Correcting Under-reporting in Survey Income

Jonathan Rothbaum, Ph.D.
Income Statistics Branch
Social, Economic, and Housing Statistics Division

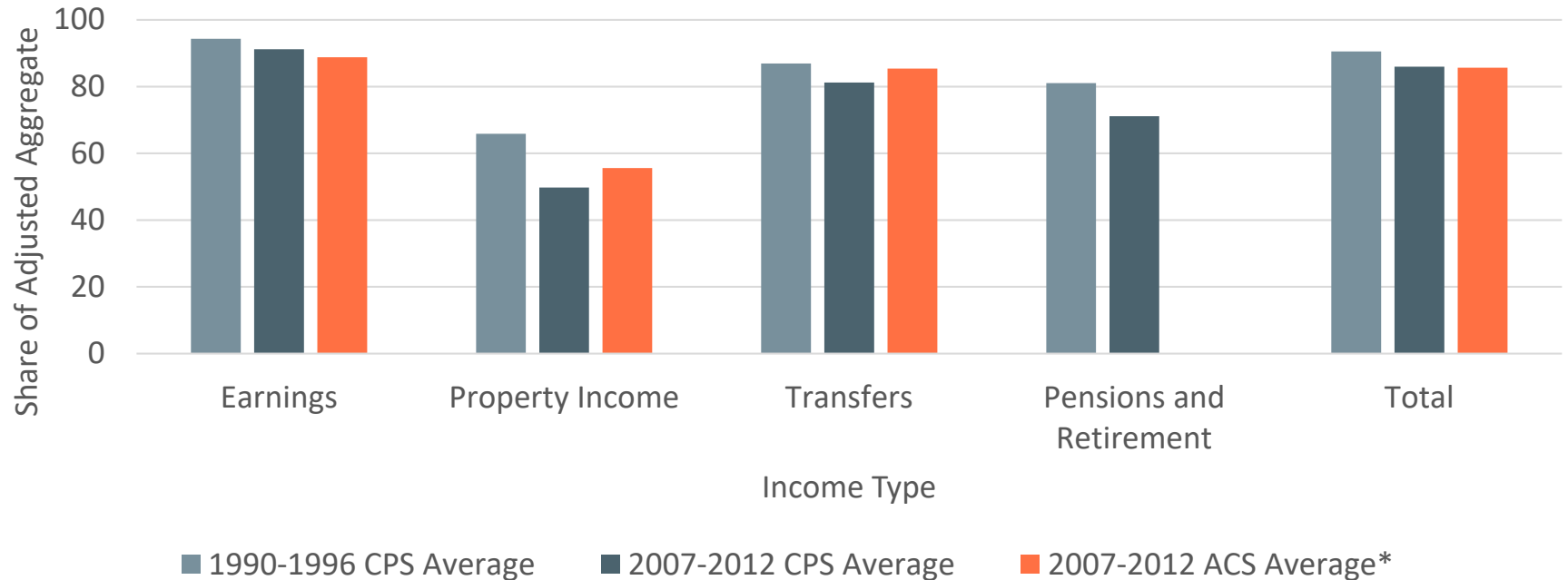BEA Advisory Committee Meeting
May 15, 2020

Biases in Survey Income

- Measurement error
  - Misreporting
  - Under-reporting

- Non-response/Imputation bias
  - Non-response increasing
  - Imputations could bias results if assumptions of imputation model are not true

Underreporting – Looking at the Aggregates

Share of NIPA Aggregate in Census Surveys

Source: Rothbaum (2015)
* ACS Transfers includes both Transfers, Pension, and Retirement Income due to the
  lower level of detail in the questionnaire.

Possible Options

1. **Modeling based on external aggregates**
   - Examples – Urban Institute's TRIM Microsimulation Model, CBO's regression-based adjustments, and Fixler and Johnson (2014)
   - Potential for mis-allocation based on reported characteristics

