

Addressing Nonresponse Bias in Household Surveys Using Linked Administrative Data

Jonathan Rothbaum

US Census Bureau

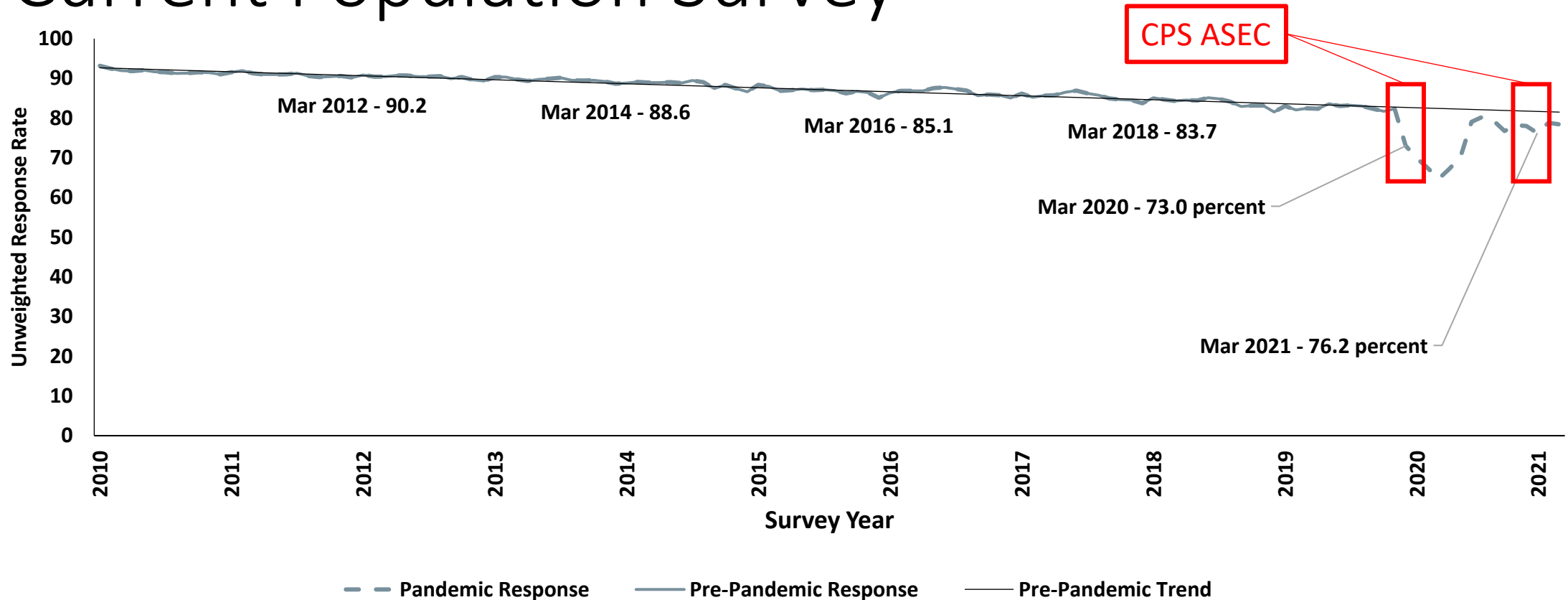
Society of Government Economists, ASSA 2022 Annual Meeting
January 7, 2022

This presentation 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, CBDRB-FY21-POP001-0199, and CBDRB-FY21-POP001-0238. The public-use weights are approved for release under approval CBDRB-FY21-126.

