Discussion: Progress on Building an Integrated System of Data and Statistics on Household Income, Consumption, and Wealth

Katharine G. Abraham, University of Maryland

Federal Committee on Statistical Methodology Conference

October 23, 2024

Why develop a system of ICW data and statistics?

- Improve understanding disparities in economic well-being and intergenerational economic mobility
- Assess retirement security
- Evaluate effects of economic shocks and economic policy changes
  - Impact of the COVID-19 pandemic and associated cash transfers
  - Consequences of shifts to alternative work arrangements
- Improve budget and tax modeling
- Create checks on existing data, helping to identify potential measurement problems

What should official statistics measure?

- Great appeal to having a core set of published statistics that satisfy some basic principles
  - Consistent with basic accounting identities:
    - $S = I - C$
    - $\Delta W = S$
  - Grounded in NIPAs
- Different definitions may be appropriate for different purposes
  - Household well being (Bee et al.; Garner, Matsumoto and Schild) versus distribution of national output (Gindelsky and Martin)
- Adoption of a set of standardized terms and definitions will increase value of published data

Many challenging conceptual issues

- Example: Valuation of health insurance
  - Gindelsky and Martin: Employer contributions and public health spending allocated to households as part of personal income and counted in consumption
  - Bee et al.: Don't currently have data, but discussion in paper suggests that, if available, would count cost of health insurance as part of income
  - Garner, Matsumoto and Schild: Allow for possibility that recipients value a dollar of health insurance at less than a dollar
- Example: Treatment of retirement savings and benefits
  - Gindelsky and Martin: Retirement benefits other than Social Security not included in income; part of reason for negative savings rates at older ages
  - Bee et al.: Retirement benefits included in income
  - Garner, Matsumoto and Schild: NA
- Preferred approach may depend on purpose for which data are to be used

Many challenging measurement issues

- All three empirical projects start with survey data. Measurement challenges include
  - Mismeasurement
  - Item nonresponse
  - Absence of needed variables on survey
  - Unit nonresponse
- Projects have made progress through imputation and record linkage
  - Gindelsky and Martin: Develop rules for allocating more than 70 components of Personal Income to CPS-ASEC sample; construct comparable estimates of income in CEX; impute consumption to CPS-ASEC sample
  - Garner, Matsumoto and Schild: Start with CEX; impute values for types of consumption survey does not measure directly
  - Bee et al: Link to multiple sources of information on income; since ability to link not random, reweight after linking

Many challenging measurement issues (continued)

- Imputations can be very sensitive to assumptions. A few examples:
  - How should income estimated to be missing from tax returns be allocated across households?
  - What is the right interest rate or rates to use for capitalizing flows of income from financial assets as reported in tax data?
- Administrative data linked to survey data may themselves suffer from measurement error
- Bee et al provide a nice illustration of how assumptions matter
  - NEWS and CID both seek to improve the measurement of income by incorporating administrative data
  - With same concepts, poverty rate considerably lower in CID than NEWS
    - Differences in how discrepancies between survey and administrative income treated
    - Differences in adjustments for shortfalls in aggregate EITC amounts

Discrepancies across measures

- Expect eventual convergence to a standard set of concepts and measurement best practices
  - Statistical system leadership can help
- Differences in source data still likely will mean differences in statistics published by different agencies
  - Important to reconcile alternative series and explain why they differ
  - A possible model: Periodic reports prepared by BLS and BEA staff to explain the differences between the CPI and the PCE deflator

Creating a microdata infrastructure file

- Household survey sample (CPS-ASEC, CEX, SCF)
  - Pros: By design, representative of the population; lots of demographic information; great work already well underway
  - Cons: Sample size limits potential for disaggregation

- Census Bureau Master Address File (MAF) or tax data
  - Pros: Cover (almost) the entire population (not just a sample)
  - Cons: Limited demographic information

Access to underlying microdata

- Glad to see that code and (most of the) data generated through the Bee et al. project will be available through the FSRDCs

- Creation and dissemination of distributional microdata could be a terrific NSDS pilot project
  - Privacy protecting data linkages
  - Provision of secure access
  - Methods of review for proposed research outputs
  - Possible creation of synthetic data files

Katharine G. Abraham
University of Maryland
kabraham@umd.edu