2. **Corrections based on parameters from linked data**
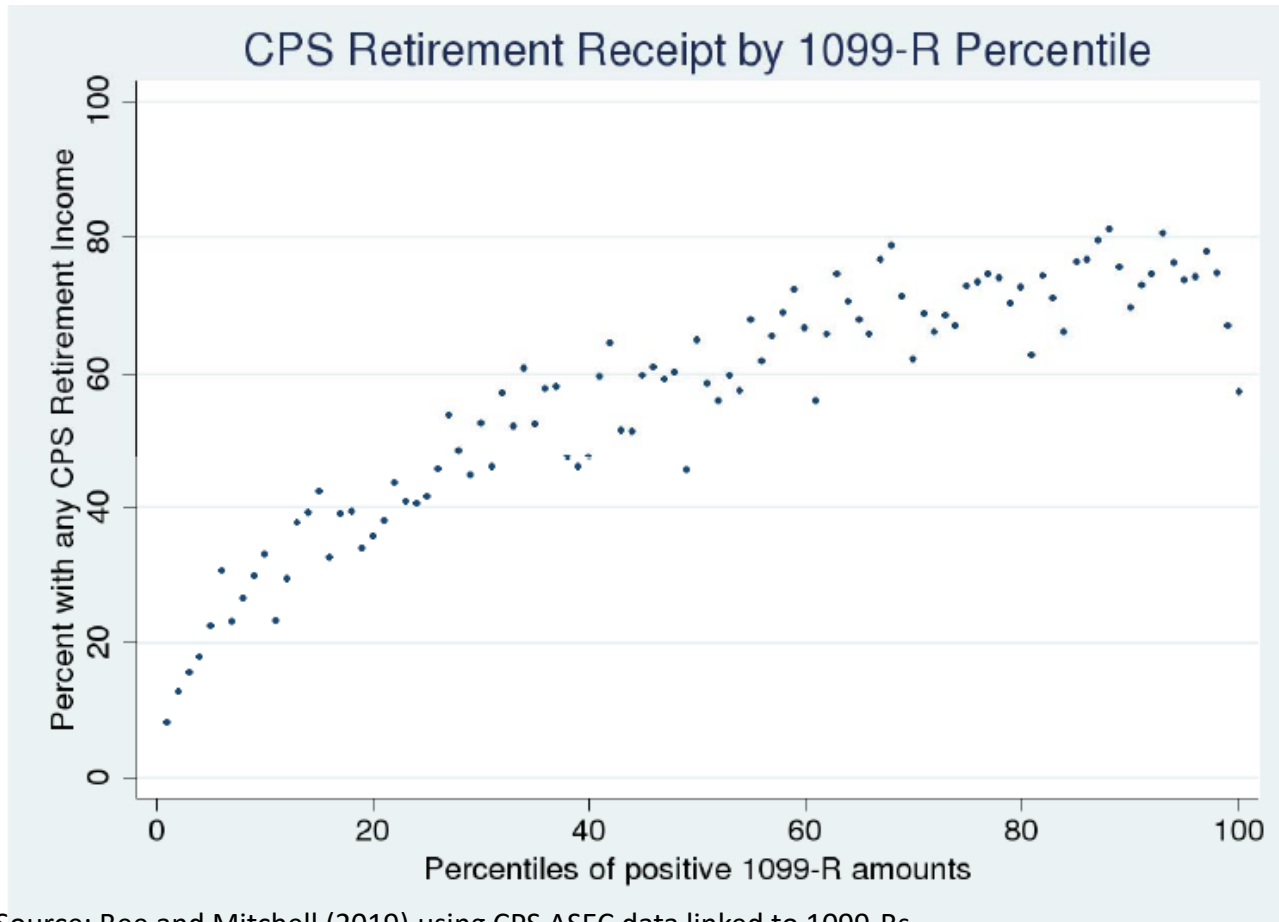   - Estimate relationship of survey reports to "true" income using linked data
   - Release parameters that summarize that relationship
   - Examples – Mittag (2019) SNAP under-reporting adjustment and Fixler, Gindelsky, and Johnson (2019) Pareto adjustment

