.01473** (0.007297)	-0.0141** (0.006294)	-0.007778 (0.006114)	0.001973 (0.005902)
Respondents - Nonrespondents	-0.02206*** (0.005712)	-0.0142*** (0.005199)	-0.009173* (0.004795)	0.01898*** (0.004823)	0.007264 (0.004852)	0.007856 (0.00746)	0.005027 (0.006604)	0.02815*** (0.006238)	-0.01171* (0.00642)
1099R									
Respondents	0.3329*** (0.002643)	0.3374*** (0.00252)	0.3342*** (0.002456)	0.2261*** (0.00215)	0.2379*** (0.002509)	0.004502* (0.002714)	-0.003161 (0.002761)	-0.1081*** (0.002675)	0.01177*** (0.00293)
Nonrespondents	0.2711*** (0.005178)	0.2763*** (0.005119)	0.2708*** (0.004891)	0.1548*** (0.003345)	0.1655*** (0.003141)	0.005221 (0.006475)	-0.005457 (0.006036)	-0.116*** (0.005695)	0.01067** (0.004227)
Respondents - Nonrespondents	0.06181*** (0.005116)	0.06109*** (0.005028)	0.06338*** (0.004787)	0.0713*** (0.003249)	0.07241*** (0.003463)	-0.000719 (0.006818)	0.002296 (0.006185)	0.00792 (0.005749)	0.001107 (0.004537)
1040									
Respondents	0.7429*** (0.002759)	0.7403*** (0.002573)	0.7304*** (0.002757)	0.7565*** (0.002556)	0.7538*** (0.002325)	-0.002518 (0.002593)	-0.009947*** (0.002707)	0.02609*** (0.002507)	-0.00272 (0.002561)
Nonrespondents	0.7148*** (0.005585)	0.7124*** (0.005328)	0.6936*** (0.004713)	0.6737*** (0.004435)	0.6878*** (0.004694)	-0.002396 (0.00706)	-0.01883*** (0.006589)	-0.01991*** (0.006105)	0.01414** (0.005614)
Respondents - Nonrespondents	0.02805*** (0.005222)	0.02793*** (0.005312)	0.03681*** (0.004905)	0.0828*** (0.004651)	0.06594*** (0.00477)	-0.0001225 (0.007043)	0.008878 (0.006956)	0.04599*** (0.006414)	-0.01686*** (0.006319)
2010 Census									
Respondents	0.7713*** (0.002574)	0.7706*** (0.002395)	0.756*** (0.002686)	0.7746*** (0.002847)	0.7728*** (0.002431)	-0.0006161 (0.002384)	-0.01461*** (0.002554)	0.01858*** (0.002557)	-0.001848 (0.002589)
Nonrespondents	0.7178*** (0.00524)	0.7066*** (0.005369)	0.6929*** (0.004752)	0.6733*** (0.004834)	0.6839*** (0.004644)	-0.01118 (0.006953)	-0.01367** (0.006289)	-0.01957*** (0.006093)	0.01061* (0.005473)
Respondents - Nonrespondents	0.05351*** (0.004855)	0.06407*** (0.004953)	0.06313*** (0.00481)	0.1013*** (0.004546)	0.08883*** (0.004405)	0.01056 (0.006844)	-0.0009447 (0.006451)	0.03816*** (0.006291)	-0.01246** (0.00582)
ACS									
Respondents	0.2224*** (0.002129)	0.2226*** (0.002122)	0.2184*** (0.002031)	0.2252*** (0.002251)	0.2252*** (0.002088)	0.0001277 (0.002293)	-0.004171** (0.002127)	0.00678*** (0.002275)	0.00003234 (0.002145)
Nonrespondents	0.1863*** (0.004637)	0.1767*** (0.004057)	0.1716*** (0.003738)	0.18*** (0.003587)	0.1793*** (0.003426)	-0.009546 (0.005825)	-0.005138 (0.004584)	0.008451* (0.004433)	-0.0007026 (0.004139)
Respondents - Nonrespondents	0.03618*** (0.004613)	0.04586*** (0.004023)	0.04682*** (0.00389)	0.04515*** (0.003929)	0.04589*** (0.003602)	0.009674 (0.006117)	0.0009676 (0.00508)	-0.001671 (0.004869)	0.0007349 (0.004587)

Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This table shows the unconditional link rate between housing units in the full CPS ASEC sample and each data set in Table 1. The initial link is at the address level to the 1099 IRMF file of information returns. Each subsequent is conditional on the 1099 IRMF link at the housing unit level, and all subsequent links are at the person level, using PIKs. For person-/PIK-based links, a housing unit is classified as linked if at least one PIK can be linked. Standard errors are shown in parenthesis. ***, **, and * indicate statistical significance at the 1-, 5-, and 10-percent levels respectively, but asterisks are only shown for differences as all estimates for respondents and nonrespondents are significant at the 1-percent level.

Table 3: Shares of Characteristics of the CPS ASEC Sample from Linked Data for Respondent and Nonrespondent Households

Characteristic	Year					Difference			
	2017 (1)	2018 (2)	2019 (3)	2020 (4)	2021 (5)	2018-2017 (6)	2019-2018 (7)	2020-2019 (8)	2021-2020 (9)
Race									
Black									
Respondents	0.1346*** (0.002732)	0.1351*** (0.002495)	0.1343*** (0.002792)	0.1339*** (0.002672)	0.1412*** (0.002808)	0.0005482 (0.002313)	-0.0008427 (0.002457)	-0.0003569 (0.002567)	0.007209*** (0.002288)
Nonrespondents	0.1603*** (0.00547)	0.1672*** (0.005189)	0.1678*** (0.004515)	0.17*** (0.004825)	0.174*** (0.004548)	0.006914 (0.006216)	0.0005049 (0.005836)	0.002292 (0.005562)	0.003977 (0.005202)
Respondents - Nonrespondents	-0.02574*** (0.004726)	-0.03211*** (0.004624)	-0.03346*** (0.004101)	-0.0361*** (0.004617)	-0.03287*** (0.003623)	-0.006366 (0.006011)	-0.001348 (0.005752)	-0.002649 (0.005462)	0.003232 (0.005461)
White									
Respondents	0.8238*** (0.00265)	0.8247*** (0.002753)	0.8291*** (0.003036)	0.8252*** (0.002845)	0.8226*** (0.00288)	0.0009772 (0.002588)	0.004374 (0.002727)	-0.003901 (0.002806)	-0.002619 (0.002502)
Nonrespondents	0.809*** (0.005653)	0.8064*** (0.005382)	0.8016*** (0.00513)	0.7901*** (0.005239)	0.7943*** (0.004567)	-0.002527 (0.006817)	-0.00486 (0.006179)	-0.01142* (0.00592)	0.004169 (0.005564)
Respondents - Nonrespondents	0.01481*** (0.005273)	0.01832*** (0.004814)	0.02755*** (0.004708)	0.03507*** (0.004809)	0.02828*** (0.003813)	0.003504 (0.006625)	0.009233 (0.006028)	0.00752 (0.005805)	-0.006788 (0.005686)
Hispanic									
Respondents	0.1323*** (0.002136)	0.1341*** (0.002664)	0.1383*** (0.002455)	0.1365*** (0.002416)	0.1385*** (0.002579)	0.001735 (0.002475)	0.004197* (0.002448)	-0.001798 (0.002651)	0.002012 (0.002442)
Nonrespondents	0.1145*** (0.004388)	0.1171*** (0.00448)	0.1283*** (0.004521)	0.1522*** (0.004117)	0.149*** (0.004798)	0.002521 (0.005229)	0.01126** (0.005666)	0.02385*** (0.005161)	-0.003225 (0.00493)
Respondents - Nonrespondents	0.0178*** (0.004211)	0.01701*** (0.004347)	0.009949** (0.003949)	-0.0157*** (0.003738)	-0.01046** (0.004181)	-0.000786 (0.0054)	-0.007062 (0.005691)	-0.02565*** (0.004885)	0.005237 (0.004893)
Citizenship									
Native Born									
Respondents	0.9269*** (0.001543)	0.9245*** (0.001599)	0.9215*** (0.001657)	0.9246*** (0.001639)	0.9241*** (0.001717)	-0.002374 (0.001801)	-0.003045* (0.001711)	0.003089* (0.001756)	-0.0004551 (0.001742)
Nonrespondents	0.9332*** (0.003027)	0.9228*** (0.00326)	0.9278*** (0.003279)	0.9161*** (0.002978)	0.9193*** (0.002936)	-0.01038*** (0.003904)	0.004958 (0.004149)	-0.01172*** (0.003758)	0.003217 (0.003618)
Respondents - Nonrespondents	-0.0063** (0.00276)	0.001701 (0.002986)	-0.006302** (0.003151)	0.00851*** (0.00275)	0.004838* (0.002746)	0.008001** (0.00387)	-0.008003* (0.004415)	0.01481*** (0.003744)	-0.003672 (0.003751)
Foreign Born									
Respondents	0.09922*** (0.001878)	0.1047*** (0.002144)	0.1076*** (0.002112)	0.1026*** (0.001871)	0.1034*** (0.001982)	0.005461** (0.002209)	0.002927 (0.002066)	-0.004964** (0.002197)	0.0007353 (0.001967)
Nonrespondents	0.09121*** (0.003643)	0.1001*** (0.003738)	0.1034*** (0.003809)	0.1169*** (0.003825)	0.1088*** (0.003418)	0.008914* (0.0046)	0.003289 (0.0045)	0.01346*** (0.004495)	-0.008098* (0.004465)
Respondents - Nonrespondents	0.008009** (0.003486)	0.004556 (0.003492)	0.004194 (0.003585)	-0.01423*** (0.003617)	-0.005393 (0.00337)	-0.003453 (0.004518)	-0.0003617 (0.004851)	-0.01842*** (0.00446)	0.008833* (0.004806)
Non-citizen									
Respondents	0.05524*** (0.001219)	0.05677*** (0.00133)	0.05734*** (0.001518)	0.05528*** (0.001435)	0.05471*** (0.001446)	0.001521 (0.001573)	0.0005718 (0.001525)	-0.002053 (0.001691)	-0.0005696 (0.001364)
Nonrespondents	0.05356*** (0.002745)	0.04869*** (0.002596)	0.05125*** (0.002471)	0.05789*** (0.002579)	0.05779*** (0.00246)	-0.004872 (0.003497)	0.002559 (0.003051)	0.00664** (0.003174)	-0.00009749 (0.00302)
Respondents - Nonrespondents	0.001685 (0.002588)	0.008079*** (0.002392)	0.006091** (0.002462)	-0.002601 (0.002279)	-0.003074 (0.002209)	0.006393* (0.003347)	-0.001987 (0.00325)	-0.008693*** (0.003238)	-0.0004721 (0.003006)
Education									
High School Diploma (or above)									
Respondents	0.8832*** (0.003073)	0.8726*** (0.003014)	0.8666*** (0.003197)	0.8635*** (0.003497)	0.8503*** (0.003258)	-0.01064*** (0.0038)	-0.006029 (0.003962)	-0.00308 (0.004122)	-0.01324*** (0.003958)
Nonrespondents	0.8781*** (0.008419)	0.8944*** (0.007178)	0.8497*** (0.00854)	0.8167*** (0.007535)	0.8043*** (0.007526)	0.01629 (0.01038)	-0.0447*** (0.01057)	-0.03304*** (0.01145)	-0.01234 (0.009799)
Respondents - Nonrespondents	0.005123 (0.009429)	-0.02181*** (0.007941)	0.01687* (0.008876)	0.04683*** (0.008142)	0.04593*** (0.008428)	-0.02693** (0.01141)	0.03867*** (0.01176)	0.02996** (0.01232)	-0.0008912 (0.01119)
Bachelor's Degree (or above)									
Respondents	0.3523*** (0.005294)	0.3491*** (0.005068)	0.3565*** (0.004909)	0.3645*** (0.005132)	0.3434*** (0.005264)	-0.003183 (0.005653)	0.00742 (0.005655)	0.008027 (0.005252)	-0.02111*** (0.005831)
Nonrespondents	0.324*** (0.01196)	0.3469*** (0.01238)	0.3129*** (0.01194)	0.2836*** (0.009781)	0.2758*** (0.01018)	0.02286 (0.01511)	-0.03393** (0.0161)	-0.02933** (0.01258)	-0.007852 (0.01271)
Respondents - Nonrespondents	0.02825** (0.01256)	0.002204 (0.01207)	0.04356*** (0.01212)	0.08091*** (0.01008)	0.06766*** (0.01017)	-0.02605* (0.01574)	0.04135** (0.01732)	0.03735*** (0.01396)	-0.01326 (0.01432)

Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This table shows the summary statistics for respondents and nonrespondents in the full CPS ASEC sample conditional on linkage to the source linked data set. Race and Hispanic-origin information is from the 2010 decennial census, citizenship information is from the Numident, and education information is from the ACS. Standard errors are shown in parenthesis. ***, **, and * indicate statistical significance at the 1-, 5-, and 10-percent levels respectively; but asterisks are only shown for differences as all estimates for respondents and nonrespondents are significant at the 1-percent level.

Table 4: Administrative Income for Linked CPS ASEC Respondent and Nonrespondent Households