Results from Two Papers

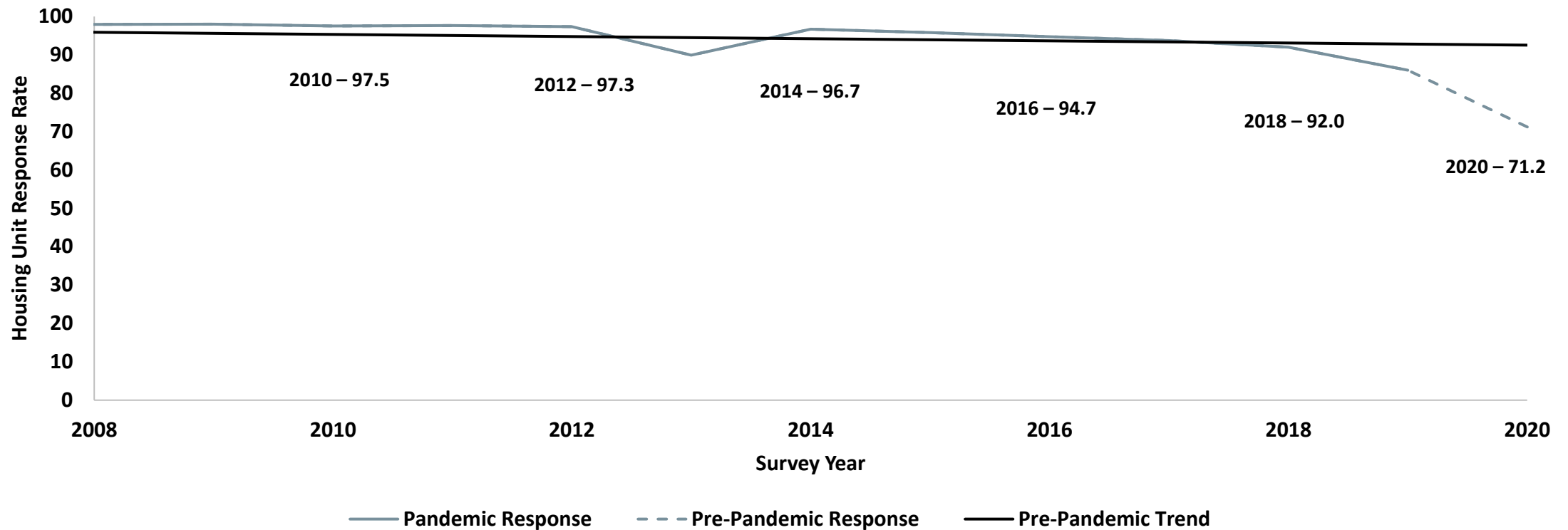
1. CPS ASEC - [Addressing Nonresponse Bias in Household Surveys using Linked Administrative Data](#) with Adam Bee (earlier version)
2. ACS - [Addressing Nonresponse Bias in the American Community Survey During the Pandemic Using Administrative Data](#) with Jonathan Eggleston, Adam Bee, Mark Klee, and Brian Mendez-Smith

Response Rate Current Population Survey



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

Response Rate American Community Survey



Source: U.S. Census Bureau – <https://www.census.gov/acs/www/methodology/sample-size-and-data-quality/response-rates/>