3. **Estimates using linked data**
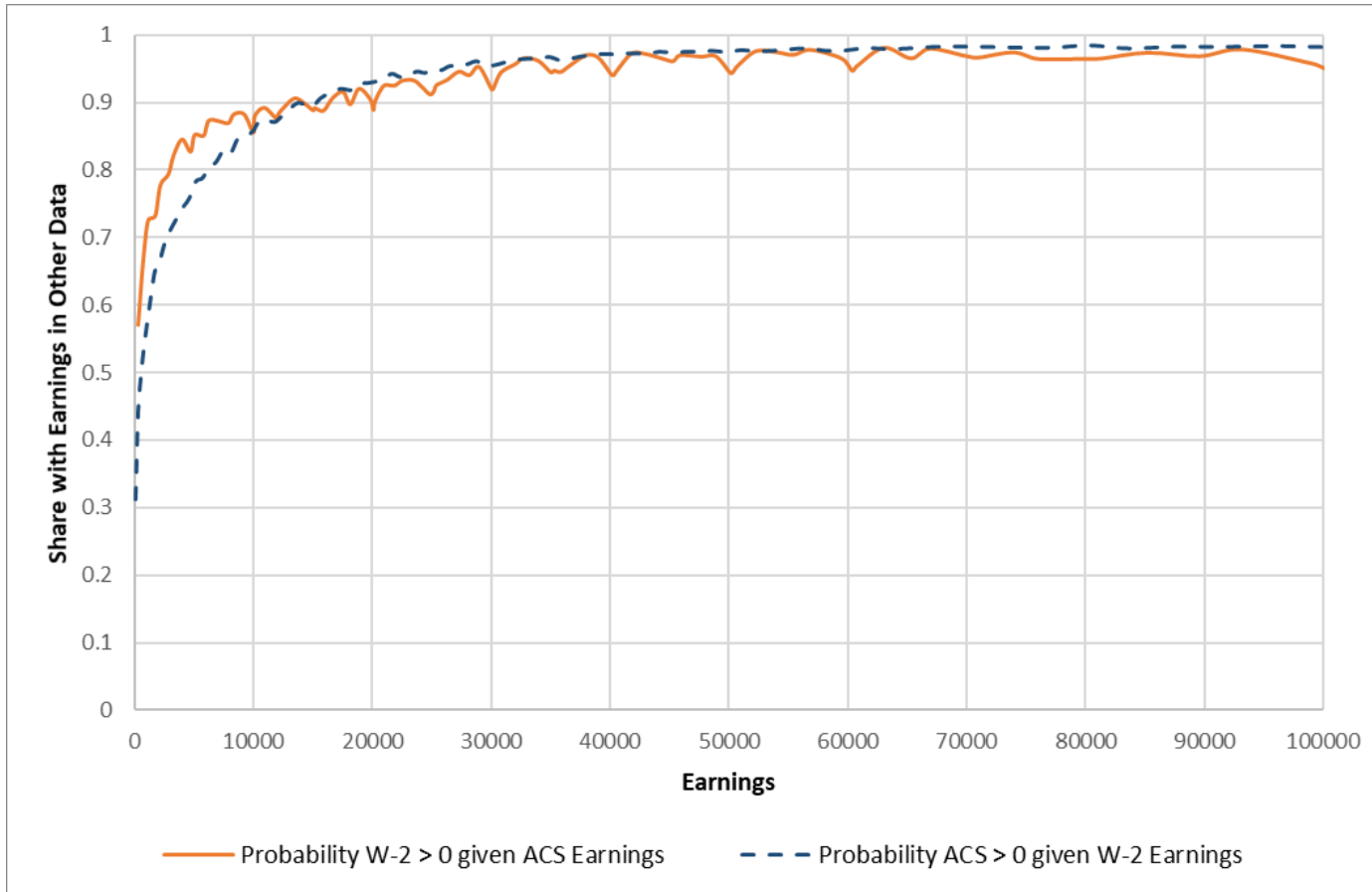
Modeling Based on External Aggregates

- Under-reporting is often a problem at the extensive margin
  - Scaling up gives missing income to the wrong people

Extensive Margin Under-reporting Retirement Income

Source: Bee and Mitchell (2019) using CPS ASEC data linked to 1099-Rs.

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

Extensive Margin Disagreement
Wage and Salary Earnings

Source: Bee, Mitchell, and Rothbaum (2020) using ACS data linked to W2s.

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

Earnings – Extensive Margin Agrees, How About the Intensive?