Characteristic	Year					Difference			
	2017 (1)	2018 (2)	2019 (3)	2020 (4)	2021 (5)	2018-2017 (6)	2019-2018 (7)	2020-2019 (8)	2021-2020 (9)
W2									
Mean									
Respondents	96,360*** (1,391)	94,680*** (1,003)	97,100*** (1,141)	100,700*** (1,144)	101,700*** (1,607)	-1,677 (1,325)	2,421* (1,416)	3,615*** (1,252)	1,015 (1,753)
Nonrespondents	94,710*** (2,182)	95,610*** (1,880)	96,910*** (2,215)	93,880*** (2,732)	96,010*** (3,959)	900 (2,626)	1,297 (2,502)	-3,028 (3,462)	2,132 (4,616)
Respondents - Nonrespondents	1,645 (2,528)	-932 (1,943)	193 (2,318)	6,836** (2,861)	5,719 (3,964)	-2,577 (3,019)	1,125 (2,862)	6,643* (3,710)	-1,117 (4,731)
10th Percentile									
Respondents	11,840*** (235)	11,810*** (273)	11,480*** (244)	13,250*** (242)	11,210*** (248)	-28 (328)	-330 (329)	1,776*** (316)	-2,044*** (322)
Nonrespondents	12,710*** (571)	12,920*** (464)	13,150*** (556)	12,880*** (464)	11,020*** (501)	210 (687)	240 (678)	-277 (720)	-1,861*** (695)
Respondents - Nonrespondents	-870 (568)	-1,107** (527)	-1,677*** (581)	376 (504)	194 (529)	-238 (741)	-569 (776)	2,053*** (987)	-182 (755)
25th Percentile									
Respondents	32,160*** (356)	32,280*** (365)	32,530*** (355)	34,840*** (310)	33,110*** (374)	127 (424)	245 (411)	2,307*** (399)	-1,720*** (451)
Nonrespondents	32,180*** (667)	32,860*** (711)	34,190*** (739)	31,500*** (518)	30,650*** (599)	679 (967)	1,322 (959)	-2,689*** (862)	-843 (696)
Respondents - Nonrespondents	-27 (672)	-580 (761)	-1,657** (796)	3,339*** (547)	2,461*** (635)	-553 (1,014)	-1,077 (1,047)	4,996*** (987)	-878 (831)
Median									
Respondents	67,300*** (497)	67,320*** (540)	68,200*** (493)	71,730*** (574)	70,010*** (552)	-18 (572)	881 (557)	3,523*** (603)	-1,714** (689)
Nonrespondents	64,710*** (947)	66,200*** (888)	68,710*** (885)	64,140*** (787)	64,770*** (813)	1,486 (1,196)	2,514** (1,140)	-4,571*** (1,092)	627 (1,062)
Respondents - Nonrespondents	2,593** (1,013)	1,125 (927)	-508 (917)	7,586*** (793)	5,245*** (859)	-1,468 (1,300)	-1,632 (1,253)	8,094*** (1,281)	-2,341** (1,118)
75th Percentile									
Respondents	120,100*** (903)	119,600*** (977)	121,800*** (835)	126,200*** (1,029)	126,300*** (978)	-481 (1,007)	2,184** (982)	4,447*** (998)	42 (1,096)
Nonrespondents	114,000*** (1,697)	118,700*** (1,711)	118,700*** (1,743)	114,200*** (1,568)	115,200*** (1,527)	4,743** (2,116)	-18 (2,154)	-4,439** (2,093)	979 (1,966)
Respondents - Nonrespondents	6,129*** (1,821)	905 (1,764)	3,107* (1,820)	11,990*** (1,560)	11,050*** (1,620)	-5,224** (2,308)	2,202 (2,293)	8,886*** (2,350)	-938 (2,212)
90th Percentile									
Respondents	190,700*** (1,848)	189,200*** (1,698)	192,000*** (1,794)	200,200*** (1,877)	203,900*** (1,860)	-1,439 (1,757)	2,772 (2,043)	8,178*** (1,972)	3,764* (2,116)
Nonrespondents	186,300*** (4,885)	192,100*** (3,353)	189,700*** (3,467)	182,800*** (2,561)	186,600*** (2,741)	5,822 (5,348)	-2,431 (4,141)	-6,954* (3,827)	3,880 (3,405)
Respondents - Nonrespondents	4,329 (4,945)	-2,932 (3,394)	2,271 (3,521)	17,400*** (2,785)	17,290*** (2,873)	-7,261 (5,433)	5,203 (4,649)	15,130*** (4,396)	-116 (3,699)
1040									
Mean									
Respondents	116,800*** (3,022)	113,100*** (2,315)	115,200*** (1,514)	125,100*** (3,173)	124,200*** (2,798)	-3,733 (3,555)	2,102 (2,583)	9,947*** (3,381)	-912 (4,118)
Nonrespondents	131,200*** (9,067)	116,700*** (4,281)	115,900*** (9,274)	118,000*** (9,132)	120,600*** (5,445)	-14,460 (9,697)	-825 (9,632)	2,125 (11,600)	2,565 (10,300)
Respondents - Nonrespondents	-14,400 (9,193)	-3,667 (4,638)	-740 (9,297)	7,083 (9,688)	3,605 (6,235)	10,730 (10,630)	2,927 (10,510)	7,822 (12,450)	-3,477 (11,440)
10th Percentile									
Respondents	16,010*** (237)	16,090*** (215)	16,560*** (212)	16,830*** (243)	15,770*** (276)	77 (283)	469* (269)	276 (288)	-1,060*** (339)
Nonrespondents	15,880*** (509)	15,720*** (505)	17,250*** (460)	15,290*** (316)	14,140*** (474)	-163 (672)	1,530** (678)	-1,962*** (539)	-1,152** (546)
Respondents - Nonrespondents	129 (525)	369 (538)	-693 (465)	1,545*** (383)	1,638*** (523)	240 (715)	-1,061 (725)	2,238*** (600)	93 (642)
25th Percentile									
Respondents	36,830*** (359)	36,730*** (416)	37,510*** (345)	39,240*** (357)	38,170*** (369)	-102 (429)	786* (428)	1,727*** (424)	-1,068** (436)
Nonrespondents	35,250*** (833)	36,360*** (656)	37,100*** (669)	33,280*** (588)	33,490*** (609)	1,104 (977)	746 (929)	-3,829*** (865)	219 (775)
Respondents - Nonrespondents	1,576* (814)	370 (735)	410 (681)	5,966*** (645)	4,679*** (642)	-1,206 (1,032)	40 (968)	5,556*** (945)	-1,287 (907)
Median									
Respondents	75,510*** (587)	75,100*** (628)	76,120*** (585)	79,610*** (651)	78,690*** (638)	-404 (661)	1,019 (622)	3,487*** (726)	-924 (670)
Nonrespondents	71,910*** (1,024)	73,220*** (877)	72,840*** (935)	68,690*** (843)	70,910*** (1,072)	1,306 (1,261)	-372 (1,176)	-4,155** (1,171)	2,220* (1,197)
Respondents - Nonrespondents	3,598*** (1,084)	1,888* (987)	3,279*** (927)	10,920*** (933)	7,778*** (1,112)	-1,711 (1,360)	1,392 (1,274)	7,642*** (1,351)	-3,143** (1,347)
75th Percentile									
Respondents	132,300*** (1,011)	129,700*** (949)	133,100*** (1,024)	137,900*** (1,006)	138,600*** (1,180)	-2,632** (1,072)	3,401*** (1,046)	4,772*** (1,142)	756 (1,178)
Nonrespondents	127,000*** (1,944)	129,500*** (1,804)	127,900*** (1,556)	122,700*** (1,754)	126,600*** (1,767)	2,447 (2,221)	-1,592 (2,156)	-5,197** (2,052)	3,880 (2,219)
Respondents - Nonrespondents	5,328*** (2,018)	249 (1,821)	5,242*** (1,682)	15,210*** (1,807)	12,090*** (1,791)	-5,079** (2,294)	4,993** (2,298)	9,969*** (2,356)	-3,124 (2,322)
90th Percentile									
Respondents	218,600*** (2,073)	215,000*** (2,112)	218,400*** (2,137)	227,300*** (2,102)	232,200*** (2,490)	-3,603* (2,185)	3,439 (2,254)	8,844*** (2,501)	4,941* (2,762)
Nonrespondents	215,900*** (5,107)	217,900*** (3,349)	220,400*** (4,527)	204,400*** (3,004)	217,800*** (4,129)	1,988 (5,846)	2,474 (5,191)	-16,030*** (4,777)	13,400*** (4,885)
Respondents - Nonrespondents	2,657 (5,123)	-2,934 (3,501)	-1,969 (4,254)	22,900*** (3,262)	14,440*** (4,143)	-5,591 (6,047)	965 (5,284)	24,870*** (5,114)	-8,463* (4,985)

Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This table shows income estimates and the difference in income by address between respondents and nonrespondents in the full CPS ASEC sample. The top half shows total W-2 earnings at that address in the reference year of the survey. The bottom half shows total 1040 AGI in the prior year for linked individuals at the survey address. A value of greater than zero indicates higher income for respondents than nonrespondents for that statistic and year. Standard errors are shown in parenthesis. ***, **, and * indicate statistical significance at the 1-, 5-, and 10-percent levels respectively, but asterisks are only shown for differences as all estimates for respondents and nonrespondents are significant at the 1-percent level.

Table 5: Probability of Response by Total W-2 Earnings at Address

A. No Controls												
	Regression						Comparison					
	2017 (1)	2018 (2)	2019 (3)	Pooled (2017-2019) (4)	2020 (5)	2021 (6)	2018 - 2017 (7)	2019 - 2018 (8)	2020 - 2019 (9)	2021 - 2020 (10)	2020 - Pooled (11)	2021 - Pooled (12)
Has W-2	-0.02041*** (0.004426)	-0.01102** (0.004379)	-0.01558*** (0.004382)	-0.01510*** (0.002722)	-0.008350 (0.005627)	-0.01194** (0.005252)	0.009390 (0.005829)	-0.004564 (0.006215)	0.007234 (0.007065)	-0.003586 (0.007578)	0.006753 (0.006194)	0.003167 (0.005695)
0-25,000	0.01120** (0.005547)	0.006967 (0.005630)	0.02195*** (0.006053)	0.01342*** (0.003394)	0.007422 (0.006408)	0.008095 (0.006765)	-0.004238 (0.007721)	0.01499* (0.008151)	-0.01453* (0.008650)	0.0006725 (0.009088)	-0.006003 (0.007250)	-0.005330 (0.007719)
50,000-75,000	0.0009069 (0.005627)	0.00003118 (0.005166)	0.006789 (0.005358)	0.002484 (0.002987)	0.01885*** (0.007015)	0.008741 (0.006287)	-0.0008757 (0.007553)	0.006758 (0.007876)	0.01207 (0.008962)	-0.01011 (0.009867)	0.01637** (0.007476)	0.006257 (0.007018)
75,000-100,000	0.009000 (0.005899)	0.003237 (0.005557)	0.003223 (0.006942)	0.005085 (0.003607)	0.005085 (0.006897)	0.02771*** (0.007091)	-0.005763 (0.008390)	-0.00001407 (0.009157)	0.02448** (0.009901)	-0.007601 (0.009438)	0.02262*** (0.007651)	0.01502** (0.007634)
100,000-150,000	0.01469*** (0.005255)	0.007415 (0.005291)	0.01050* (0.005703)	0.01057*** (0.003530)	0.03455*** (0.007066)	0.02184*** (0.006231)	-0.007277 (0.007270)	0.003085 (0.007037)	0.02405*** (0.008945)	-0.01270 (0.008908)	0.02398*** (0.007856)	0.01127* (0.006837)
150,000-200,000	0.02980*** (0.007087)	0.007100 (0.007363)	0.01817*** (0.006910)	0.01781*** (0.004047)	0.04749*** (0.008334)	0.04178*** (0.008240)	-0.02270** (0.009947)	0.01107 (0.01028)	0.02932*** (0.01079)	-0.005717 (0.01103)	0.02968*** (0.009382)	0.02396** (0.009316)
ge 200,000	0.01432** (0.007118)	0.0005016 (0.007063)	0.01536** (0.007527)	0.01004** (0.004545)	0.06031*** (0.007713)	0.04722*** (0.007207)	-0.01382 (0.008620)	0.01486 (0.009850)	0.04495*** (0.01088)	-0.01309 (0.01020)	0.05026*** (0.008915)	0.03717*** (0.008184)
Constant	0.8761*** (0.002505)	0.8647*** (0.002272)	0.8439*** (0.002386)	0.8608*** (0.001492)	0.7577*** (0.003348)	0.7958*** (0.003090)	-0.01142*** (0.003120)	-0.02076*** (0.003107)	-0.08620*** (0.003672)	0.03802*** (0.003857)	-0.1031*** (0.003552)	-0.06504*** (0.003416)
R-Squared	0.00	0.00	0.00	0.00	0.00	0.00						
Observations	81,000	79,500	82,000	242,000	79,500	79,500						

B. With Full Controls												
	Regression						Comparison					
	2017 (1)	2018 (2)	2019 (3)	Pooled (2017-2019) (4)	2020 (5)	2021 (6)	2018 - 2017 (7)	2019 - 2018 (8)	2020 - 2019 (9)	2021 - 2020 (10)	2020 - Pooled (11)	2021 - Pooled (12)
0-25,000	0.01018* (0.005578)	0.005000 (0.005474)	0.01832*** (0.006123)	0.01141*** (0.003368)	0.002713 (0.006411)	0.003553 (0.006734)	-0.005177 (0.007429)	0.01332 (0.008298)	-0.01561* (0.008719)	0.0008397 (0.009283)	-0.008692 (0.007324)	-0.007852 (0.007661)
50,000-75,000	0.001130 (0.005528)	-0.0009626 (0.005290)	0.004465 (0.005312)	0.001336 (0.002945)	0.01677** (0.006875)	0.008142 (0.006326)	-0.002093 (0.007566)	0.005427 (0.007948)	0.01231 (0.008925)	-0.008632 (0.009730)	0.01544*** (0.007378)	0.006807 (0.007008)
75,000-100,000	0.008198 (0.005900)	0.001123 (0.005864)	-0.0009271 (0.007169)	0.002470 (0.003663)	0.02398*** (0.007081)	0.01954*** (0.007267)	-0.007075 (0.008594)	-0.002051 (0.009486)	0.02491** (0.01040)	-0.004439 (0.009479)	0.02151*** (0.007874)	0.01707** (0.007840)
100,000-150,000	0.01294** (0.005431)	0.004141 (0.005503)	0.003807 (0.006023)	0.006430* (0.003519)	0.02985*** (0.006625)	0.02190*** (0.006361)	-0.008795 (0.007720)	-0.0003348 (0.007719)	0.02604*** (0.009085)	-0.007942 (0.008640)	0.02342*** (0.007455)	0.01548** (0.007086)
150,000-200,000	0.02756*** (0.007336)	0.002676 (0.007430)	0.01100 (0.007429)	0.01290*** (0.004103)	0.04216*** (0.008249)	0.04092*** (0.008395)	-0.02488** (0.01032)	0.008320 (0.01064)	0.03116*** (0.01123)	-0.001241 (0.01104)	0.02926*** (0.009334)	0.02802*** (0.009690)
ge 200,000	0.01027 (0.007289)	-0.003848 (0.007636)	0.006404 (0.008027)	0.003764 (0.004649)	0.05138*** (0.008167)	0.04388*** (0.007671)	-0.01412 (0.009073)	0.01025 (0.01101)	0.04497*** (0.01159)	-0.007500 (0.01035)	0.04761*** (0.009181)	0.04011*** (0.008672)
R-Squared	0.01	0.01	0.01	0.01	0.02	0.02						
Observations	81,000	79,500	82,000	242,000	79,500	79,500						

Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This table shows the coefficient estimates from a regression of housing unit response on W-2 earnings at that address for the full CPS ASEC sample. Positive values indicate individuals in that income range are more likely to respond than the baseline group (25,000–50,000). Panel A shows the results without controls for linkage rates and available demographic and socioeconomic information (such as race, Hispanic origin, citizenship, etc.). Panel B shows the results with those controls included. ***, **, and * indicate statistical significance at the 1-, 5-, and 10-percent levels respectively.

Table 6: Two-Stage Entropy Balance Reweighting Procedure

Stage/Step	Moment Variables	Moment Sample	Reweighted Sample
1. Housing-unit level	Linked survey, administrative, and census variables	Non-vacant housing units in March Basic CPS (respondents and nonrespondents)	Respondent housing units
2. Person level			
A. Preserve distribution of housing unit characteristics	Linked survey, administrative, and census variables	Householders, using the housing-unit level weights from Stage 1	All individuals, weighted by 1/number of people in each household
B. Spousal equivalence	Linked survey, administrative, and census variables	Married couples and cohabiting partners	Married couples and cohabiting partners
C. External population targets	State-level population estimates by race, Hispanic-origin, gender, and age	External population estimates	All individuals
D. Full CPS ASEC only: Match distribution of household characteristics in March Basic Sample	Subset of linked survey, administrative, and census variables and state-level population controls	Householders and householder partners in the March Basic File	Householders and householder partners in the full CPS ASEC sample

Notes: This table describes the two-stage entropy balance reweighting procedure. In the first stage, respondent housing units are reweighted to control for selection into response. This is done by reweighting them to match the characteristics of the target population – all nonvacant housing units in sample. In the second stage, we estimate individual weights that preserve the distribution of housing-unit characteristics from the first stage, while also matching external population totals and approximating the spousal equivalence of weights that are a part of the existing CPS ASEC weights.