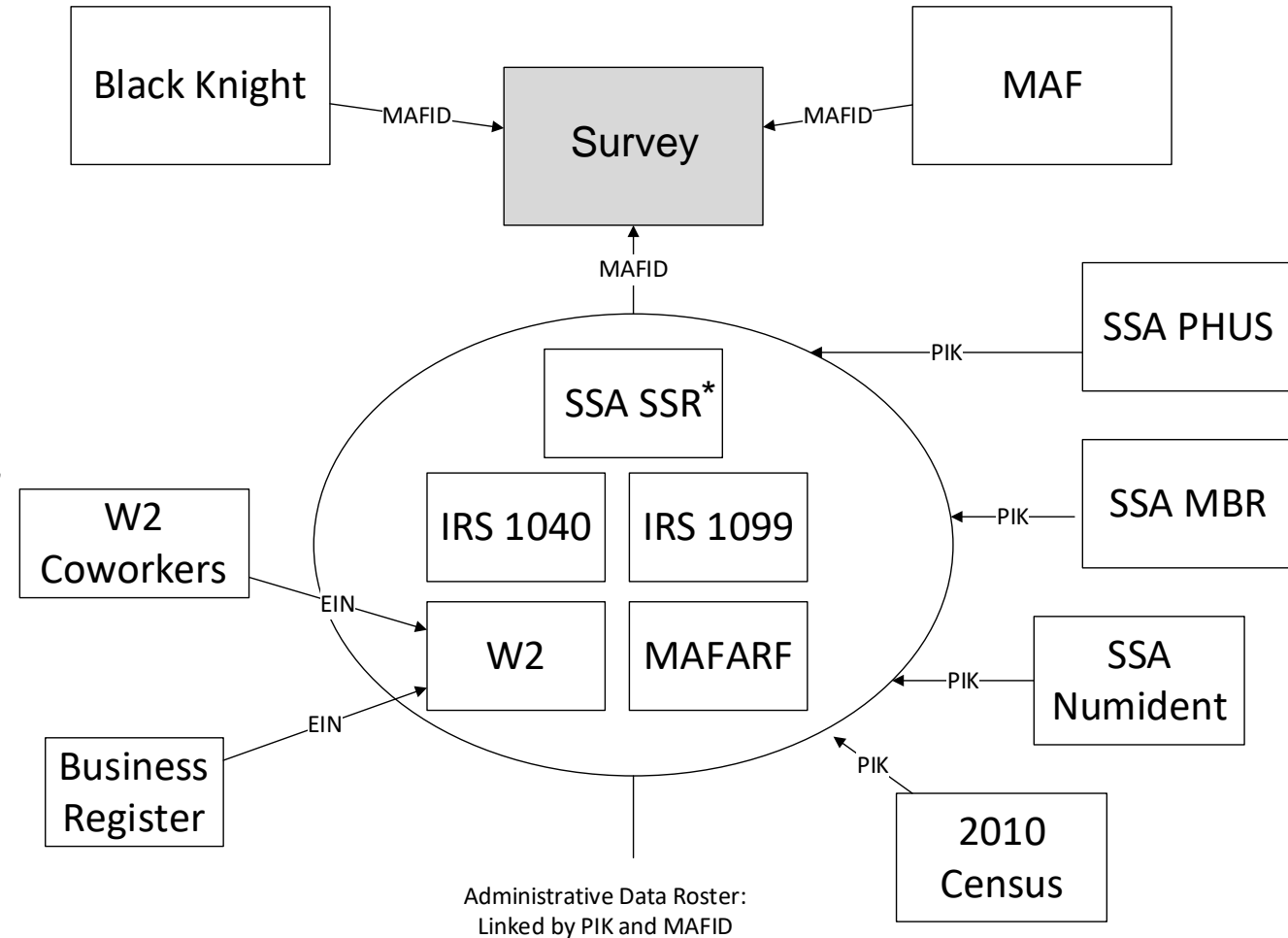
What We Do

- Use linked administrative, census, and prior survey data to evaluate nonresponse bias
- Reweight the sample to correct for nonresponse bias
 - Condition on linked data and external population controls
- Create public-use weights that do not condition on administrative data (CPS ASEC)

Additional Data

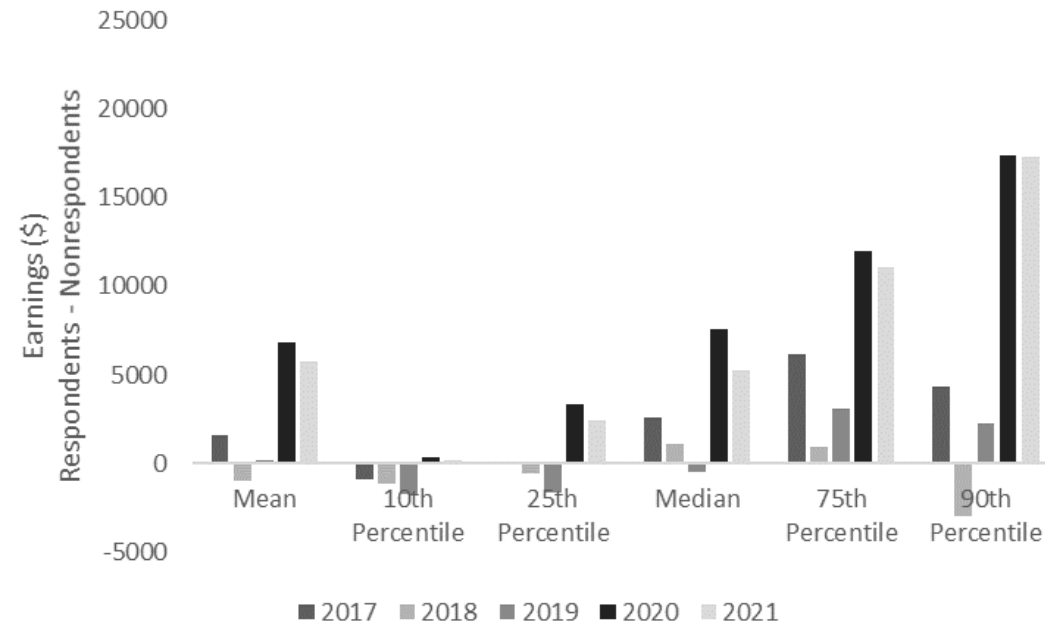
- Data Linkage
 - Link administrative data to addresses of both **respondents** *and* **nonrespondents**
- What We Have
 - Housing – MAF, Black Knight*
 - Income/Employment – W2, 1040, 1099s, PHUS*, SSR*, MBR*, Business Register*
 - Demographics – Numident, 2010 Census. Includes race, Hispanic-origin, citizenship, age, gender
 - Household composition/marital status – 1040
- What We Don't Have
 - Direct measures of education, health insurance status, among other stats
 - Any information in linked data for households that can't be linked by MAFID