Comparing Survey to Administrative Earnings
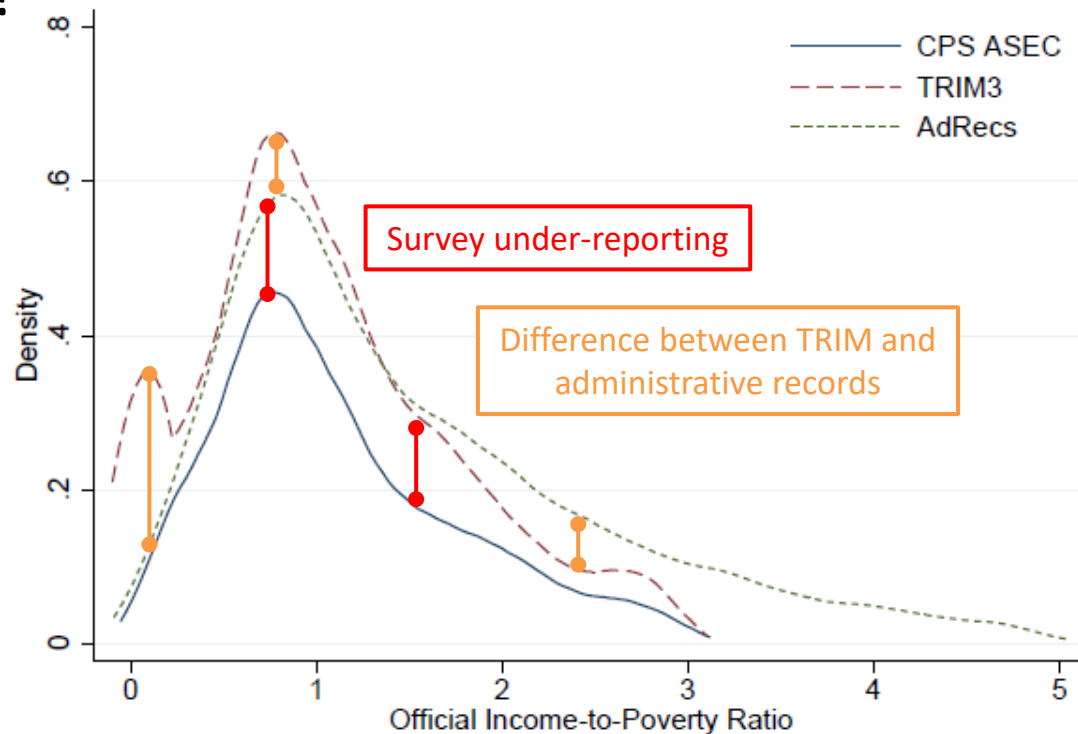By Percentile of Administrative Earnings
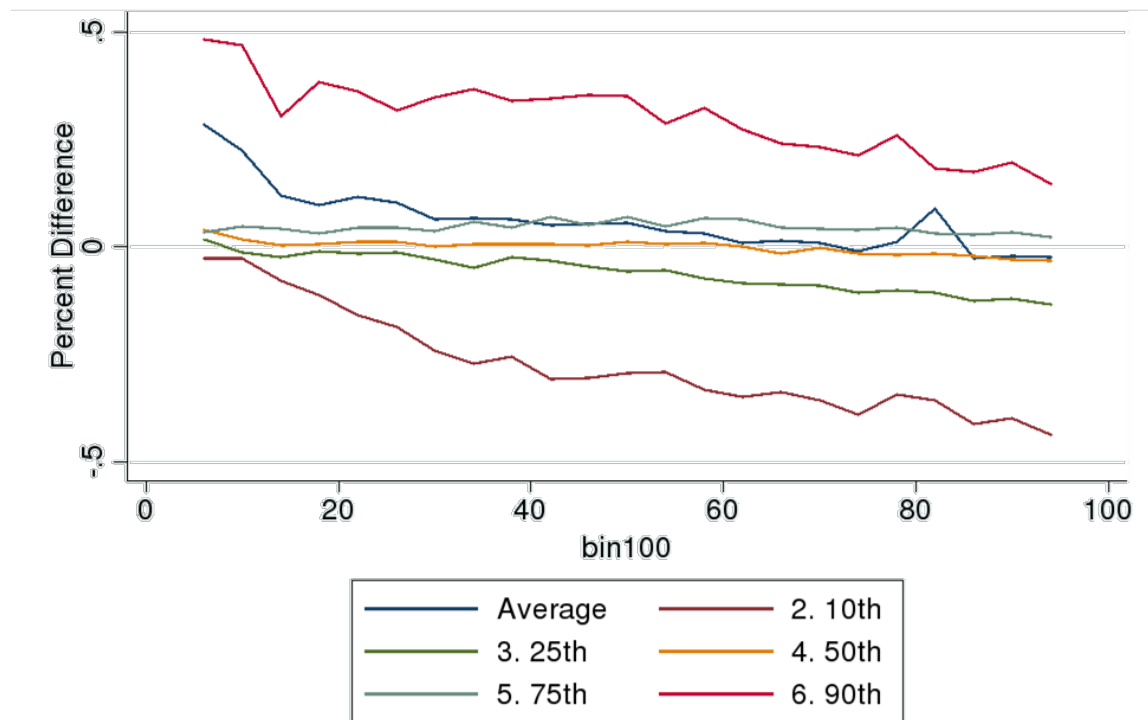
Source: Bee, Mitchell, and Rothbaum, 2020.
2013 CPS ASEC linked with 2012 DER, 2012 W-2s, and 2012 LEHD.
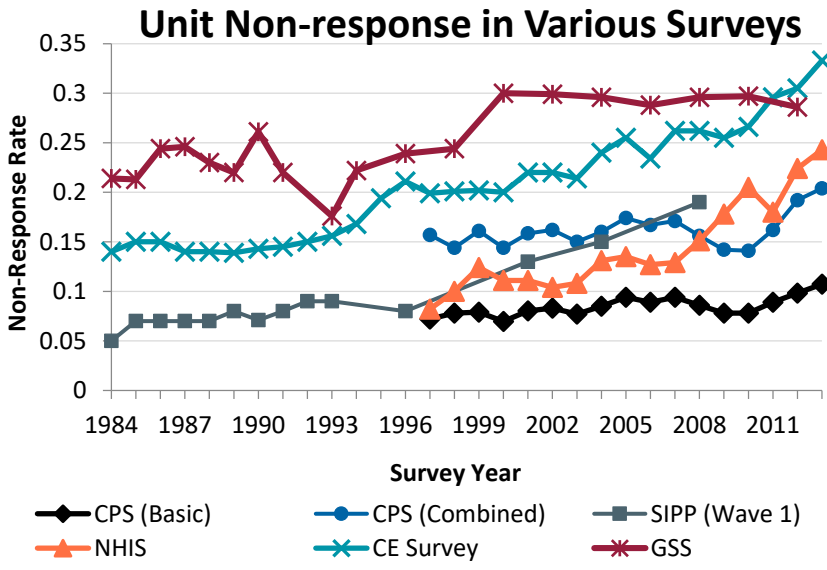
Correcting for Under-reporting in SNAP Who Gets Benefits with Survey, TRIM, and Adrecs?