Table 7: Before Entropy Balance Weighting — Linked Data Summary Statistics using the Base Weights

Characteristic	March Base Weights					Full CPS ASEC Respondents (No Oversample Adjustment) - March Base Weights					March Basic CPS Respondents - March Base Weights				
	2017	2018	2019	2020	2021	2017	2018	2019	2020	2021	2017	2018	2019	2020	2021
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Percentage of Housing Units															
Age (At Least One Individual in Range)															
18-24	14.97 (0.18)	14.67 (0.16)	14.31 (0.16)	14.48 (0.16)	14.21 (0.17)	1.298*** (0.122)	1.51*** (0.125)	1.239*** (0.114)	0.962*** (0.138)	0.839*** (0.121)	-0.014 (0.060)	-0.005 (0.067)	-0.092 (0.070)	-0.446*** (0.100)	-0.398*** (0.088)
25-34	22.24 (0.21)	22.23 (0.20)	21.55 (0.21)	22.26 (0.22)	22.53 (0.21)	1.326*** (0.138)	1.133*** (0.135)	0.981*** (0.135)	0.696*** (0.153)	0.853*** (0.153)	-0.23*** (0.078)	-0.195*** (0.075)	-0.316*** (0.091)	-0.546*** (0.109)	-0.546*** (0.108)
35-44	20.72 (0.20)	20.64 (0.19)	20.30 (0.19)	20.85 (0.18)	21.06 (0.20)	2.845*** (0.147)	2.828*** (0.136)	2.566*** (0.144)	2.53*** (0.139)	2.632*** (0.144)	-0.207*** (0.078)	-0.176* (0.079)	-0.200* (0.084)	-0.254* (0.106)	-0.201* (0.096)
45-54	21.93 (0.22)	21.29 (0.21)	20.33 (0.18)	20.00 (0.21)	19.69 (0.19)	1.263*** (0.139)	1.537*** (0.136)	1.326*** (0.134)	1.812*** (0.153)	1.681*** (0.144)	-0.232*** (0.085)	-0.204* (0.091)	-0.208* (0.084)	-0.009 (0.117)	-0.081 (0.097)
55-64	22.36 (0.20)	21.73 (0.19)	21.65 (0.20)	21.96 (0.20)	21.55 (0.19)	-1.221*** (0.128)	-0.83*** (0.127)	-0.975*** (0.118)	0.023 (0.129)	-0.387*** (0.128)	0.17* (0.077)	0.292*** (0.085)	0.436*** (0.084)	1.187*** (0.118)	0.587*** (0.105)
65 Plus	27.08 (0.23)	28.11 (0.24)	27.86 (0.21)	28.17 (0.21)	29.02 (0.22)	-2.328*** (0.127)	-1.988*** (0.146)	-1.453*** (0.124)	0.084 (0.157)	-0.765*** (0.124)	1.542*** (0.077)	1.685*** (0.081)	1.936*** (0.090)	3.321*** (0.125)	2.87*** (0.112)
Native Born	76.02 (0.27)	75.55 (0.25)	74.29 (0.26)	75.32 (0.28)	75.73 (0.26)	-1.307*** (0.153)	-0.739*** (0.154)	-0.725*** (0.147)	0.775*** (0.168)	0.256 (0.170)	0.531*** (0.078)	0.801*** (0.090)	0.888*** (0.094)	2.422*** (0.123)	1.996*** (0.109)
Education (Most Educated Linked Individual)															
High School Diploma															
19.19 (0.20)	18.89 (0.20)	18.28 (0.19)	18.37 (0.20)	18.26 (0.18)	-0.174 (0.114)	0.043 (0.110)	0.205* (0.115)	0.73*** (0.136)	0.422*** (0.125)	0.571*** (0.071)	0.66*** (0.071)	0.84*** (0.073)	1.386*** (0.109)	1.257*** (0.089)	
Bachelor's Degree															
7.60 (0.14)	7.54 (0.13)	7.40 (0.11)	7.49 (0.13)	7.17 (0.12)	0.058 (0.078)	0.044 (0.068)	0.251*** (0.075)	0.636*** (0.083)	0.401*** (0.085)	0.299*** (0.047)	0.29*** (0.045)	0.452*** (0.045)	0.903*** (0.068)	0.721*** (0.054)	
Graduate Degree															
3.29 (0.09)	3.12 (0.09)	3.03 (0.09)	3.11 (0.09)	3.06 (0.08)	0.048 (0.054)	0.058 (0.048)	0.134*** (0.047)	0.329*** (0.054)	0.195*** (0.050)	0.187*** (0.031)	0.134*** (0.029)	0.22*** (0.031)	0.406*** (0.042)	0.348*** (0.037)	
Hispanic															
9.95 (0.16)	10.04 (0.19)	10.21 (0.18)	10.50 (0.19)	10.60 (0.20)	3.996*** (0.143)	4.056*** (0.139)	3.594*** (0.124)	3.346*** (0.134)	3.32*** (0.138)	0.073 (0.059)	0.039 (0.066)	0.003 (0.061)	-0.365*** (0.094)	-0.276*** (0.083)	
Race															
Black															
10.53 (0.21)	10.61 (0.19)	10.39 (0.20)	10.63 (0.20)	11.11 (0.22)	1.703*** (0.133)	1.795*** (0.141)	1.4*** (0.116)	1.367*** (0.148)	1.374*** (0.133)	-0.294*** (0.060)	-0.334*** (0.065)	-0.531*** (0.064)	-0.681*** (0.099)	-0.511*** (0.071)	
White															
62.82 (0.31)	62.64 (0.33)	61.55 (0.33)	61.43 (0.32)	61.71 (0.30)	-2.854*** (0.173)	-2.696*** (0.181)	-2.135*** (0.173)	-0.408* (0.202)	-1.248*** (0.182)	1.073*** (0.098)	1.377*** (0.099)	1.792*** (0.102)	3.475*** (0.137)	2.856*** (0.111)	
Income Statistics															
W-2 Earnings Percentile															
10th															
11,970 (238)	11,900 (258)	11,790 (251)	13,200 (230)	11,190 (253)	532*** (157)	600*** (151)	443*** (165)	934*** (174)	937*** (175)	-114 (95)	-257* (105)	-290*** (101)	22 (152)	-5 (132)	
25th															
32,280 (369)	32,430 (337)	32,870 (356)	34,070 (312)	32,450 (364)	662*** (219)	679*** (202)	676*** (232)	1,926*** (260)	2,190*** (277)	-47 (132)	-186 (160)	-148 (152)	778*** (178)	588*** (203)	
50th															
66,770 (451)	67,070 (520)	68,360 (483)	70,060 (561)	68,940 (519)	1,965*** (324)	1,725*** (308)	1,442*** (338)	3,821*** (395)	3,177*** (367)	426* (202)	93 (195)	111 (220)	1,788*** (280)	1,297*** (261)	
75th															
119,100 (748)	119,600 (947)	121,500 (814)	123,500 (989)	123,800 (924)	2,604*** (509)	1,895*** (582)	2,002*** (570)	5,611*** (677)	5,239*** (630)	995*** (332)	-41 (334)	889*** (342)	3,064*** (484)	2,876*** (431)	
90th															
190,100 (1,646)	189,300 (1,662)	191,700 (1,713)	196,500 (1,778)	200,900 (1,844)	1,764* (1033)	1,444 (1045)	1,960* (1163)	7,096*** (1233)	6,163*** (1363)	467 (568)	-902 (585)	1,062 (658)	4,140*** (808)	4,276*** (874)	
1040 AGI Percentile															
10th															
15,940 (237)	16,010 (218)	16,710 (218)	16,430 (218)	15,480 (275)	612*** (133)	585*** (139)	345*** (178)	1,230*** (185)	1,074*** (185)	-45 (80)	18 (93)	-113 (92)	362* (146)	418*** (158)	
25th															
36,560 (377)	36,860 (389)	37,550 (334)	37,900 (349)	37,270 (363)	691*** (232)	426* (193)	387* (206)	2,245*** (282)	1,740*** (245)	159 (136)	70 (145)	161 (121)	1,644*** (205)	1,186*** (205)	
50th															
75,030 (562)	74,810 (570)	75,550 (562)	77,190 (575)	77,190 (603)	783* (351)	787* (318)	1,495*** (345)	3,874*** (405)	2,650*** (405)	501* (210)	257 (192)	801*** (213)	2,884*** (316)	1,821*** (298)	
75th															
131,700 (908)	129,500 (916)	132,400 (988)	134,900 (1,033)	136,500 (1,076)	697 (577)	1,315* (577)	1,207* (624)	4,136*** (640)	3,798*** (633)	900* (374)	-247 (324)	1,375*** (402)	3,801*** (453)	3,025*** (442)	
90th															
218,600 (2,286)	215,300 (1,973)	218,800 (2,298)	223,400 (2,025)	230,700 (2,293)	-901 (1443)	-1,451 (1128)	-1,188 (1376)	5,807*** (1312)	2,912* (1460)	619 (748)	-1,032 (680)	239 (751)	5,392*** (1015)	4,089*** (1035)	

Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This table shows various demographic and socioeconomic summary statistics at the household level using the base weights with no adjustment for oversampling or selection into response. In Columns (1)-(5), we show estimates using the base weights on the March Basic CPS sample, including responding and nonrespondent housing units. These estimates should best represent the distribution of the linked characteristics in the population and are therefore the target distribution for the first-stage EBW adjustment. In Columns (5)-(8), we show the difference between the estimates for respondents in the full CPS ASEC Sample and using the base weights for all nonvacant units, as in (1)-(5). Significant differences in (6)-(10) reflect both oversampling by observable characteristics and nonrandom nonresponse. In Columns (11)-(15), we show the difference between the estimates for respondents in the March Basic CPS ASEC Sample and using the base weights for all nonvacant units, as in (1)-(5), which should reflect nonrandom nonresponse. Standard errors are shown in parenthesis. Education requires a link to the ACS, and the reported values are unconditional. ***, **, and * indicate statistical significance at the 1-, 5-, and 10-percent levels respectively, but asterisks are only shown for differences as all estimates for respondents and nonrespondents are significant at the 1-percent level.

Table 8: After Entropy Balance Weighting — Linked Data Summary Statistics using the Base Weights

Characteristic	Full EBW - March Base Weights					March EBW - March Base Weights				
	2017 (1)	2018 (2)	2019 (3)	2020 (4)	2021 (5)	2017 (6)	2018 (7)	2019 (8)	2020 (9)	2021 (10)
Percentage of Housing Units										
Age (At Least One Individual in Range)										
18-24	-0.039*** (0.003)	0.008*** Z	0.007*** Z	0.013*** Z	0.007*** Z	-0.039*** (0.003)	0.007*** Z	0.006*** Z	0.014*** Z	0.007*** Z
25-34	0.027*** (0.001)	0.003*** (0.001)	0.009*** Z	0.009*** (0.001)	0.012*** Z	0.036*** (0.001)	0.002* (0.001)	0.006*** Z	0.006*** (0.001)	0.014*** Z
35-44	0.03*** (0.001)	0.009*** Z	0.009*** (0.001)	0.016*** Z	0.011*** (0.001)	0.032*** (0.001)	0.007*** Z	0.009*** Z	0.019*** (0.001)	0.013*** Z
45-54	0.033*** (0.001)	0.007*** (0.001)	0.01*** Z	Z (0.001)	-0.001* (0.001)	0.032*** (0.001)	0.01*** Z	0.008*** Z	-0.001 (0.001)	-0.007*** (0.001)
55-64	0.022*** (0.002)	0.011*** Z	-0.007*** (0.001)	0.011*** (0.001)	0.012*** Z	0.022*** (0.002)	0.009*** Z	0.001 (0.001)	0.014*** (0.001)	0.012*** Z
65 Plus	0.023*** (0.002)	0.003*** (0.001)	0.008*** (0.001)	0.018*** Z	0.007*** Z	0.021*** (0.002)	Z (0.001)	0.007*** Z	0.024*** (0.001)	0.006*** (0.001)
Native Born	0.024*** (0.001)	0.019*** (0.001)	0.009*** (0.001)	0.026*** (0.001)	0.021*** (0.001)	0.013*** (0.001)	0.012*** (0.001)	0.015*** (0.001)	0.03*** (0.001)	0.019*** (0.001)
Education (Most Educated Linked Individual)										
High School Diploma	0.091 (0.065)	-0.007 (0.068)	0.017 (0.072)	0.057 (0.073)	0.054 (0.069)	0.054 (0.065)	-0.088 (0.068)	0.006 (0.072)	0.052 (0.073)	0.082 (0.069)
Bachelor's Degree	0.172* (0.102)	0.036 (0.104)	0.161* (0.095)	0.178* (0.099)	0.162 (0.102)	0.081 (0.102)	-0.022 (0.104)	0.077 (0.095)	0.183* (0.099)	0.154 (0.101)
Graduate Degree	0.093 (0.078)	0.071 (0.079)	0.099 (0.077)	0.084 (0.080)	0.069 (0.075)	0.037 (0.077)	0.008 (0.079)	0.055 (0.077)	0.06 (0.080)	0.08 (0.074)
Hispanic	0.013*** (0.001)	0.005*** Z	0.002*** Z	0.009*** Z	0.007*** Z	0.018*** (0.001)	0.004*** Z	Z Z	0.011*** Z	0.007*** Z
Race										
Black	-0.007*** (0.002)	0.005*** Z	0.002*** Z	0.005*** Z	0.007*** Z	-0.016*** (0.002)	0.004*** Z	0.005*** Z	0.008*** Z	0.007*** Z
White	0.029*** (0.002)	0.015*** (0.001)	0.005*** (0.001)	0.028*** (0.001)	0.018*** (0.001)	0.029*** (0.002)	0.007*** (0.001)	0.009*** (0.001)	0.031*** (0.001)	0.015*** (0.001)
Income Statistics										
W-2 Earnings Percentile										
10th	86 (123)	131 (138)	-8 (128)	-6 (130)	-11 (134)	166 (122)	-8 (139)	1 (127)	60 (129)	72 (132)
25th	116 (146)	-3 (127)	-43 (130)	-9 (131)	159 (135)	182 (144)	7 (129)	46 (130)	62 (130)	98 (135)
50th	200 (144)	147 (161)	-147 (155)	-18 (180)	-120 (164)	109 (144)	75 (160)	-84 (156)	185 (181)	-86 (164)
75th	309 (282)	-230 (290)	-77 (297)	263 (270)	395 (291)	256 (275)	36 (285)	164 (296)	246 (272)	262 (288)
90th	156 (409)	257 (431)	-81 (463)	-18 (284)	-43 (147)	47 (367)	84 (433)	220 (447)	187 (280)	17 (119)
1040 AGI Percentile										
10th	136 (114)	65 (102)	99 (120)	51 (134)	13 (162)	128 (114)	67 (103)	87 (121)	12 (134)	46 (162)
25th	123 (135)	-211 (146)	23 (150)	50 (149)	92 (152)	83 (136)	-55 (146)	218 (152)	165 (149)	105 (151)
50th	86*** (17)	56 (82)	61 (108)	58 (92)	73 (117)	91*** (23)	46 (81)	54 (108)	-30 (91)	102 (117)
75th	210 (313)	88 (327)	-211 (340)	90 (348)	28 (336)	331 (307)	-177 (328)	256 (335)	280 (348)	111 (336)
90th	-142 (819)	-109 (565)	-654 (747)	-962 (708)	-1089 (864)	-34 (805)	-165 (598)	-700 (763)	-408 (697)	-152 (870)

Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This table shows comparisons of summary statistics at the household level using the EBW to the base-weighted March Basic sample (including respondent and nonrespondent households). In Columns (1)-(5), we compare the EBW estimates for respondents in the March Basic sample. Columns (6)-(10) compare the EBW estimates for respondents in the March Basic CPS ASEC sample. Standard errors are shown in parenthesis. Education requires a link to the ACS, and the reported values are unconditional. Z indicates an estimate rounds to 0. ***, **, and * indicate statistical significance at the 1-, 5-, and 10-percent levels respectively.