Note: * indicates used with ACS only



Nonrandom Nonresponse in the CPS ASEC

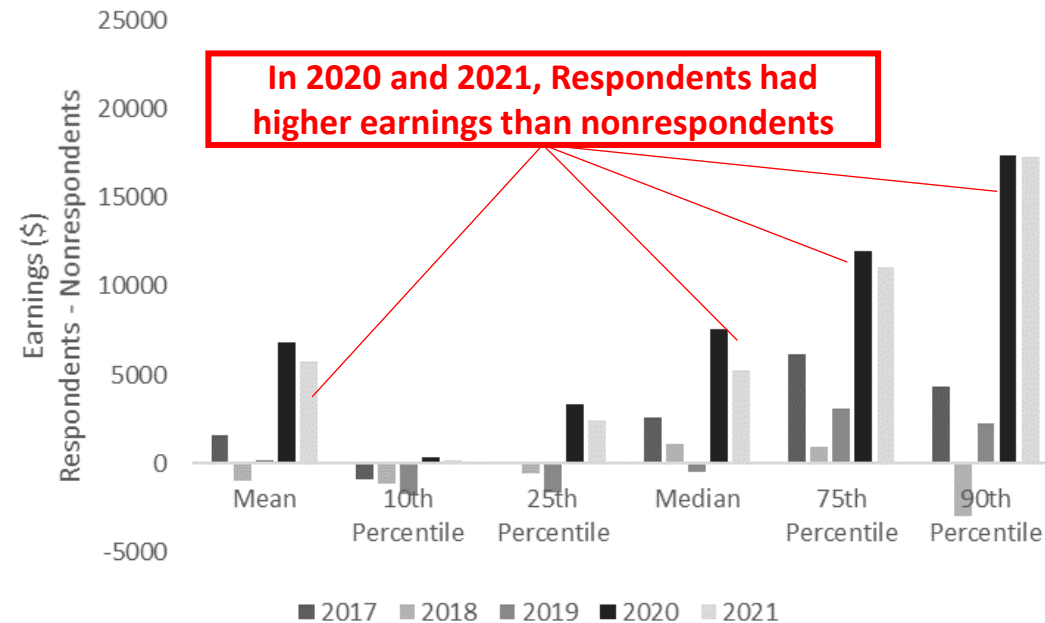
Income Difference between Respondents and Nonrespondents



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. This shows total W-2 earnings at that address in the reference year of the survey.