Survey under-reporting
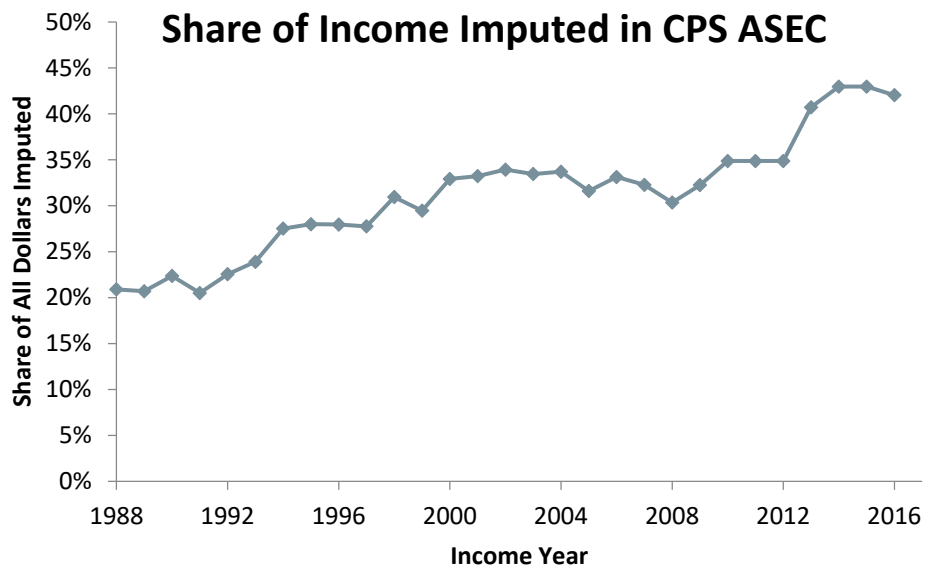
Difference between TRIM and administrative records

Source: Shantz and Fox, 2019 using the 2010–2016 Annual Social and Economic Supplements (CPS ASEC), Transfer Income Model version 3 (TRIM3), and state Supplemental Nutrition Assistance Program (SNAP) administrative records (AdRecs).

Note: Adjusted using IPW, excluding full line imputes, excluding imputed SNAP receipt and amount, and excluding the top and bottom five percent of observations. The densities have been scaled based on the rates of SNAP receipt. The density for the administrative records curve is one. The unit of analysis is the SPM unit. Values are conditional on positive SNAP benefits in each data source. For information on confidentiality protection, sampling error, non-sampling error, and definitions, see https://www2.census.gov/programs-surveys/cps/techdocs/cpsmar16.pdf.

Non-response over time

Unit Non-response in Various Surveys

Source: Bee, Gathright, and Meyer (2016)

Legend:
- CPS (Basic)
- CPS (Combined)
- SIPP (Wave 1)
- NHIS
- CE Survey
- GSS

Share of Income Imputed in CPS ASEC

Source: Hokayem, Raghunathan, and Rothbaum (2019)

United States Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov
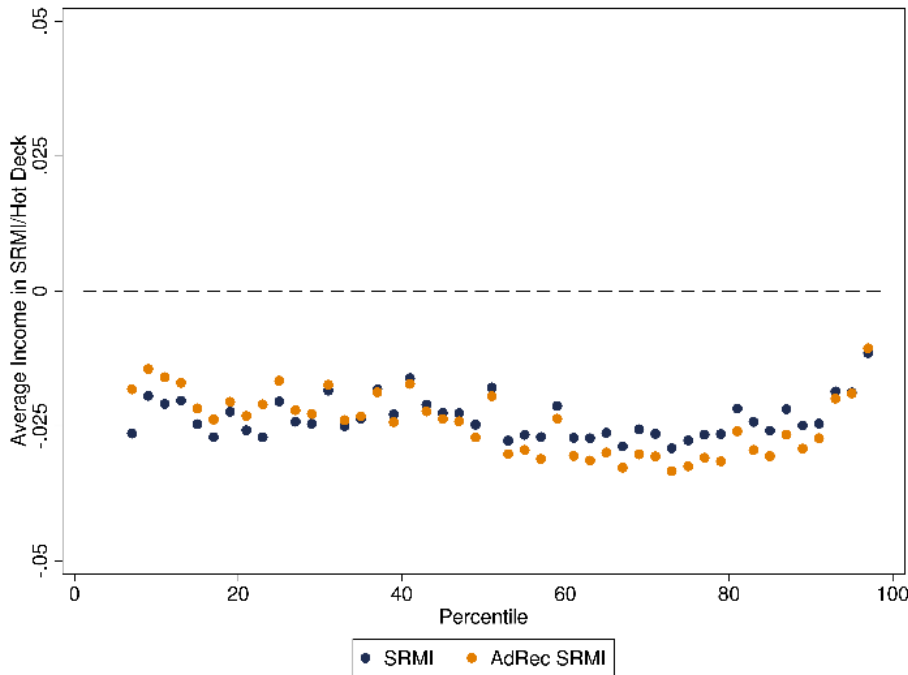
10

Non-response – is it random?