Table 9: Percentage of People by Characteristic using Survey and Alternative Weights

Characteristic	Survey Weights					EBW - Survey					March EBW - Survey				
	2017 (1)	2018 (2)	2019 (3)	2020 (4)	2021 (5)	2017 (6)	2018 (7)	2019 (8)	2020 (9)	2021 (10)	2017 (11)	2018 (12)	2019 (13)	2020 (14)	2021 (15)
Race															
Asian	6.5 (0.04)	6.7 (0.05)	6.8 (0.05)	6.9 (0.05)	7.0 (0.06)	-0.1 (0.1)	Z (0.1)	-0.2*** (0.1)	-0.1 (0.1)	-0.1* (0.1)	-0.1 (0.1)	Z (0.1)	-0.2*** (0.1)	-0.1 (0.1)	-0.1* (0.1)
American Indian or Alaskan Native	2.1 (0.05)	2.1 (0.06)	2.1 (0.06)	2.0 (0.05)	2.1 (0.06)	-0.4*** (0.1)	-0.3*** (0.1)	-0.2*** (0.1)	-0.2*** (0.1)	-0.2*** (0.1)	-0.4*** (0.1)	-0.3*** (0.1)	-0.2*** (0.1)	-0.2*** (0.1)	-0.2*** (0.1)
Black	14.3 (0.04)	14.4 (0.05)	14.5 (0.05)	14.6 (0.05)	14.7 (0.05)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)
Hawaiian or Pacific Islander	0.5 (0.03)	0.5 (0.04)	0.6 (0.04)	0.5 (0.04)	0.6 (0.03)	-0.1* (0.1)	-0.1* (0.1)	Z (0.1)	-0.1** (0.1)	Z (0.1)	-0.1* (0.1)	-0.1* (0.1)	Z (0.1)	-0.1** (0.1)	Z (0.1)
White	79.2 (0.06)	79.1 (0.07)	79.0 (0.06)	78.9 (0.06)	78.7 (0.06)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)
Hispanic	18.0 (0.06)	18.3 (0.07)	18.5 (0.06)	18.7 (0.06)	18.8 (0.06)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)	Z (0.1)
Citizenship															
Native	86.3 (0.12)	85.9 (0.13)	85.9 (0.13)	86.2 (0.13)	86.3 (0.12)	0.5*** (0.1)	0.5*** (0.1)	0.4*** (0.1)	0.2* (0.1)	0.3*** (0.1)	0.4*** (0.1)	0.4*** (0.1)	0.4*** (0.1)	0.2 (0.1)	0.3** (0.1)
Foreign-Born Citizen	6.4 (0.08)	6.8 (0.09)	6.9 (0.09)	7.0 (0.09)	6.9 (0.09)	-0.3*** (0.1)	-0.4*** (0.1)	-0.4*** (0.1)	-0.3*** (0.1)	-0.3*** (0.1)	-0.3*** (0.1)	-0.4*** (0.1)	-0.4*** (0.1)	-0.4*** (0.1)	-0.3*** (0.1)
Non-citizen	7.3 (0.09)	7.3 (0.10)	7.3 (0.11)	6.8 (0.10)	6.8 (0.10)	-0.2* (0.1)	-0.1 (0.1)	-0.1 (0.1)	0.1 (0.1)	Z (0.1)	-0.1 (0.1)	Z (0.1)	Z (0.1)	0.2* (0.1)	Z (0.1)
Education															
High School	10.4 (0.12)	10.2 (0.12)	9.9 (0.12)	9.1 (0.13)	8.9 (0.12)	Z (0.1)	0.1 (0.1)	0.1 (0.1)	0.4*** (0.1)	0.3** (0.1)	0.1 (0.1)	0.1 (0.1)	0.1 (0.1)	0.3*** (0.1)	0.3** (0.1)
High School	28.8 (0.21)	28.5 (0.20)	28.1 (0.21)	27.6 (0.20)	27.9 (0.20)	0.2 (0.2)	0.1 (0.2)	0.2 (0.2)	0.4* (0.2)	0.5*** (0.2)	0.5** (0.2)	0.5** (0.2)	0.3* (0.2)	0.6*** (0.2)	0.7*** (0.2)
Some College	26.6 (0.17)	26.3 (0.16)	25.9 (0.19)	25.8 (0.17)	25.4 (0.16)	0.2 (0.2)	0.2 (0.2)	0.2 (0.2)	0.2 (0.2)	0.2 (0.2)	Z (0.2)	-0.1 (0.2)	0.1 (0.2)	0.3* (0.2)	0.1 (0.2)
Bachelor's	21.3 (0.16)	21.9 (0.18)	22.6 (0.18)	23.4 (0.18)	23.5 (0.18)	-0.2 (0.2)	-0.2 (0.2)	-0.2 (0.2)	-0.6*** (0.2)	-0.5*** (0.2)	-0.3* (0.2)	-0.3 (0.2)	-0.3 (0.2)	-0.7*** (0.2)	-0.5*** (0.2)
Advanced Degree	12.8 (0.15)	13.1 (0.15)	13.5 (0.14)	14.1 (0.14)	14.4 (0.16)	-0.2 (0.1)	-0.2 (0.1)	-0.3** (0.1)	-0.4*** (0.1)	-0.4*** (0.1)	-0.3** (0.1)	-0.3** (0.1)	-0.3*** (0.1)	-0.5*** (0.1)	-0.6*** (0.1)

Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This table shows various demographic and socioeconomic summary statistics at the person level using the survey weights and EBW. In Columns (1)-(5), we show estimates using the official survey weights. In Columns (6)-(10), we show the difference between the Full EBW estimates and the survey. In Columns (11)-(15), we show the difference between the March EBW estimates and the survey. Standard errors are shown in parenthesis. Z indicates an estimate rounds to 0. ***, **, and * indicate statistical significance at the 1-, 5-, and 10-percent levels respectively, but asterisks are only shown for differences as all estimates for respondents and nonrespondents are significant at the 1-percent level.

Table 10: Household Income Estimates using Survey and Alternative Weights (in 2019 dollars)

Percentile	Survey Weights					Full CPS ASEC Sample (EBW)					Basic March CPS (March EBW)				
	2017 (1)	2018 (2)	2019 (3)	2020 (4)	2021 (5)	2017 (6)	2018 (7)	2019 (8)	2020 (9)	2021 (10)	2017 (11)	2018 (12)	2019 (13)	2020 (14)	2021 (15)
5	8,917 (126)	8,986 (146)	8,926 (155)	9,798 (180)	9,113 (182)	8,973 (38)	9,087 (41)	8,974 (41)	9,489 (45)	9,005 (45)	8,947 (39)	9,223 (44)	8,897 (43)	9,529 (47)	9,070 (45)
10	14,800 (160)	15,110 (173)	15,190 (191)	16,270 (159)	15,650 (173)	14,800 (70)	15,090 (73)	15,240 (72)	15,860 (58)	15,380 (58)	14,870 (75)	15,160 (75)	15,240 (77)	15,840 (60)	15,300 (59)
15	20,370 (194)	20,740 (203)	20,940 (176)	22,440 (204)	21,320 (198)	20,400 (105)	20,680 (106)	20,980 (84)	21,840 (86)	20,880 (82)	20,530 (105)	20,700 (107)	21,090 (87)	21,880 (87)	20,820 (81)
20	26,050 (227)	26,300 (230)	26,700 (197)	28,480 (269)	27,140 (235)	26,110 (131)	26,200 (130)	26,730 (98)	27,670 (108)	26,580 (108)	26,180 (128)	26,210 (132)	26,830 (97)	27,790 (97)	26,480 (108)
25	31,640 (276)	32,070 (194)	32,470 (189)	34,690 (311)	33,140 (338)	31,710 (161)	31,980 (110)	32,460 (104)	33,570 (162)	32,330 (112)	31,770 (166)	32,070 (111)	32,530 (104)	33,540 (159)	32,310 (113)
30	37,500 (339)	37,600 (226)	38,300 (224)	41,170 (229)	39,620 (338)	37,510 (201)	37,500 (138)	38,210 (126)	40,150 (195)	38,700 (182)	37,560 (197)	37,610 (142)	38,310 (128)	40,320 (204)	38,640 (176)
35	43,680 (259)	43,490 (256)	44,550 (369)	47,400 (288)	46,030 (303)	43,690 (156)	43,380 (158)	44,370 (211)	46,270 (161)	45,160 (176)	43,710 (151)	43,490 (159)	44,450 (214)	46,340 (167)	45,160 (175)
40	49,870 (279)	49,980 (303)	51,640 (276)	53,980 (441)	52,320 (258)	49,910 (175)	49,820 (181)	51,530 (198)	52,530 (159)	51,520 (154)	49,860 (178)	49,980 (185)	51,630 (151)	52,670 (159)	51,490 (153)
45	56,080 (278)	56,720 (454)	57,980 (309)	61,620 (307)	60,000 (373)	56,090 (169)	56,460 (299)	57,820 (198)	60,090 (284)	58,510 (273)	56,020 (165)	56,570 (302)	57,970 (197)	60,250 (276)	58,480 (282)
50	63,870 (490)	64,560 (340)	65,130 (433)	69,570 (551)	67,520 (475)	63,930 (318)	64,340 (223)	64,900 (276)	67,470 (268)	66,210 (238)	63,880 (322)	64,450 (224)	65,120 (275)	67,620 (270)	66,110 (241)
55	71,550 (382)	73,010 (587)	73,380 (338)	78,150 (435)	76,080 (409)	71,640 (245)	72,640 (400)	73,250 (233)	76,330 (271)	74,630 (309)	71,490 (243)	72,630 (404)	73,380 (229)	76,500 (274)	74,770 (315)
60	80,630 (521)	81,630 (434)	82,110 (493)	87,850 (507)	85,400 (542)	80,630 (345)	81,330 (280)	81,900 (337)	85,630 (415)	83,290 (399)	80,320 (339)	81,290 (282)	82,020 (332)	85,670 (423)	83,330 (409)
65	89,810 (621)	91,450 (518)	92,350 (592)	98,890 (613)	96,370 (595)	89,750 (410)	91,080 (360)	92,000 (426)	96,500 (370)	94,150 (418)	89,590 (391)	90,880 (356)	91,820 (417)	96,320 (402)	94,110 (435)
70	101,700 (614)	103,500 (668)	103,900 (381)	111,100 (774)	107,900 (782)	101,600 (420)	103,000 (474)	103,600 (282)	107,900 (398)	105,600 (416)	101,300 (429)	102,600 (427)	103,400 (278)	107,900 (395)	105,600 (444)
75	114,600 (570)	117,200 (583)	116,600 (739)	125,300 (898)	122,600 (792)	114,500 (386)	116,800 (411)	115,900 (413)	122,500 (366)	120,300 (441)	114,500 (396)	116,300 (453)	115,900 (410)	122,200 (362)	120,500 (434)
80	131,200 (548)	134,000 (756)	134,000 (758)	144,000 (878)	141,500 (796)	131,100 (368)	133,600 (513)	133,200 (601)	140,400 (729)	138,500 (645)	131,200 (367)	133,000 (518)	133,500 (591)	140,000 (695)	139,000 (656)
85	154,000 (997)	158,500 (936)	156,300 (605)	168,000 (988)	165,700 (1,039)	153,700 (583)	158,000 (796)	155,800 (406)	164,000 (584)	162,100 (601)	153,400 (442)	157,000 (901)	155,600 (421)	163,600 (576)	162,200 (589)
90	187,300 (1,269)	192,100 (1,303)	190,100 (1,424)	203,200 (1,043)	201,600 (1,006)	186,900 (898)	191,700 (817)	188,700 (857)	198,400 (866)	198,000 (931)	185,900 (662)	190,500 (775)	188,500 (827)	197,500 (859)	198,200 (968)
95	251,600 (2,405)	257,500 (2,377)	256,700 (2,250)	275,300 (2,250)	273,600 (2,754)	251,500 (1,443)	257,700 (1,645)	254,900 (1,059)	268,500 (1,174)	269,300 (2,005)	250,200 (1,111)	255,900 (1,214)	254,500 (816)	268,200 (1,324)	267,200 (1,911)

Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This table shows the estimates of income using various weights. In Columns (1)-(5), we show estimates of income at each percentile, consistent with the estimates in each year's Income and Poverty Report, except in 2017 and 2018 estimates, which use the 2017 Research File and 2018 Bridge File, respectively (Semega et al., 2019). In Columns (6)-(10), we show the estimates using the EBW with the full CPS sample. In Columns (11)-(15), we show the estimates using the March Basic sample (avoiding the challenge of adjusting the base weights for oversampling of Hispanics and households with children). Standard errors are shown in parenthesis.

Table 11: Percent Difference of Household Income using Survey and Alternative Weights

Percentile	Full CPS ASEC Sample (EBW)					Basic March CPS (March EBW)				
	- Survey					- Survey				
	2017	2018	2019	2020	2021	2017	2018	2019	2020	2021
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	
5	0.63 (1.41)	1.13 (1.61)	0.53 (1.73)	-3.16* (1.74)	-1.18 (1.89)	0.33 (1.40)	2.64 (1.64)	-0.33 (1.72)	-2.75 (1.76)	-0.46 (1.90)
10	0.01 (1.01)	-0.11 (1.05)	0.34 (1.20)	-2.50*** (0.89)	-1.70* (1.02)	0.47 (1.01)	0.30 (1.05)	0.33 (1.20)	-2.62*** (0.89)	-2.20** (1.01)
15	0.15 (0.91)	-0.30 (0.90)	0.21 (0.78)	-2.67*** (0.84)	-2.06** (0.86)	0.76 (0.91)	-0.21 (0.90)	0.72 (0.79)	-2.49*** (0.85)	-2.35*** (0.85)
20	0.26 (0.85)	-0.40 (0.81)	0.12 (0.69)	-2.85*** (0.86)	-2.05** (0.80)	0.53 (0.85)	-0.33 (0.80)	0.50 (0.69)	-2.43*** (0.87)	-2.40*** (0.79)
25	0.22 (0.86)	-0.26 (0.56)	-0.03 (0.55)	-3.22*** (0.82)	-2.46*** (0.92)	0.43 (0.86)	0.00 (0.56)	0.20 (0.55)	-3.30*** (0.82)	-2.52*** (0.92)
30	0.01 (0.87)	-0.27 (0.54)	-0.23 (0.54)	-2.46*** (0.55)	-2.33*** (0.76)	0.14 (0.87)	0.02 (0.55)	0.02 (0.54)	-2.05*** (0.56)	-2.48*** (0.76)
35	0.03 (0.55)	-0.24 (0.52)	-0.41 (0.75)	-2.39*** (0.52)	-1.88*** (0.60)	0.07 (0.55)	0.01 (0.53)	-0.23 (0.76)	-2.24*** (0.52)	-1.88*** (0.60)
40	0.08 (0.53)	-0.32 (0.52)	-0.22 (0.48)	-2.68*** (0.71)	-1.53*** (0.45)	-0.02 (0.53)	-0.01 (0.53)	-0.03 (0.47)	-2.43*** (0.71)	-1.58*** (0.45)
45	0.01 (0.45)	-0.46 (0.71)	-0.28 (0.47)	-2.48*** (0.48)	-2.47*** (0.58)	-0.10 (0.45)	-0.27 (0.72)	-0.01 (0.48)	-2.23*** (0.47)	-2.53*** (0.59)
50	0.08 (0.70)	-0.33 (0.46)	-0.35 (0.58)	-3.01*** (0.68)	-1.94*** (0.61)	0.02 (0.70)	-0.17 (0.46)	-0.01 (0.58)	-2.81*** (0.68)	-2.09*** (0.61)
55	0.12 (0.49)	-0.50 (0.69)	-0.18 (0.41)	-2.32*** (0.48)	-1.91*** (0.45)	-0.09 (0.49)	-0.52 (0.68)	0.00 (0.41)	-2.11*** (0.48)	-1.72*** (0.46)
60	0.00 (0.58)	-0.36 (0.43)	-0.25 (0.52)	-2.53*** (0.53)	-2.47*** (0.53)	-0.39 (0.58)	-0.41 (0.43)	-0.11 (0.52)	-2.48*** (0.53)	-2.43*** (0.53)
65	-0.07 (0.66)	-0.41 (0.47)	-0.38 (0.56)	-2.42*** (0.52)	-2.30*** (0.51)	-0.25 (0.66)	-0.63 (0.46)	-0.57 (0.55)	-2.60*** (0.52)	-2.35*** (0.52)
70	-0.06 (0.61)	-0.46 (0.52)	-0.30 (0.32)	-2.85*** (0.59)	-2.12*** (0.59)	-0.34 (0.61)	-0.85* (0.52)	-0.44 (0.31)	-2.87*** (0.58)	-2.09*** (0.59)
75	-0.10 (0.49)	-0.33 (0.42)	-0.60 (0.54)	-2.23*** (0.62)	-1.85*** (0.54)	-0.04 (0.50)	-0.79* (0.43)	-0.63 (0.54)	-2.46*** (0.61)	-1.68*** (0.54)
80	-0.03 (0.40)	-0.27 (0.49)	-0.62 (0.51)	-2.50*** (0.56)	-2.12*** (0.50)	-0.01 (0.40)	-0.71 (0.48)	-0.38 (0.51)	-2.79*** (0.55)	-1.77*** (0.51)
85	-0.24 (0.62)	-0.27 (0.55)	-0.33 (0.34)	-2.41*** (0.50)	-2.16*** (0.58)	-0.44 (0.60)	-0.93 (0.58)	-0.46 (0.34)	-2.64*** (0.50)	-2.11*** (0.58)
90	-0.21 (0.63)	-0.22 (0.60)	-0.74 (0.67)	-2.39*** (0.49)	-1.79*** (0.53)	-0.75 (0.60)	-0.82 (0.60)	-0.86 (0.67)	-2.82*** (0.49)	-1.67*** (0.55)
95	-0.04 (0.85)	0.05 (0.80)	-0.70 (0.80)	-2.47*** (0.74)	-1.57 (1.07)	-0.56 (0.84)	-0.64 (0.78)	-0.84 (0.79)	-2.58*** (0.75)	-2.32** (1.04)

Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This table shows the annual percent difference in median household income estimates using inverse probability weights compared to the survey weights. In Columns (1)-(5), we show estimates using the EBW with the full CPS ASEC sample in each year. In Columns (6)-(10), we show the estimates using the EBW with only the March Basic CPS sample (avoiding the challenge of adjusting the base weights for oversampling of Hispanics and households with children). Standard errors are shown in parenthesis. ***, **, and * indicate statistical significance at the 1-, 5-, and 10-percent levels respectively.