Nonrandom Nonresponse Income Difference between Respondents and Nonrespondents

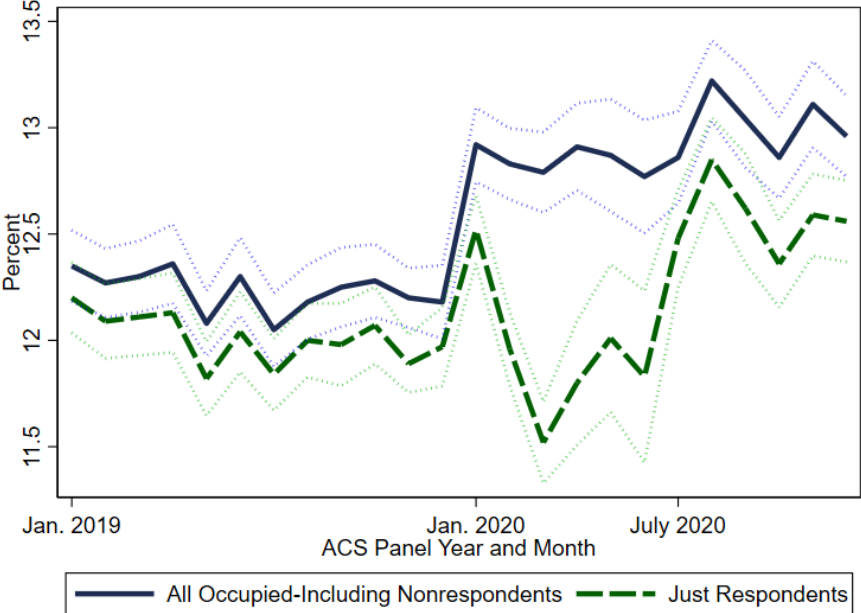


Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This figure shows the difference in income by address between respondents and nonrespondents. This shows total W-2 earnings at that address in the reference year of the survey.

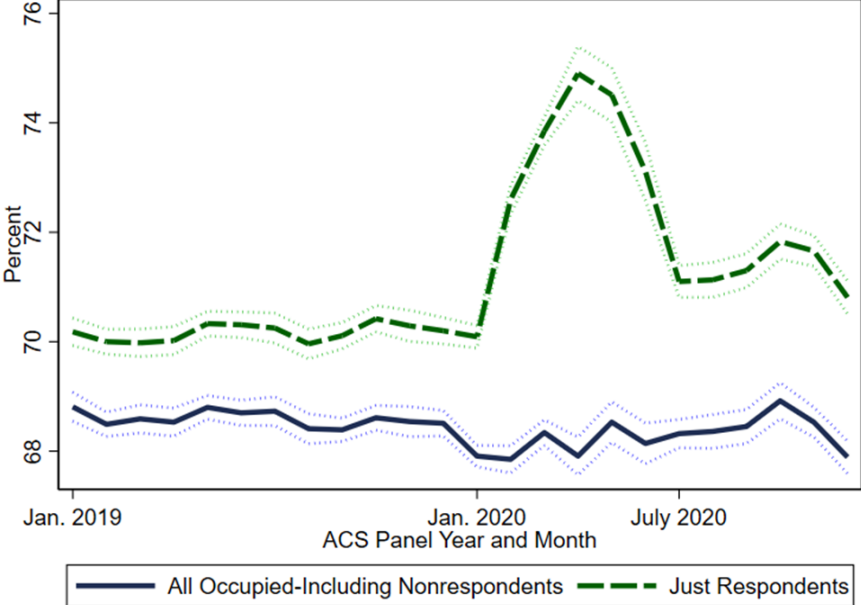
Nonrandom Nonresponse in the ACS

Percent of Households with W-2 Earnings Between \$1 and \$25,000



Source: U.S. Census Bureau, 2019 and 2020 American Community Survey 1-Year Data matched to IRS Form W-2 records.

Percent of Households in Single-Family Homes



Source: U.S. Census Bureau, 2019 and 2020 American Community Survey 1-year data matched to the Master Address File.



U.S. Department of Commerce
 Economics and Statistics Administration
 U.S. CENSUS BUREAU
 census.gov

Weighting in the CPS ASEC and 1-Year ACS

- CPS ASEC
 - Noninterview adjustment – rural/urban response adjustment
 - Pop controls – race/Hispanic-origin x age x gender x state
- ACS
 - Household Noninterview Adjustment – tract and building type (e.g., single-family home vs. apartment) response adjustment
 - Pop controls – adjust to occupied housing unit counts and pop controls (race/Hispanic-origin x age x gender) at substate level (counties/group of counties)