- Trouble in the Tails (Bollinger et al., 2018)
  - High/Low earners most likely to be non-respondents
  - Result in biased income distribution statistics
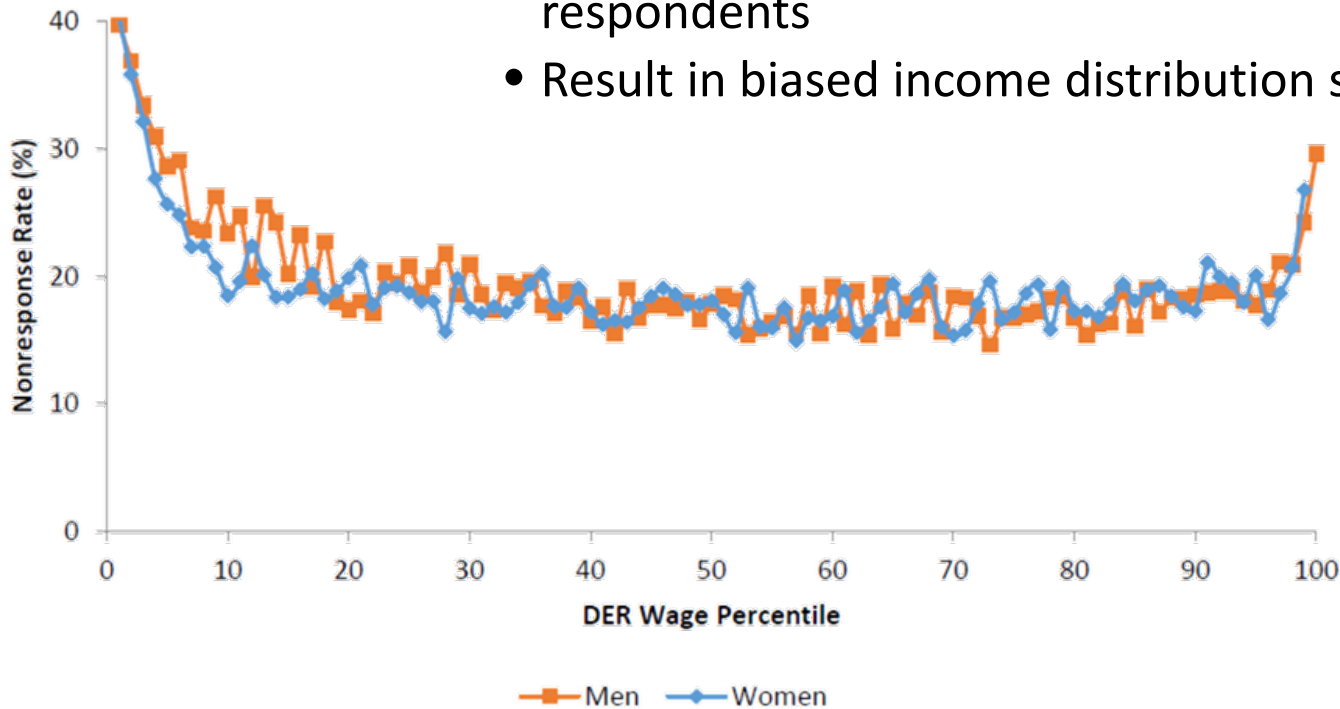
Source: Bollinger et al. (2018) from CPS ASEC linked to W2 records

United States Census Bureau

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

Non-random Non-response – how it affects income estimates?

- Imputation with administrative records (Hokayem, Rothbaum, Raghunathan, 2019)
  - Address trouble in the tails non-random non-response/imputation bias
- Results – correcting for bias
  - Poverty ↑
  - Median household income ↓
  - Inequality ↑
  - Adrecs help with precision, not necessary for non-response bias

Source: Hokayem, Raghunathan, and Rothbaum, 2019. 2009 CPS ASEC linked with administrative records.

Administrative Data has Error Too

- Under-reporting of earnings
  - Wages – under-the-table
  - Self-employment
    - I'll defer to Splinter's comments on this, mostly
    - FYI – adjusting surveys is made even harder by wage/self-employment category reporting differences in survey and administrative data
      - Abraham et al. (2018)
        - $p(\text{CPS SE} \neq 0 | \text{DER SE} > 0) = 0.49$
        - $p(\text{DER SE} > 0 | \text{CPS SE} \neq 0) = 0.35$

Earnings – Which Source is Right when They Disagree?