Table 12: Percent Year-to-Year Income Growth using Survey and Alternative Weights (in 2019 dollars)

Percentile	Survey Weights				Basic March CPS (March EBW)				March EBW - Survey			
	2018 (1)	2019 (2)	2020 (3)	2021 (4)	2018 (5)	2019 (6)	2020 (7)	2021 (8)	2018 (9)	2019 (10)	2020 (11)	2021 (12)
5	0.76 (1.93)	-0.67 (2.14)	9.78*** (2.69)	-7.00*** (2.31)	3.09*** (0.47)	-3.54*** (0.48)	7.10*** (0.56)	-4.81*** (0.49)	2.32 (1.93)	-2.87 (2.12)	-2.67 (2.72)	2.19 (2.26)
10	2.13 (1.41)	0.50 (1.56)	7.11*** (1.57)	-3.81*** (1.29)	1.95*** (0.51)	0.53** (0.48)	3.96*** (0.47)	-3.39*** (0.37)	-0.17 (1.36)	0.03 (1.50)	-3.15** (1.54)	0.41 (1.26)
15	1.82 (1.26)	0.94 (1.12)	7.17*** (1.22)	-5.00*** (1.11)	0.84*** (0.51)	1.88*** (0.45)	3.75*** (0.42)	-4.86*** (0.36)	-0.98 (1.21)	0.94 (1.08)	-3.42*** (1.19)	0.14 (1.09)
20	0.98 (1.22)	1.51 (0.99)	6.67*** (1.13)	-4.73*** (1.08)	0.12 (0.48)	2.36*** (0.42)	3.56*** (0.37)	-4.70*** (0.36)	-0.87 (1.16)	0.85 (0.96)	-3.11*** (1.12)	0.03 (1.07)
25	1.36 (1.00)	1.24* (0.74)	6.84*** (1.05)	-4.46*** (1.10)	0.94** (0.45)	1.44*** (0.30)	3.11*** (0.42)	-3.69*** (0.39)	-0.43 (0.97)	0.20 (0.72)	-3.73*** (1.02)	0.77 (1.09)
30	0.27 (1.00)	1.85** (0.73)	7.49*** (0.80)	-3.76*** (0.82)	0.15 (0.46)	1.85*** (0.32)	5.26*** (0.46)	-4.18*** (0.44)	-0.12 (0.96)	0.00 (0.70)	-2.23*** (0.80)	-0.42 (0.84)
35	-0.43 (0.76)	2.45*** (0.86)	6.40*** (0.97)	-2.90*** (0.67)	-0.49** (0.34)	2.21*** (0.40)	4.25*** (0.44)	-2.55*** (0.34)	-0.06 (0.72)	-0.24 (0.84)	-2.15** (0.94)	0.36 (0.66)
40	0.23 (0.72)	3.32*** (0.65)	4.53*** (0.89)	-3.08*** (0.75)	0.24 (0.35)	3.30*** (0.31)	2.02*** (0.29)	-2.24*** (0.27)	0.01 (0.69)	-0.02 (0.61)	-2.51*** (0.84)	0.84 (0.72)
45	1.15 (0.81)	2.21*** (0.80)	6.29*** (0.68)	-2.64*** (0.58)	0.98 (0.45)	2.48*** (0.43)	3.93*** (0.40)	-2.93*** (0.42)	-0.17 (0.77)	0.26 (0.78)	-2.36*** (0.65)	-0.30 (0.60)
50	1.07 (0.81)	0.88 (0.65)	6.82*** (0.98)	-2.94*** (0.78)	0.88 (0.43)	1.04** (0.34)	3.84*** (0.40)	-2.23*** (0.34)	-0.19 (0.76)	0.16 (0.63)	-2.98*** (0.91)	0.71 (0.75)
55	2.04** (0.86)	0.51 (0.73)	6.50*** (0.68)	-2.65*** (0.56)	1.59*** (0.48)	1.04* (0.44)	4.25*** (0.33)	-2.26*** (0.35)	-0.45 (0.78)	0.53 (0.70)	-2.25*** (0.64)	0.39 (0.53)
60	1.23* (0.73)	0.60 (0.61)	6.99*** (0.78)	-2.79*** (0.62)	1.21** (0.38)	0.90** (0.33)	4.45*** (0.44)	-2.74*** (0.43)	-0.02 (0.70)	0.30 (0.60)	-2.54*** (0.74)	0.05 (0.58)
65	1.83** (0.77)	0.98 (0.65)	7.08*** (0.84)	-2.55*** (0.63)	1.44*** (0.41)	1.03*** (0.37)	4.90*** (0.43)	-2.29*** (0.38)	-0.39 (0.75)	0.05 (0.64)	-2.18*** (0.77)	0.26 (0.58)
70	1.78** (0.74)	0.35 (0.58)	6.97*** (0.75)	-2.88*** (0.75)	1.26*** (0.41)	0.76 (0.33)	4.36*** (0.32)	-2.10*** (0.35)	-0.52 (0.73)	0.41 (0.54)	-2.61*** (0.68)	0.78 (0.69)
75	2.29*** (0.58)	-0.51 (0.62)	7.44*** (0.95)	-2.18*** (0.75)	1.53*** (0.36)	-0.35** (0.34)	5.47*** (0.34)	-1.40*** (0.30)	-0.76 (0.59)	0.16 (0.62)	-1.97** (0.88)	0.78 (0.69)
80	2.11*** (0.59)	0.06 (0.66)	7.45*** (0.84)	-1.75*** (0.65)	1.39*** (0.34)	0.39 (0.39)	4.85*** (0.47)	-0.71*** (0.45)	-0.72 (0.59)	0.33 (0.64)	-2.60*** (0.77)	1.04 (0.64)
85	2.87*** (0.74)	-1.36** (0.60)	7.49*** (0.68)	-1.40* (0.74)	2.36*** (0.49)	-0.89*** (0.47)	5.14*** (0.32)	-0.86*** (0.35)	-0.51 (0.80)	0.47 (0.61)	-2.36*** (0.63)	0.54 (0.73)
90	2.60*** (0.82)	-1.04 (0.82)	6.89*** (0.87)	-0.80 (0.63)	2.53*** (0.37)	-1.08*** (0.40)	4.78*** (0.45)	0.37 (0.47)	-0.07 (0.83)	-0.04 (0.79)	-2.11*** (0.81)	1.18* (0.71)
95	2.35** (1.16)	-0.32 (1.10)	7.26*** (1.18)	-0.63 (1.20)	2.26*** (0.44)	-0.52* (0.41)	5.37*** (0.47)	-0.37 (0.66)	-0.09 (1.11)	-0.20 (1.02)	-1.88* (1.12)	0.26 (1.30)

Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This table shows the estimates of the percent change in household income using various weights. In Columns (1)-(4), we show the year-to-year percent change in income at each percentile, consistent with the estimates in each year's Income and Poverty Report, except for the 2017 estimates, which use the 2017 Research File (Semega et al., 2020). In Columns (5)-(8), we show the change in income estimated using the EBW with the March Basic CPS sample (avoiding the challenge of adjusting the base weights for oversampling of Hispanics and households with children). Columns (9)-(12) show the difference in differences. Standard errors are shown in parenthesis. ***, **, and * indicate statistical significance at the 1-, 5-, and 10-percent levels respectively.

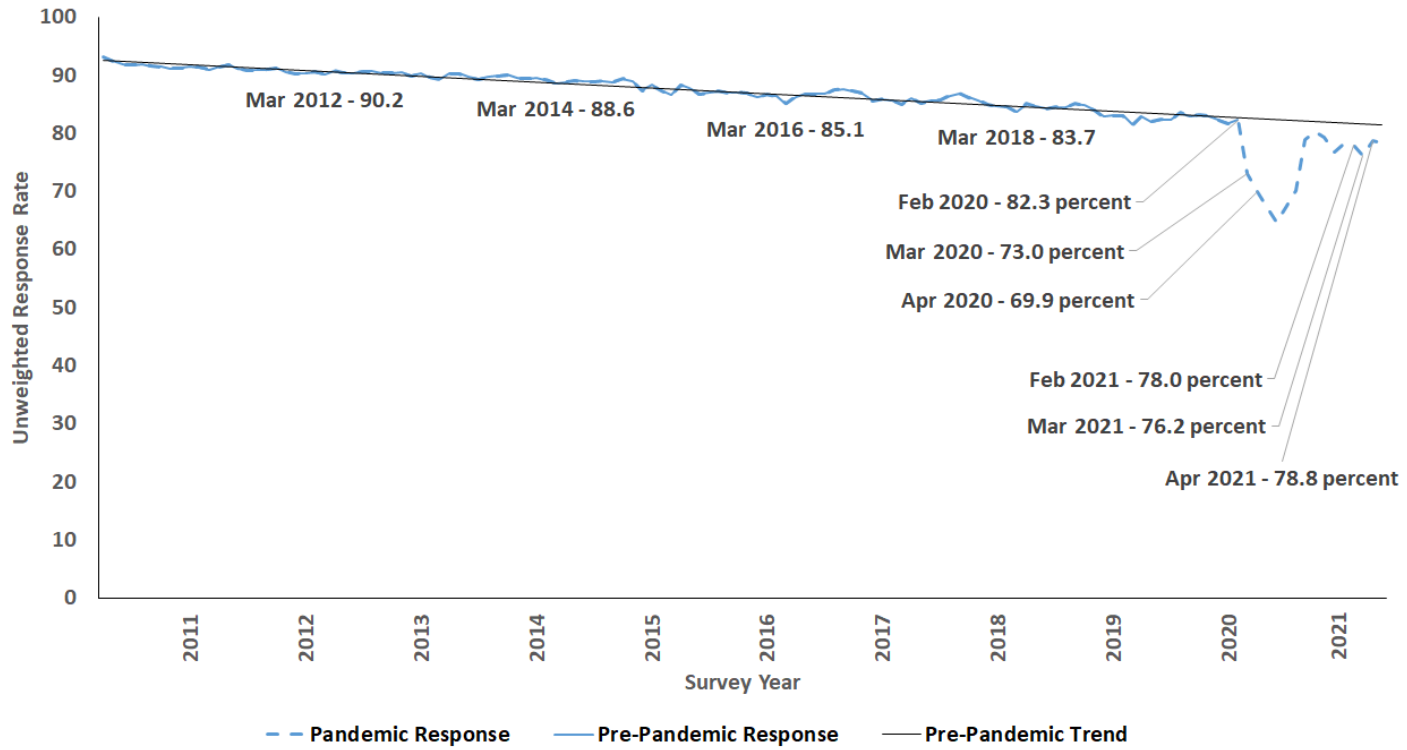
Table 13: Poverty Estimates using Survey and Alternative Weights

	Survey Weights					Full CPS (Full EBW)					Full EBW - Survey				
	2017 (1)	2018 (2)	2019 (3)	2020 (4)	2021 (5)	2017 (6)	2018 (7)	2019 (8)	2020 (9)	2021 (10)	2017 (11)	2018 (12)	2019 (13)	2020 (14)	2021 (15)
Official Poverty Measure															
Overall	12.8 (0.14)	12.3 (0.17)	11.8 (0.15)	10.5 (0.15)	11.4 (0.17)	12.9 (0.06)	12.4 (0.06)	11.9 (0.06)	10.9 (0.05)	11.9 (0.05)	0.15 (0.14)	0.09 (0.17)	0.14 (0.14)	0.44*** (0.15)	0.46*** (0.16)
White	11.1 (0.15)	10.5 (0.18)	10.1 (0.15)	9.1 (0.15)	10.1 (0.16)	11.2 (0.06)	10.6 (0.06)	10.2 (0.06)	9.4 (0.05)	10.5 (0.06)	0.06 (0.15)	0.09 (0.17)	0.09 (0.15)	0.37** (0.15)	0.37** (0.16)
Black	21.8 (0.56)	21.7 (0.51)	20.8 (0.59)	18.8 (0.55)	19.5 (0.58)	22.5 (0.15)	22.2 (0.14)	21.5 (0.22)	19.7 (0.14)	20.7 (0.16)	0.68 (0.55)	0.47 (0.52)	0.75 (0.57)	0.95* (0.54)	1.11* (0.60)
Hispanic	19.8 (0.44)	18.3 (0.47)	17.6 (0.41)	15.8 (0.44)	17.0 (0.47)	20.2 (0.12)	18.5 (0.12)	18.1 (0.11)	16.4 (0.11)	17.8 (0.11)	0.41 (0.44)	0.18 (0.49)	0.54 (0.39)	0.68 (0.43)	0.80* (0.45)
Supplemental Poverty Measure															
Overall	13.5 (0.16)	13.0 (0.19)	12.8 (0.16)	11.7 (0.17)	8.8 (0.14)	13.3 (0.05)	12.8 (0.03)	12.6 (0.03)	11.8 (0.03)	8.8 (0.02)	-0.21 (0.16)	-0.21 (0.18)	-0.15 (0.16)	0.06 (0.16)	-0.02 (0.14)
White	12.0 (0.16)	11.5 (0.20)	11.2 (0.16)	10.5 (0.16)	7.9 (0.14)	11.9 (0.05)	11.3 (0.03)	11.1 (0.03)	10.6 (0.03)	7.9 (0.03)	-0.15 (0.17)	-0.16 (0.19)	-0.11 (0.16)	0.09 (0.16)	-0.01 (0.14)
Black	21.0 (0.53)	20.6 (0.54)	20.4 (0.61)	18.4 (0.58)	14.1 (0.47)	20.6 (0.08)	20.1 (0.07)	20.0 (0.06)	18.2 (0.07)	14.2 (0.05)	-0.39 (0.53)	-0.52 (0.53)	-0.42 (0.60)	-0.12 (0.57)	0.07 (0.48)
Hispanic	22.5 (0.47)	20.5 (0.55)	20.3 (0.45)	18.9 (0.47)	13.5 (0.42)	22.1 (0.20)	20.1 (0.07)	20.2 (0.07)	18.7 (0.08)	13.2 (0.05)	-0.40 (0.49)	-0.38 (0.55)	-0.08 (0.44)	-0.20 (0.45)	-0.26 (0.41)

Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This table shows the estimates of poverty using various weights. In Columns (1)-(5), we show estimates using survey weights, consistent with the estimates in each year’s Income and Poverty Report (except for the 2017 estimates, which use the 2017 Research File). In Columns (6)-(10), we show the estimates using the EBW with only the March Basic CPS sample (avoiding the challenge of adjusting the base weights for oversampling of Hispanics and households with children). Columns (11)-(15) shows the difference between the EBW and survey estimates each year. Standard errors are shown in parenthesis. ***, **, and * indicate statistical significance at the 1-, 5-, and 10-percent levels respectively for the estimates in (11)-(15) only.