Reweighting with Linked Data and Pop Controls

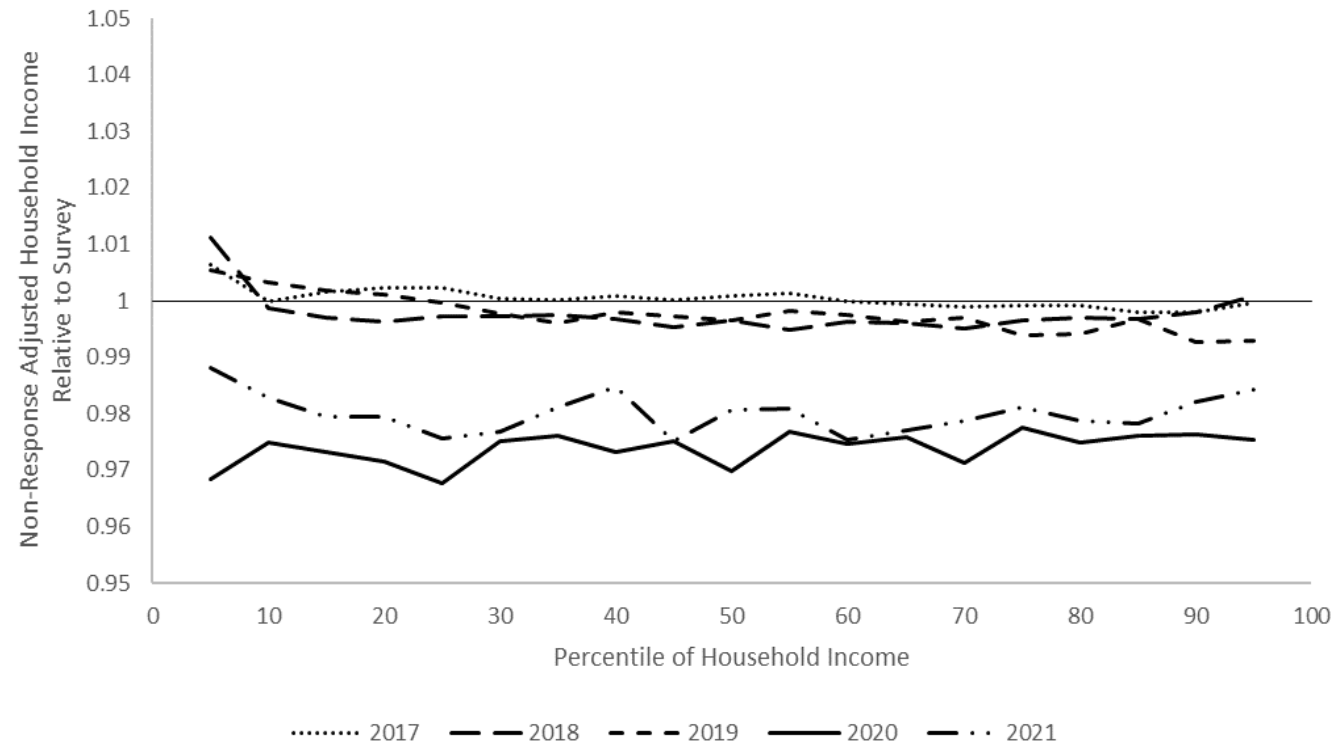
- Hainmueller, 2012. “Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies”
 - Reweight a control group to match the characteristics of a treatment group
 - Application of exponential empirical calibration (Deville and Särndal, 1992)
 - Treat reweighting as an optimization problem, rather than starting from a regression

Entropy Balance Weights

- Adjust weights to:
 - Address-level statistics from linked administrative, survey, and decennial census data
 - Person-level population controls