For low earners, a lot of earnings in surveys isn't in administrative data. Is it survey mis-reporting or error in administrative data?

DER-CPS Percentage Difference, PIK Sample
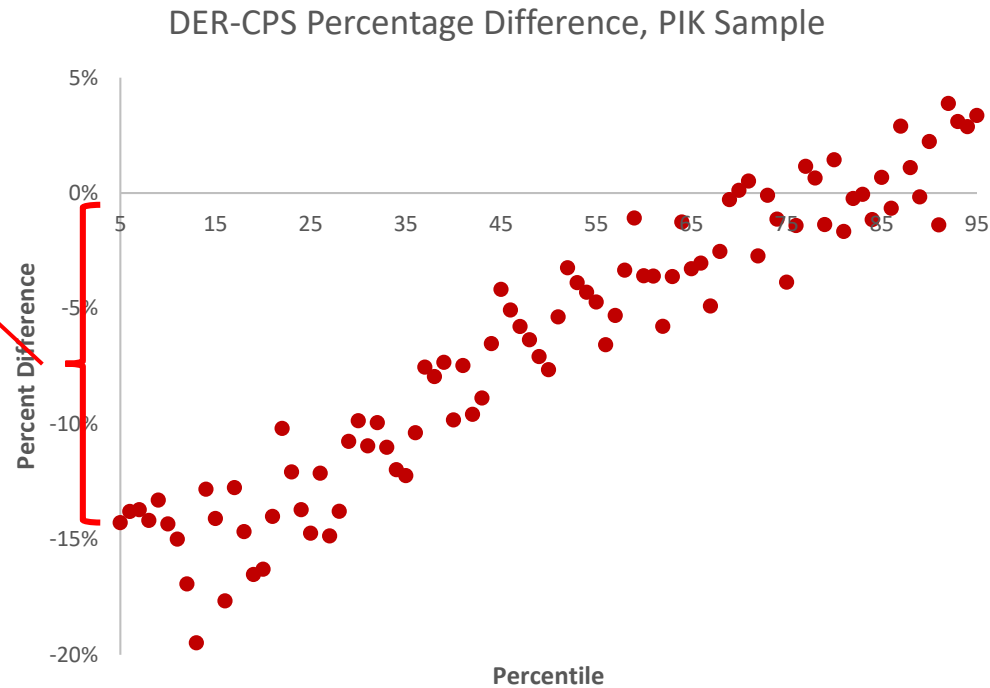
Source: Bee, Mitchell, and Rothbaum (2019) using the 2013 CSP ASEC linked to administrative data.

Our Approach - Earnings
(Bee, Mitchell, and Rothbaum 2020)