Figure 1: Basic CPS Monthly Unweighted Response Rates



Notes: This figure shows the unweighted household response rate to the Basic Monthly CPS over time.

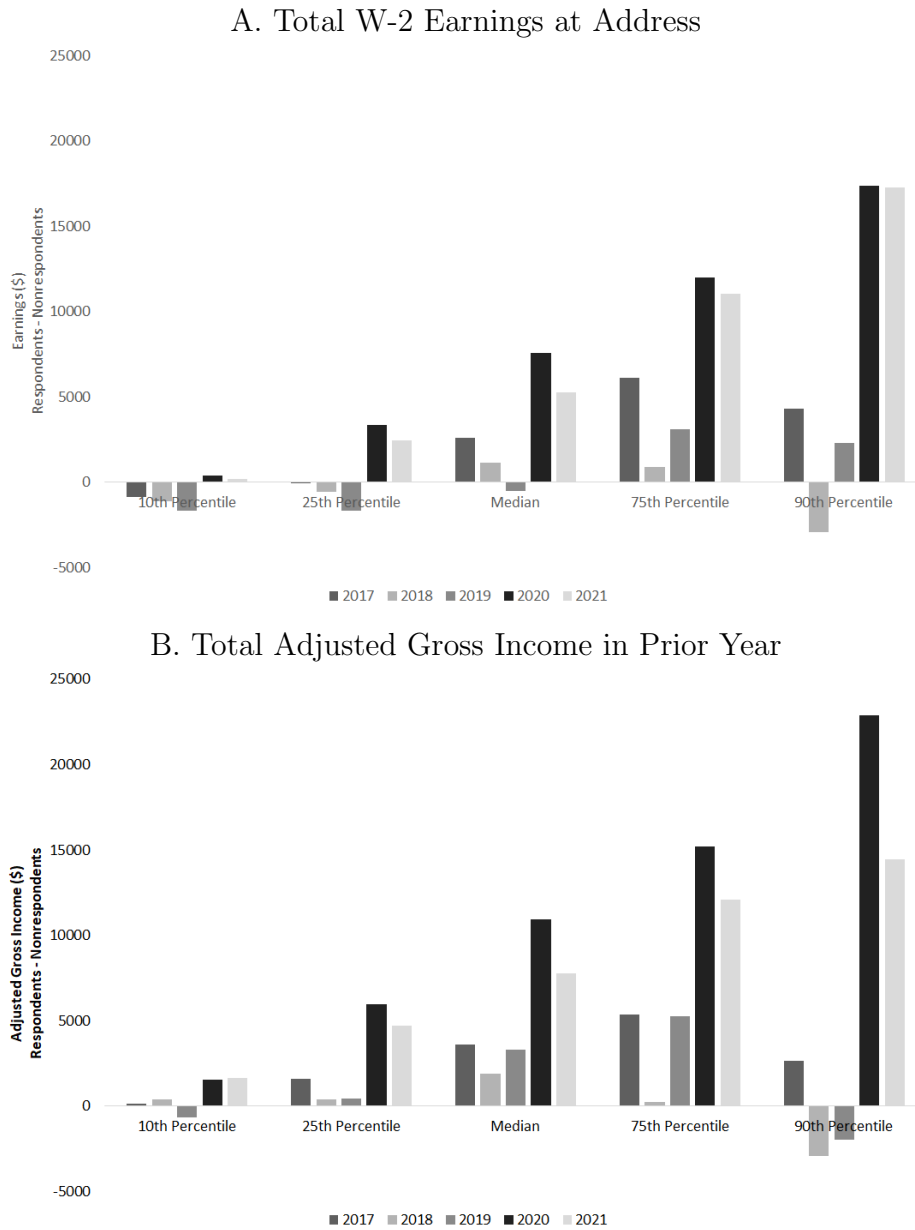
Source: Bureau of Labor Statistics, <https://www.bls.gov/osmr/response-rates/household-survey-response-rates.htm>.

Figure 2: Diagram of Data Linkage for Respondents and Nonrespondents

CPS Addresses (about 80,000 non-vacant housing units in each year)	Respondents (83 percent)	In 1099 IRMF (81 percent)	Link to W2 (78 percent) 1099R (41 percent) 1040 (90 percent) 2010 Census (93 percent) 2001-2018 ACS (27 percent)
			No link to additional data
		Not in 1099 IRMF No more links possible	
	Nonrespondents (17 percent)	In 1099 IRMF (77 percent)	Link to W2 (84), 1099R (35), 1040 (91), 2010 Census (90), 2001-2018 ACS (22)
		No link to additional data	
		Not in 1099 IRMF No more links possible	

Notes: This figure shows a diagram of the linkage process used in this paper and described in Table 1. The percent values shown in parenthesis are from the 2019 CPS ASEC. The values shown for the 1099 IRMF, W-2, 1099-R, 1040, 2010 Census and 2001-2018 ACS are linkage rates conditional on being in the group in the box to the left (i.e. for respondent housing units, 81 percent can be linked by address to the 1099 IRMF).

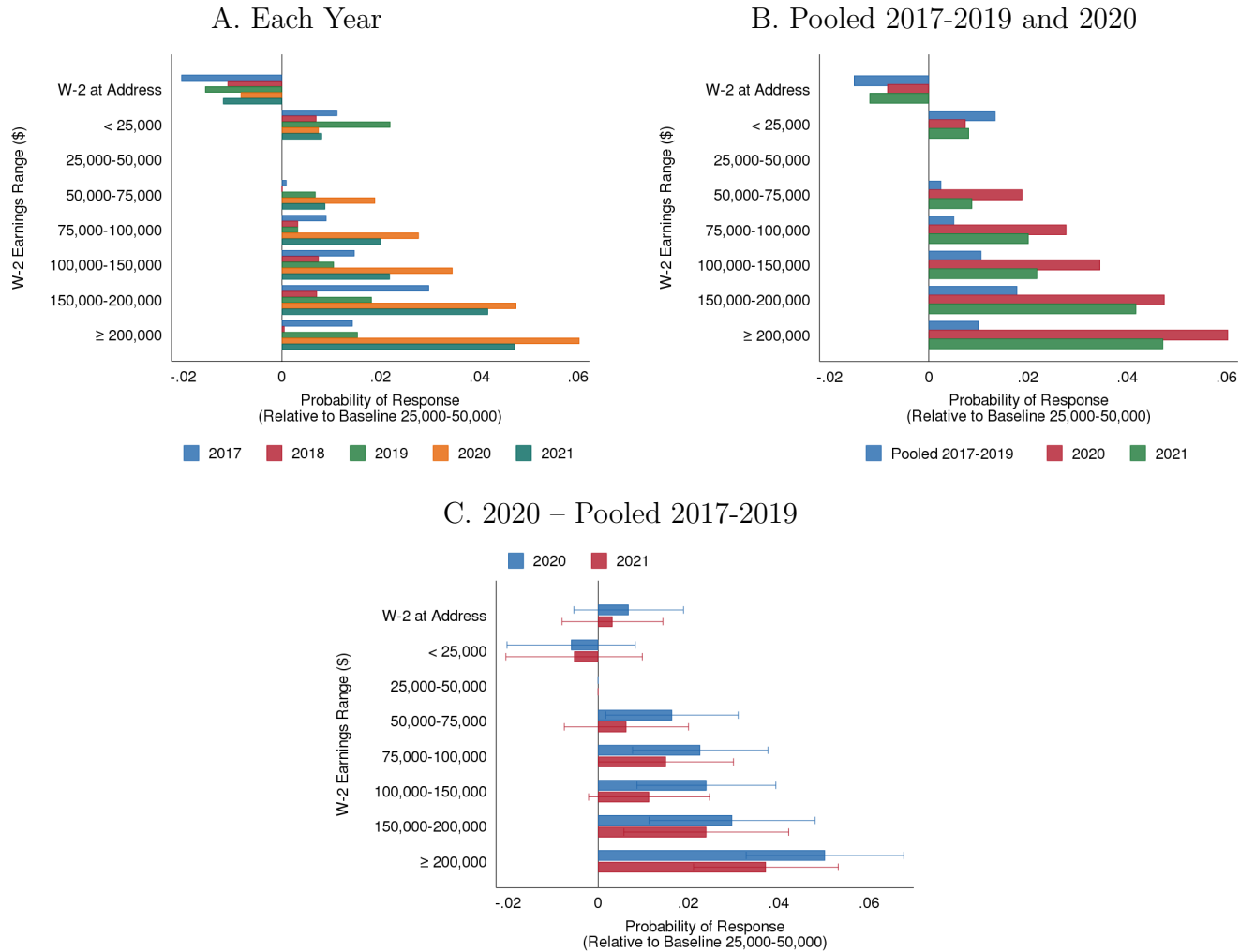
Figure 3: Income Difference between Respondents and Nonrespondents



Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This figure shows the difference in income by address between respondents and nonrespondents. Panel A shows total W-2 earnings at that address in the reference year of the survey. Panel B shows total 1040 AGI in the prior year for linked individuals at the survey address. A value of greater than zero indicates higher income for respondents than nonrespondents for that statistic and year.

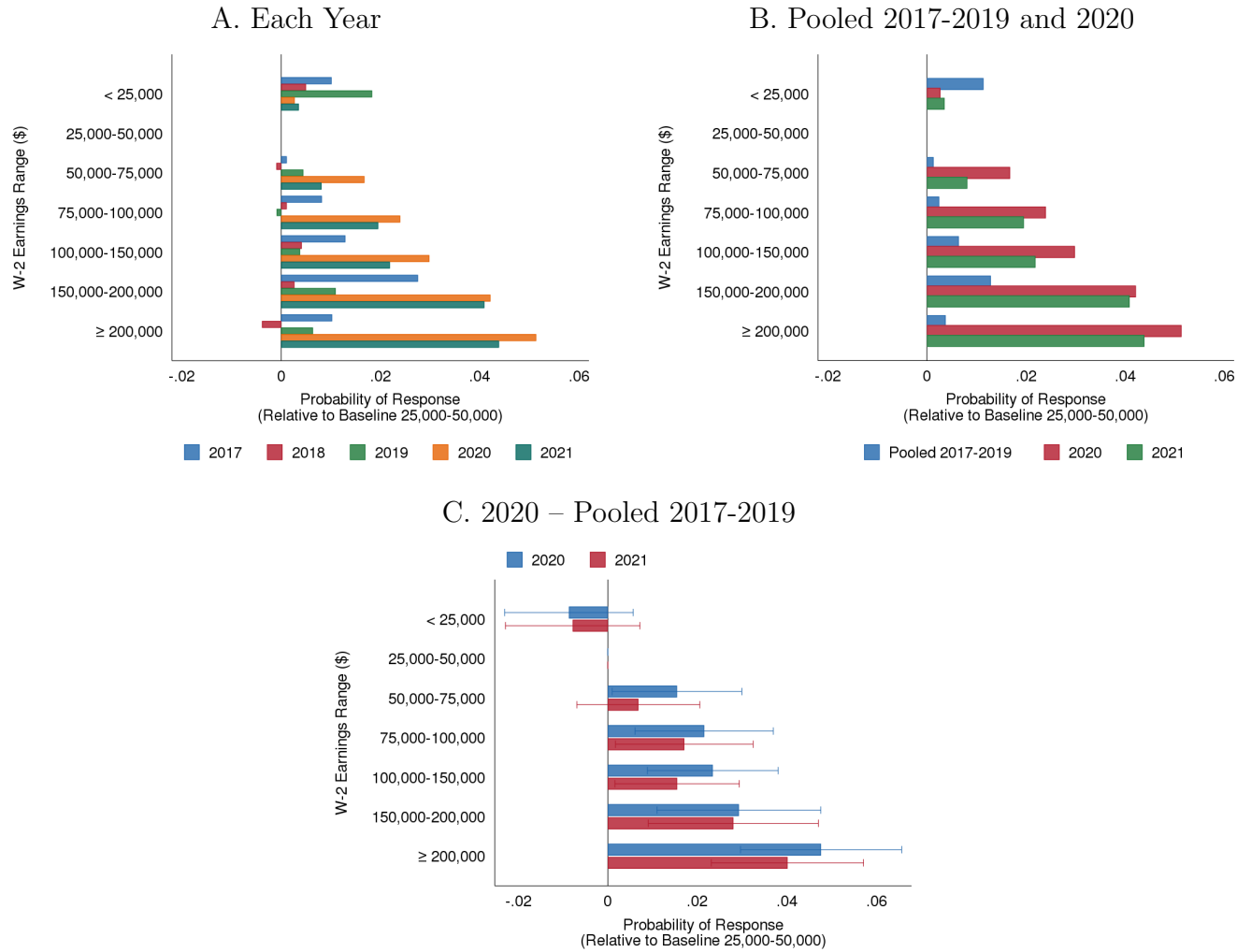
Figure 4: Probability of Response by Total W-2 Earnings at Address – No Controls



Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This figure shows the difference in income by address between respondents and nonrespondents. Panel A shows total W-2 earnings at that address in the reference year of the survey. Panel B shows total 1040 AGI in the prior year for linked individuals at the survey address. A value of greater than zero indicates higher income for respondents than nonrespondents for that statistic and year.