Results - CPS ASEC

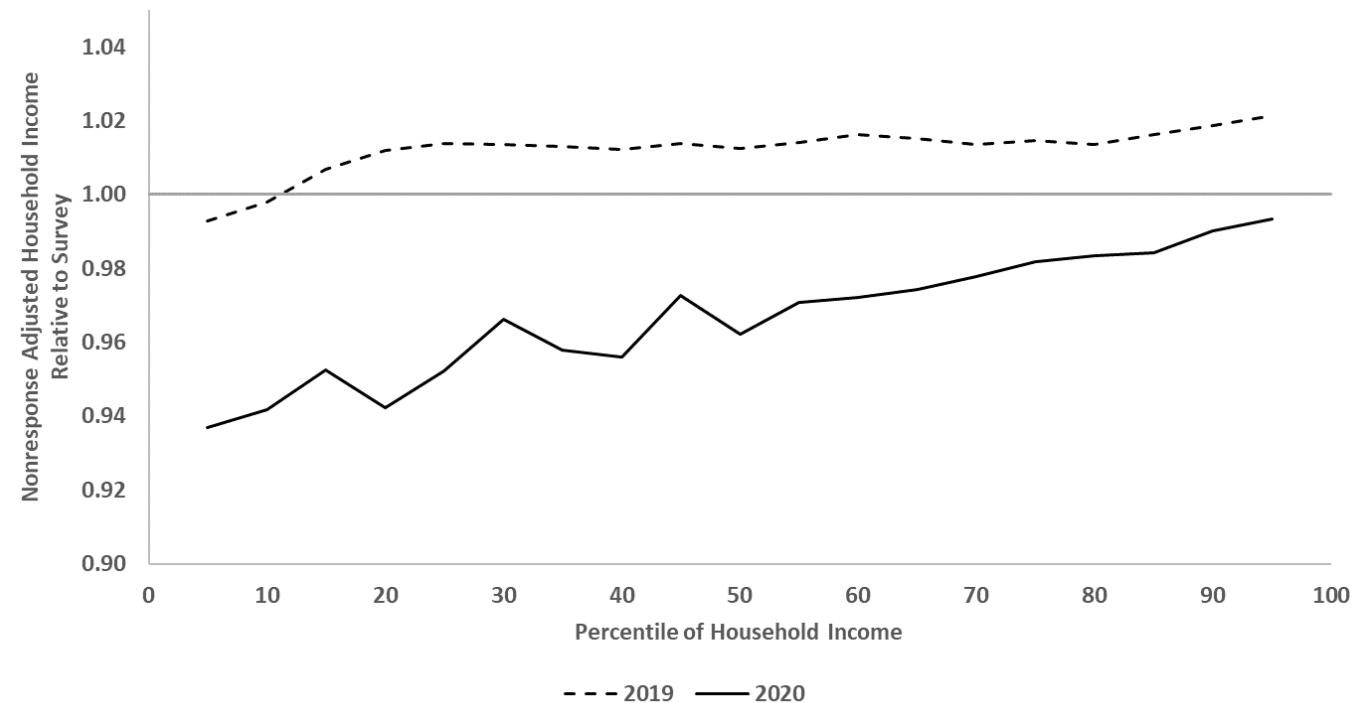
- No evidence of nonresponse bias for income or poverty from 2017-2019
- Adjustments in 2020 and 2021
 - Income ↓ by 2-3 percent across the distribution
 - Poverty ↑ by ~0.45 p.p.
 - About 1.5 million people



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.

Results – ACS

- Bias in income and poverty before and during the pandemic
- Adjustments
 - Before the pandemic (2019)
 - Income ↑
 - Poverty ↓
 - During the pandemic (2020)
 - Income ↓
 - Poverty ↑



Source: U.S. Census Bureau, 2019 and 2020 American Community Survey 1-year data. For more information on sampling and estimation methods, confidentiality protection, and sampling and nonsampling error, in the ACS, visit [2020 ACS 1-Year Experimental Data Tables \(census.gov\)](https://www.census.gov/data/tables/2020/acs/1-year-experimental-data-tables.html).

Notes: This figure shows estimates of the change in real household income (adjusted by the CPI-U-RS) using the survey and experimental entropy balance weights at each 5th percentile from the 5th to 95th. All estimates are linear interpolations across bins of \$2,500.

Public-Use Weights (CPS ASEC)

- Problem – weights were estimated by conditioning on a high-dimensional set of administrative characteristics
 - How to show that they did not disclose information about respondents?
- Solution
 - Don't
 - Make new weights that only condition on information in the survey
- Weights available at <https://www.census.gov/data/datasets/time-series/demo/income-poverty/data-extracts.html>

Future Plans

- Extend the work to other surveys, currently or planning to work on:
 - ACS 5-year file – additional administrative data and better weighting to small area controls (census tracts and counties)
 - Consumer Expenditure Survey (joint with BLS)
 - Household Pulse Survey
- Other projects that use linked survey and administrative data
 - National Experimental Wellbeing Statistics (NEWS) – part of a project to improve income/resource estimates with survey and administrative data
 - 1940 decennial census – for use with linked data to estimate long-term outcomes of children
- Stata package that wraps around ebalance.ado file to easily do survey reweighting using entropy balancing with multiple sets of moment conditions (mostly complete)

Contact Information

Jonathan Rothbaum

Research Economist

Social, Economic, and Housing Statistics Division

jonathan.l.rothbaum@census.gov

(301) 763-9681