1. Use Job-Level Information to get "best possible" administrative job-level earnings

2. Compare to 1040 to check for missing earnings (at tax-unit level)

3. Compare to survey and decide for which individuals to use survey earnings

4. Final "best" estimate of earnings for each individual/household

W-2

DER

LEHD

Best Job Earnings

1040

Best Adrec Earnings
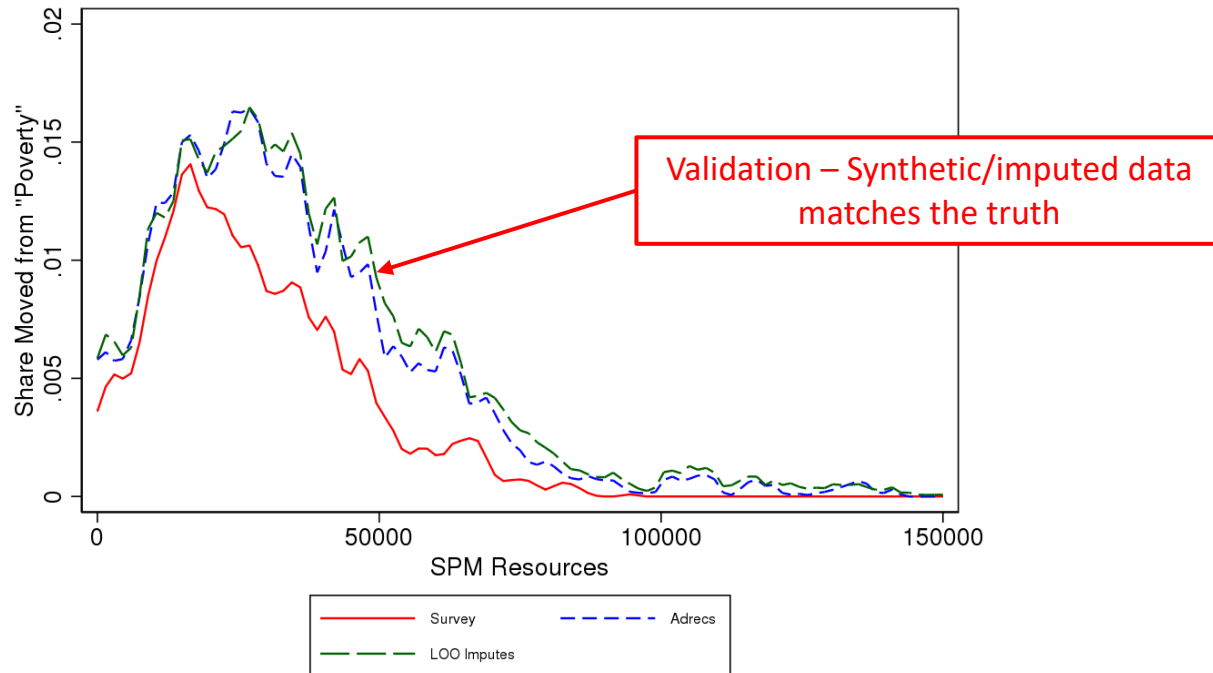
Survey

Final Earnings Estimate
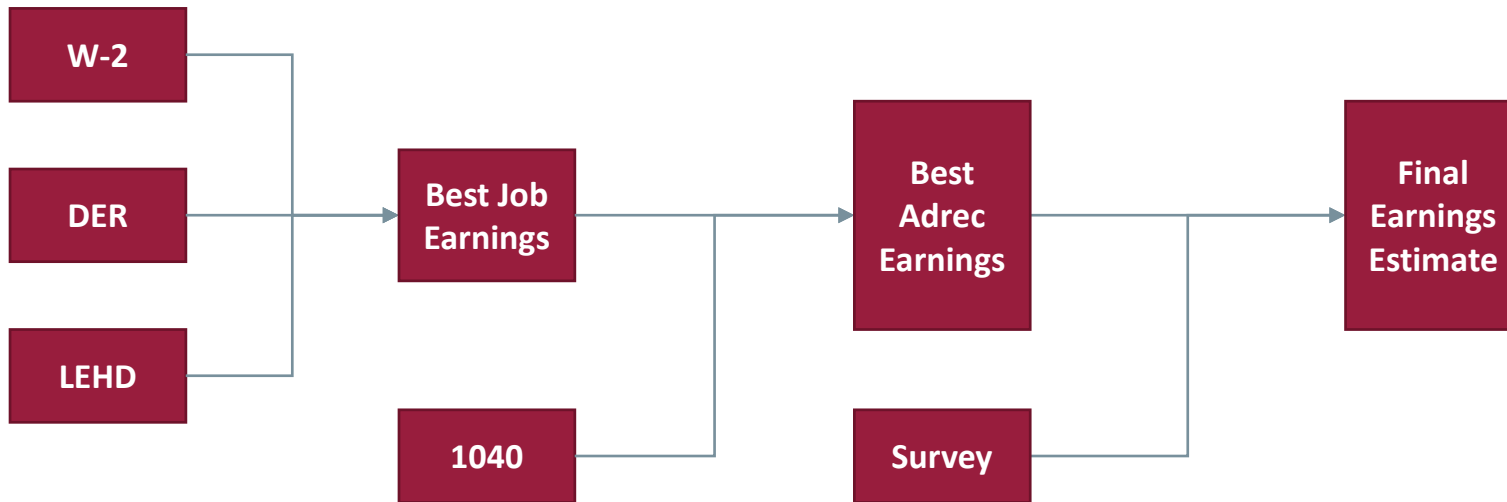
Our Approach – Under-reported Benefits
(Fox, Rothbaum, and Shantz 2020)

- Treat as a problem of missing information
  - For SNAP – data available in some states, but not all
- Impute "true" benefits to individuals with missing data
  - General method – could be used for data availability over time
  - Timely estimates – estimate for t using administrative data from prior years ($t-1$, or $t-s$)
  - Time series – estimate for t using data from $t+s$

Our Approach – Under-reported Benefits
(Fox, Rothbaum, and Shantz 2020)

Source: U.S. Census Bureau, Current Population Survey, 2014 Annual Social and Economic Supplements (CPS ASEC) and state Supplemental Nutrition Assistance Program (SNAP) administrative records.

Mobility curves (Foster and Rothbaum, 2014) trace out the share of the population that SNAP moves out of "poverty" measured at all SPM resource levels from $0 to $150,000 in $1,500 intervals. For example, if $27,000 were the poverty line, administrative reports of SNAP would move 1.7 percent of individuals out of poverty. At the same income level, the survey reports move 1.1 percent of individuals out of poverty.

Contact Information

Jonathan Rothbaum

Chief, Income Statistics Branch

Social, Economic, and Housing Statistics Division

jonathan.l.rothbaum@census.gov

(301) 763-9681

United States™ Census Bureau

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