Figure 5: Probability of Response by Total W-2 Earnings at Address – Full Controls

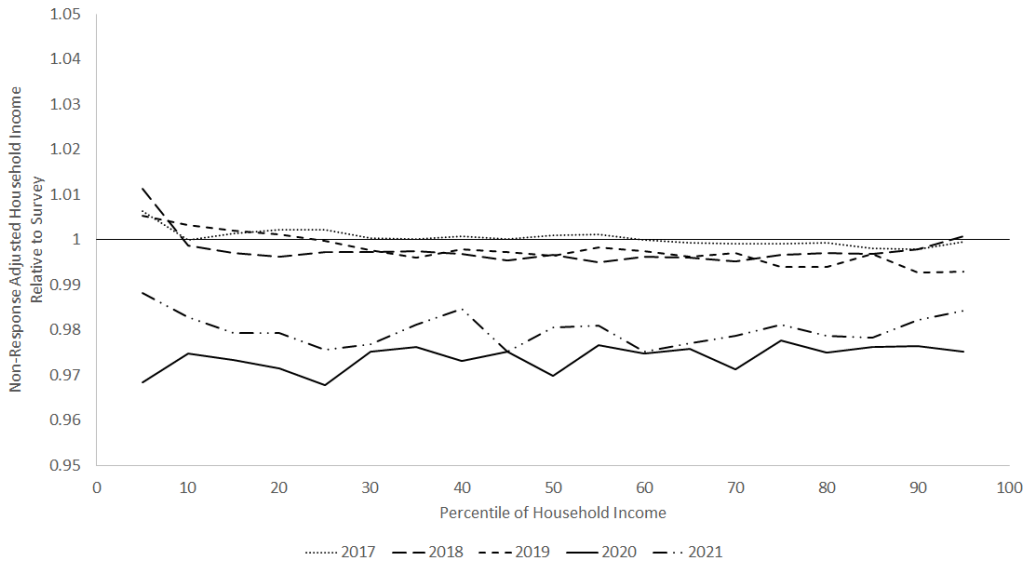


Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

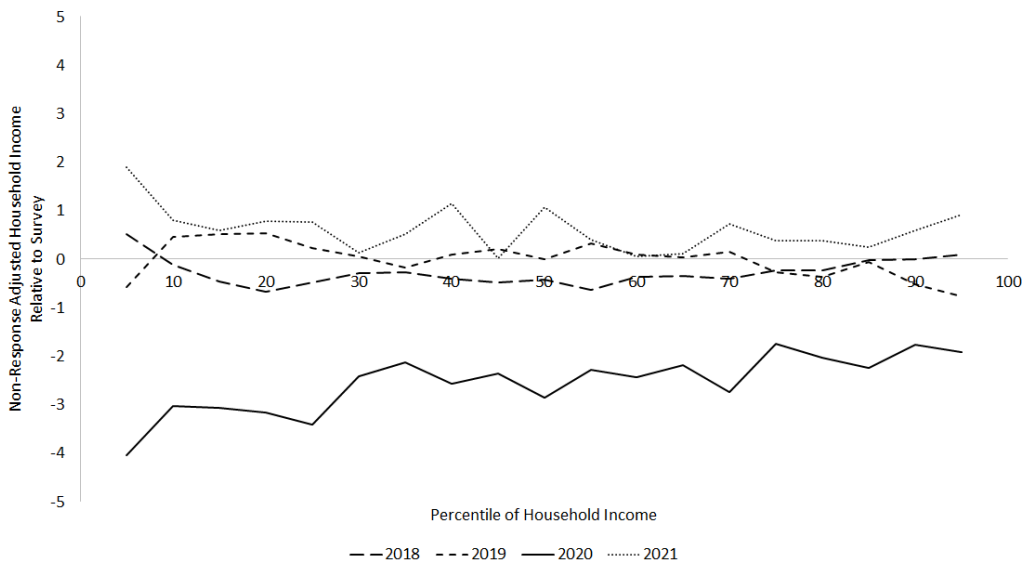
Notes: This figure shows the coefficient estimates from a regression of housing unit response on W-2 earnings at that address, with the addition of demographic and socioeconomic controls. Positive values indicate individuals in that income range are more likely to respond than the baseline group (25,000–50,000).

Figure 6: Comparing the Distribution of Household Income with Alternative Weights

A. Alternative Weight Estimate (Full EBW) Relative to Survey



B. Difference in Year-to-Year Growth with Alternative Weights (Full EBW) vs. Survey Weights



Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: Panel A shows the estimates of income using the Full EBW compared to the survey-weighted estimate as published in each year’s Income and Poverty Report (Semega et al., 2020). Panel B shows the difference in year-to-year growth in real household income with the Full EBW weights vs. the survey estimates.

Table A1: Probability of Response by Total Adjusted Gross Income in Prior Year

A. No Controls												
	Regression						Comparison					
	2017 (1)	2018 (2)	2019 (3)	Pooled (2017-2019) (4)	2020 (5)	2021 (6)	2018 - 2017 (7)	2019 - 2018 (8)	2020 - 2019 (9)	2021 - 2020 (10)	2020 - Pooled (11)	2021 - Pooled (12)
Filed 1040	0.007607* (0.004146)	0.01008** (0.004736)	0.01445*** (0.004197)	0.01134*** (0.002701)	0.04607*** (0.006013)	0.03750*** (0.006003)	0.002475 (0.006175)	0.004366 (0.005974)	0.03162*** (0.007280)	-0.008571 (0.008406)	0.03473*** (0.006657)	0.02616*** (0.007045)
0-25,000	0.006339 (0.004995)	0.007106 (0.005883)	0.01663*** (0.005055)	0.01021*** (0.003182)	0.001922 (0.006759)	-0.002755 (0.006837)	0.0007670 (0.007600)	0.009525 (0.007572)	-0.01471* (0.008530)	-0.004678 (0.009669)	-0.008290 (0.007462)	-0.01297* (0.007536)
50,000-75,000	0.007476 (0.005426)	0.006014 (0.005323)	-0.0009206 (0.005066)	0.003825 (0.003166)	0.02841*** (0.006950)	0.01768*** (0.006242)	-0.001463 (0.007746)	-0.006934 (0.006745)	0.02933*** (0.008540)	-0.01073 (0.009027)	0.02458*** (0.007570)	0.01385* (0.007188)
75,000-100,000	0.01219** (0.005331)	0.009287 (0.005822)	0.01217** (0.005412)	0.01113*** (0.003406)	0.01113*** (0.006706)	0.01570** (0.006816)	0.03744*** (0.00715)	0.01570** (0.007548)	-0.002906 (0.008196)	0.002885 (0.009418)	0.02526*** (0.009118)	0.004563 (0.008106)
100,000-150,000	0.01319*** (0.004861)	0.01573*** (0.005260)	0.01369*** (0.004887)	0.01418*** (0.003242)	0.04606*** (0.006289)	0.03012*** (0.006211)	0.002540 (0.007089)	-0.002048 (0.006569)	0.03237*** (0.008092)	-0.01594* (0.008349)	0.03188*** (0.007131)	0.01594** (0.007165)
150,000-200,000	0.01860*** (0.006095)	0.01107 (0.006774)	0.02704*** (0.006660)	0.01900*** (0.004259)	0.05720*** (0.008149)	0.03942*** (0.007490)	-0.007536 (0.008024)	0.01597* (0.008290)	0.03016*** (0.01037)	-0.01779* (0.01066)	0.03821*** (0.009215)	0.02042** (0.009286)
ge 200,000	0.01173** (0.005834)	0.003152 (0.005840)	0.01313** (0.005628)	0.009344*** (0.003484)	0.06307*** (0.006916)	0.03799*** (0.006588)	-0.008579 (0.007728)	0.009980 (0.007807)	0.04994*** (0.009073)	-0.02509*** (0.009224)	0.05373*** (0.008046)	0.02864*** (0.007650)
Constant	0.8566*** (0.002807)	0.8471*** (0.002856)	0.8227*** (0.002834)	0.8412*** (0.001780)	0.7112*** (0.004011)	0.7582*** (0.003678)	-0.009515** (0.003745)	-0.02445*** (0.003849)	-0.1115*** (0.004563)	0.04705*** (0.004777)	-0.1300*** (0.004374)	-0.08300*** (0.004148)
R-Squared	0.00	0.00	0.00	0.00	0.01	0.00						
Observations	81,000	79,500	82,000	242,000	79,500	79,500						

B. With Full Controls												
	Regression						Comparison					
	2017 (1)	2018 (2)	2019 (3)	Pooled (2017-2019) (4)	2020 (5)	2021 (6)	2018 - 2017 (7)	2019 - 2018 (8)	2020 - 2019 (9)	2021 - 2020 (10)	2020 - Pooled (11)	2021 - Pooled (12)
0-25,000	0.00009845 (0.004959)	0.002192 (0.005839)	0.01139** (0.005153)	0.004797 (0.003161)	-0.008443 (0.006801)	-0.009938 (0.006757)	0.002094 (0.007450)	0.009201 (0.007704)	-0.01984** (0.008609)	-0.001495 (0.009632)	-0.01324* (0.007562)	-0.01473** (0.007491)
50,000-75,000	0.008165 (0.005307)	0.004437 (0.005327)	-0.003075 (0.005129)	0.002712 (0.003149)	0.02496*** (0.007099)	0.01726*** (0.006201)	-0.003728 (0.007665)	-0.007512 (0.006808)	0.02804*** (0.008736)	-0.007698 (0.008975)	0.02225*** (0.007692)	0.01455** (0.007161)
75,000-100,000	0.01260** (0.005104)	0.006879 (0.005767)	0.007929 (0.005504)	0.008883*** (0.003346)	0.03250*** (0.006692)	0.01353* (0.006908)	-0.005719 (0.007523)	0.001050 (0.007683)	0.02457*** (0.008268)	-0.01897** (0.009561)	0.02362*** (0.007353)	0.004645 (0.008181)
100,000-150,000	0.01338*** (0.004791)	0.01255** (0.005436)	0.008604* (0.005140)	0.01127*** (0.003285)	0.03895*** (0.006349)	0.02731*** (0.006377)	-0.0008218 (0.007193)	-0.003950 (0.007008)	0.03034*** (0.008204)	-0.01164 (0.008306)	0.02767** (0.007117)	0.01603** (0.007367)
150,000-200,000	0.01838*** (0.006153)	0.008406 (0.006863)	0.02181*** (0.007040)	0.01607*** (0.004374)	0.04850*** (0.008213)	0.03669*** (0.007706)	-0.009978 (0.008235)	0.01340 (0.008664)	0.02669** (0.01068)	-0.01181 (0.01053)	0.03243*** (0.009322)	0.02062** (0.009502)
ge 200,000	0.01031* (0.005995)	-0.0006266 (0.006098)	0.005452 (0.006338)	0.004672 (0.003734)	0.05063*** (0.007266)	0.03302*** (0.006938)	-0.01093 (0.007981)	0.006078 (0.008661)	0.04518*** (0.009653)	-0.01761* (0.009415)	0.04596*** (0.008270)	0.02835*** (0.008171)
R-Squared	0.01	0.01	0.01	0.01	0.02	0.02						
Observations	81,000	79,500	82,000	242,000	79,500	79,500						

Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This table figure the coefficient estimates from a regression of housing unit response on total prior-year AGI for linked individuals at that address for the full CPS ASEC sample. Positive values indicate individuals in that income range are more likely to respond than the baseline group (25,000–50,000). Panel A shows the results without controls for linkage rates and available demographic and socioeconomic information (such as race, Hispanic origin, citizenship, etc.). Panel B shows the results with those controls included. ***, **, and * indicate statistical significance at the 1-, 5-, and 10-percent levels respectively.

Table A2: Probability of Response by Total W-2 Earnings at Address
MIS 1 and 5

A. No Controls									
	Regression					Comparison			
	2017 (1)	2018 (2)	2019 (3)	Pooled (2017-2019) (4)	2020 (5)	2018 - 2017 (6)	2019 - 2018 (7)	2020 - 2019 (8)	2020 - Pooled (9)
Filed 1040	0.007607* (0.004146)	0.01008** (0.004736)	0.01445*** (0.004197)	0.01134*** (0.002701)	0.04607*** (0.006013)	0.002475 (0.006175)	0.004366 (0.005974)	0.03162*** (0.007280)	0.03473*** (0.006657)
0-25,000	0.006339 (0.004995)	0.007106 (0.005883)	0.01663*** (0.005055)	0.01021*** (0.003182)	0.001922 (0.006759)	0.0007670 (0.007600)	0.009525 (0.007572)	-0.01471* (0.008530)	-0.008290 (0.007462)
50,000-75,000	0.007476 (0.005426)	0.006014 (0.005323)	-0.0009206 (0.005066)	0.003825 (0.003166)	0.02841*** (0.006950)	-0.001463 (0.007746)	-0.006934 (0.006745)	0.02933*** (0.008540)	0.02458*** (0.007570)
75,000-100,000	0.01219** (0.005331)	0.009287 (0.005822)	0.01217** (0.005412)	0.01113*** (0.003406)	0.03744*** (0.006706)	-0.002906 (0.007715)	0.002885 (0.007548)	0.02526*** (0.008196)	0.02630*** (0.007416)
100,000-150,000	0.01319*** (0.004861)	0.01573*** (0.005260)	0.01369*** (0.004887)	0.01418*** (0.003242)	0.04606*** (0.006289)	0.002540 (0.007089)	-0.002048 (0.006569)	0.03237*** (0.008092)	0.03188*** (0.007131)
150,000-200,000	0.01860*** (0.006095)	0.01107 (0.006774)	0.02704*** (0.006660)	0.01900*** (0.004259)	0.05720*** (0.008149)	-0.007536 (0.008024)	0.01597* (0.008290)	0.03016*** (0.01037)	0.03821*** (0.009215)
≥ 200,000	0.01173** (0.005834)	0.003152 (0.005840)	0.01313** (0.005628)	0.009344*** (0.003484)	0.06307*** (0.006916)	-0.008579 (0.007728)	0.009980 (0.007807)	0.04994*** (0.009073)	0.05373*** (0.008046)
Constant	0.8566*** (0.002807)	0.8471*** (0.002856)	0.8227*** (0.002834)	0.8412*** (0.001780)	0.7112*** (0.004011)	-0.009515** (0.003745)	-0.02445*** (0.003849)	-0.1115*** (0.004563)	-0.1300*** (0.004374)
R-Squared	0.00	0.00	0.00	0.00	0.01				
Observations	81,000	79,500	82,000	242,000	79,500				

B. With Full Controls									
	Regression					Comparison			
	2017 (1)	2018 (2)	2019 (3)	Pooled (2017-2019) (4)	2020 (5)	2018 - 2017 (6)	2019 - 2018 (7)	2020 - 2019 (8)	2020 - Pooled (9)
0-25,000	0.00009845 (0.004959)	0.002192 (0.005839)	0.01139** (0.005153)	0.004797 (0.003161)	-0.008443 (0.006801)	0.002094 (0.007450)	0.009201 (0.007704)	-0.01984** (0.008609)	-0.01324* (0.007562)
50,000-75,000	0.008165 (0.005307)	0.004437 (0.005327)	-0.003075 (0.005129)	0.002712 (0.003149)	0.02496*** (0.007099)	-0.003728 (0.007665)	-0.007512 (0.006808)	0.02804*** (0.008736)	0.02225*** (0.007692)
75,000-100,000	0.01260** (0.005104)	0.006879 (0.005767)	0.007929 (0.005504)	0.008883*** (0.003346)	0.03250*** (0.006692)	-0.005719 (0.007523)	0.001050 (0.007683)	0.02457*** (0.008268)	0.02362*** (0.007353)
100,000-150,000	0.01338*** (0.004791)	0.01255** (0.005436)	0.008604* (0.005140)	0.01127*** (0.003285)	0.03895*** (0.006349)	-0.0008218 (0.007193)	-0.003950 (0.007008)	0.03034*** (0.008204)	0.02767*** (0.007117)
150,000-200,000	0.01838*** (0.006153)	0.008406 (0.006863)	0.02181*** (0.007040)	0.01607*** (0.004374)	0.04850*** (0.008213)	-0.009978 (0.008235)	0.01340 (0.008664)	0.02669** (0.01068)	0.03243*** (0.009322)
≥ 200,000	0.01031* (0.005995)	-0.0006266 (0.006098)	0.005452 (0.006338)	0.004672 (0.003734)	0.05063*** (0.007266)	-0.01093 (0.007981)	0.006078 (0.008661)	0.04518*** (0.009653)	0.04596*** (0.008270)
R-Squared	0.01	0.01	0.01	0.01	0.02				
Observations	81,000	79,500	82,000	242,000	79,500				

Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This table shows the coefficient estimates from a regression of housing unit response on W-2 earnings at that address for respondents in Month-in-Sample 1 and 5. Month-in-Sample 1 and 5 response rates were particularly affected by the pandemic as those interviews are more likely to be conducted in person in non-pandemic years. Positive values indicate individuals in that income range are more likely to respond than the baseline group (25,000–50,000). Panel A shows the results without controls for linkage rates and available demographic and socioeconomic information (such as race, Hispanic origin, citizenship, etc.). Panel B shows the results with those controls included. ***, **, and * indicate statistical significance at the 1-, 5-, and 10-percent levels respectively.

Table A3: Probability of Response by Total W-2 Earnings at Address
Not MIS 1 and 5

A. No Controls									
	Regression					Comparison			
	2017 (1)	2018 (2)	2019 (3)	Pooled (2017-2019) (4)	2020 (5)	2018 - 2017 (6)	2019 - 2018 (7)	2020 - 2019 (8)	2020 - Pooled (9)
Filed 1040	0.007607* (0.004146)	0.01008** (0.004736)	0.01445*** (0.004197)	0.01134*** (0.002701)	0.04607*** (0.006013)	0.002475 (0.006175)	0.004366 (0.005974)	0.03162*** (0.007280)	0.03473*** (0.006657)
0-25,000	0.006339 (0.004995)	0.007106 (0.005883)	0.01663*** (0.005055)	0.01021*** (0.003182)	0.001922 (0.006759)	0.0007670 (0.007600)	0.009525 (0.007572)	-0.01471* (0.008530)	-0.008290 (0.007462)
50,000-75,000	0.007476 (0.005426)	0.006014 (0.005323)	-0.0009206 (0.005066)	0.003825 (0.003166)	0.02841*** (0.006950)	-0.001463 (0.007746)	-0.006934 (0.006745)	0.02933*** (0.008540)	0.02458*** (0.007570)
75,000-100,000	0.01219** (0.005331)	0.009287 (0.005822)	0.01217** (0.005412)	0.01113*** (0.003406)	0.03744*** (0.006706)	-0.002906 (0.007715)	0.002885 (0.007548)	0.02526*** (0.008196)	0.02630*** (0.007416)
100,000-150,000	0.01319*** (0.004861)	0.01573*** (0.005260)	0.01369*** (0.004887)	0.01418*** (0.003242)	0.04606*** (0.006289)	0.002540 (0.007089)	-0.002048 (0.006569)	0.03237*** (0.008092)	0.03188*** (0.007131)
150,000-200,000	0.01860*** (0.006095)	0.01107 (0.006774)	0.02704*** (0.006660)	0.01900*** (0.004259)	0.05720*** (0.008149)	-0.007536 (0.008024)	0.01597* (0.008290)	0.03016*** (0.01037)	0.03821*** (0.009215)
≥ 200,000	0.01173** (0.005834)	0.003152 (0.005840)	0.01313** (0.005628)	0.009344*** (0.003484)	0.06307*** (0.006916)	-0.008579 (0.007728)	0.009980 (0.007807)	0.04994*** (0.009073)	0.05373*** (0.008046)
Constant	0.8566*** (0.002807)	0.8471*** (0.002856)	0.8227*** (0.002834)	0.8412*** (0.001780)	0.7112*** (0.004011)	-0.009515** (0.003745)	-0.02445*** (0.003849)	-0.1115*** (0.004563)	-0.1300*** (0.004374)
R-Squared	0.00	0.00	0.00	0.00	0.01				
Observations	81,000	79,500	82,000	242,000	79,500				

B. With Full Controls									
	Regression					Comparison			
	2017 (1)	2018 (2)	2019 (3)	Pooled (2017-2019) (4)	2020 (5)	2018 - 2017 (6)	2019 - 2018 (7)	2020 - 2019 (8)	2020 - Pooled (9)
0-25,000	0.00009845 (0.004959)	0.002192 (0.005839)	0.01139** (0.005153)	0.004797 (0.003161)	-0.008443 (0.006801)	0.002094 (0.007450)	0.009201 (0.007704)	-0.01984** (0.008609)	-0.01324* (0.007562)
50,000-75,000	0.008165 (0.005307)	0.004437 (0.005327)	-0.003075 (0.005129)	0.002712 (0.003149)	0.02496*** (0.007099)	-0.003728 (0.007665)	-0.007512 (0.006808)	0.02804*** (0.008736)	0.02225*** (0.007692)
75,000-100,000	0.01260** (0.005104)	0.006879 (0.005767)	0.007929 (0.005504)	0.008883*** (0.003346)	0.03250*** (0.006692)	-0.005719 (0.007523)	0.001050 (0.007683)	0.02457*** (0.008268)	0.02362*** (0.007353)
100,000-150,000	0.01338*** (0.004791)	0.01255** (0.005436)	0.008604* (0.005140)	0.01127*** (0.003285)	0.03895*** (0.006349)	-0.0008218 (0.007193)	-0.003950 (0.007008)	0.03034*** (0.008204)	0.02767*** (0.007117)
150,000-200,000	0.01838*** (0.006153)	0.008406 (0.006863)	0.02181*** (0.007040)	0.01607*** (0.004374)	0.04850*** (0.008213)	-0.009978 (0.008235)	0.01340 (0.008664)	0.02669** (0.01068)	0.03243*** (0.009322)
≥ 200,000	0.01031* (0.005995)	-0.0006266 (0.006098)	0.005452 (0.006338)	0.004672 (0.003734)	0.05063*** (0.007266)	-0.01093 (0.007981)	0.006078 (0.008661)	0.04518*** (0.009653)	0.04596*** (0.008270)
R-Squared	0.01	0.01	0.01	0.01	0.02				
Observations	81,000	79,500	82,000	242,000	79,500				

Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This table shows the coefficient estimates from a regression of housing unit response on W-2 earnings at that address for respondents not in Month-in-Sample 1 and 5. Month-in-Sample 1 and 5 response rates were particularly affected by the pandemic as those interviews are more likely to be conducted in person in non-pandemic years. Positive values indicate individuals in that income range are more likely to respond than the baseline group (25,000–50,000). Panel A shows the results without controls for linkage rates and available demographic and socioeconomic information (such as race, Hispanic origin, citizenship, etc.). Panel B shows the results with those controls included. ***, **, and * indicate statistical significance at the 1-, 5-, and 10-percent levels respectively.

Table A4: Validation of Public-Use Weights for Median Household Income

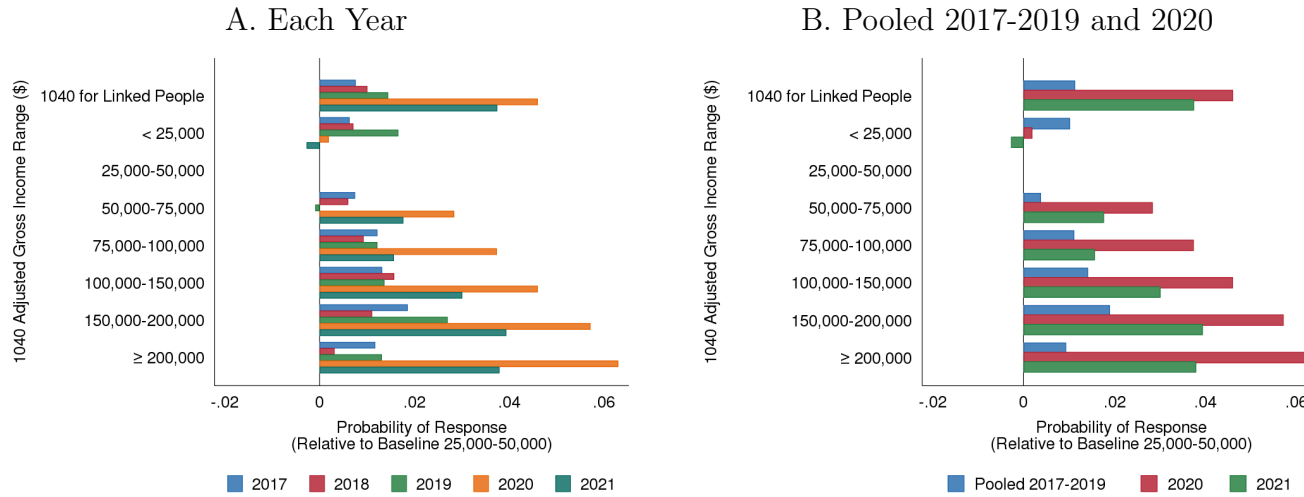
	A. Estimates of Median Household Income (Current Dollars)											
	2017			2018			2019			2020		
	Survey (1)	Full EBW (2)	Public-Use EBW (3)	Survey (4)	Full EBW (5)	Public-Use EBW (6)	Survey (7)	Full EBW (8)	Public-Use EBW (9)	Survey (10)	Full EBW (11)	Public-Use EBW (12)
All Households	59,210	59,270	59,250	61,140	60,920	60,920	63,180	63,060	63,100	68,700	66,790	66,810
Married-Couple Households	87,360	86,910	86,980	91,330	90,740	90,790	93,650	92,870	92,900	102,300	100,300	100,300
White Households	61,950	62,010	61,990	64,830	64,640	64,650	66,940	67,010	67,040	72,200	70,680	70,690
White, Non-Hispanic Households	65,440	65,490	65,480	68,190	67,870	67,900	70,640	70,770	70,810	76,060	74,370	74,410
Black Households	39,750	39,840	39,950	39,360	39,070	39,150	41,360	41,200	41,240	45,440	43,750	43,720
Hispanic Households	46,930	47,240	47,250	50,170	50,640	50,660	51,450	51,510	51,520	56,110	55,620	55,700
Householder < 65 Years Old	66,180	66,410	66,400	69,260	69,220	69,220	71,660	71,630	71,620	77,870	76,070	76,080
Householder ≥ 65 Years Old	40,530	40,060	40,070	41,300	40,700	40,690	43,700	43,320	43,350	47,360	46,200	46,210

	B. Percent Difference from Survey							
	2017		2018		2019		2020	
	Full EBW (1)	Public-Use EBW (2)	Full EBW (3)	Public-Use EBW (4)	Full EBW (5)	Public-Use EBW (6)	Full EBW (7)	Public-Use EBW (8)
All Households	0.1	0.1	-0.4	-0.4	-0.2	-0.1	-2.8	-2.8
Married-Couple Households	-0.5	-0.4	-0.6	-0.6	-0.8	-0.8	-2.0	-2.0
White Households	0.1	0.1	-0.3	-0.3	0.1	0.1	-2.1	-2.1
White, Non-Hispanic Households	0.1	0.1	-0.5	-0.4	0.2	0.2	-2.2	-2.2
Black Households	0.2	0.5	-0.7	-0.5	-0.4	-0.3	-3.7	-3.8
Hispanic Households	0.7	0.7	0.9	1.0	0.1	0.1	-0.9	-0.7
Householder < 65 Years Old	0.3	0.3	-0.1	-0.1	Z	-0.1	-2.3	-2.3
Householder ≥ 65 Years Old	-1.2	-1.1	-1.5	-1.5	-0.9	-0.8	-2.4	-2.4

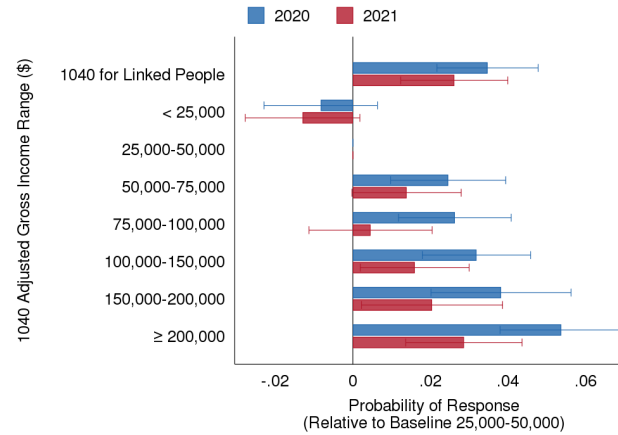
Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This table shows median household income estimates for various subgroups of households, in current dollars. Medians cannot be used as moment conditions in entropy balancing, so these serve as a simple test for the success of the public-use weights in matching untargeted, but relevant, income statistics. Survey and Full EBW estimates are generated as in all other tables, using the internal CPS ASEC file with the survey and EBW weights, respectively. Under Public-Use EBW, income is estimated using the public-use CPS ASEC file and the public-use weights discussed in Section 6. The percent differences have not been tested for statistical significance and are shown for reference only. Z indicates an estimate rounds to 0.

Figure A1: Probability of Response by Total Adjusted Gross Income in Prior Year – No Controls



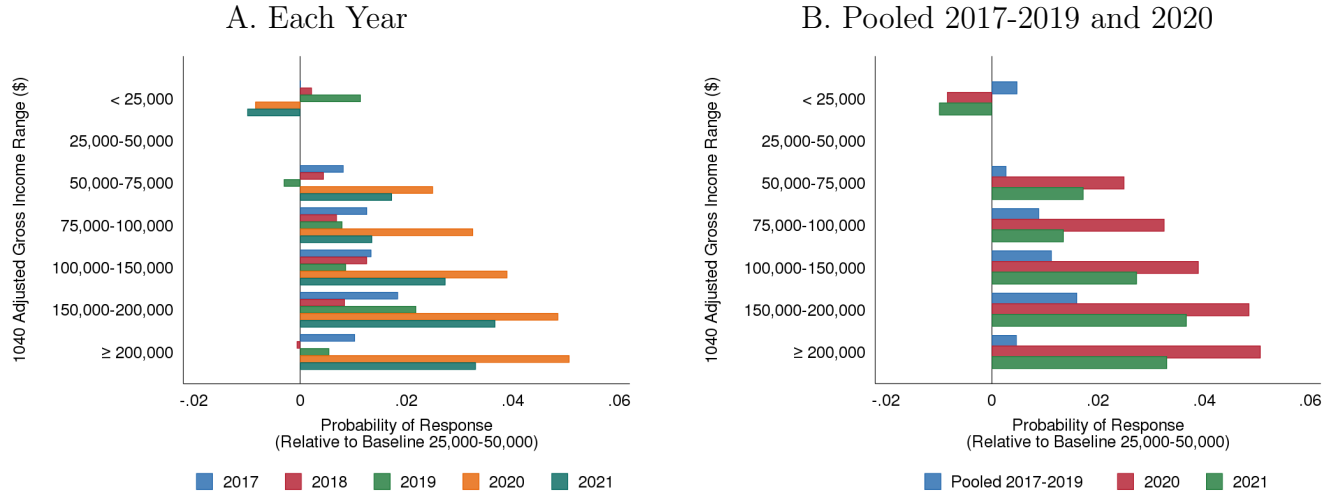
C. 2020 – Pooled 2017-2019



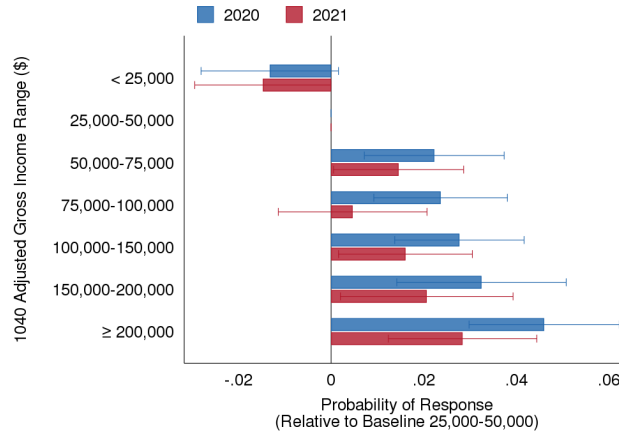
Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This figure shows the coefficient estimates from a regression of housing unit response on total prior-year AGI for linked individuals at that address. Positive values indicate individuals in that income range are more likely to respond than the baseline group (25,000–50,000).

Figure A2: Probability of Response by Total W-2 Earnings at Address – Full Controls



C. 2020 – Pooled 2017-2019



Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This figures shows the coefficient estimates from a regression of housing unit response on total prior-year AGI for linked individuals at that address, with the addition of demographic and socioeconomic controls. Positive values indicate individuals in that income range are more likely to respond than the baseline group (25,000–50,000).