Innovations in Income and Mobility Measures

Federal Economic Statistics Advisory Committee
Distributional Measures at Census, BEA, and BLS

Jonathan Rothbaum, Social, Economic, and Housing Statistics Division, Census Bureau
Nikolas Pharris-Ciurej, Center for Economic Studies, Census Bureau

June 10, 2022

Any opinions and conclusions expressed herein are those of the author(s) and do not reflect the views of the U.S. Census Bureau.
Income and Poverty in the United States: 2020

Current Population Reports

By Emily A. Sluder, Melissa Kotlar, Frances Chen, and Jessica Senege

Issued September 2021

Table A-5
Selected Measures of Equivalence-Adjusted Income Dispersion: 1967 to 2020


<table>
<thead>
<tr>
<th>Year</th>
<th>Measures of income dispersion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Shares of equivalence-adjusted income of quintiles</td>
</tr>
<tr>
<td></td>
<td>Lowerest</td>
</tr>
<tr>
<td>2020</td>
<td>3.4</td>
</tr>
<tr>
<td>2019</td>
<td>Table A-4a.</td>
</tr>
<tr>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td></td>
</tr>
</tbody>
</table>

Selected Measures of Household Income Dispersion: 1967 to 2020


<table>
<thead>
<tr>
<th>Year</th>
<th>Measures of income dispersion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Household income at selected percentiles</td>
</tr>
<tr>
<td></td>
<td>10th</td>
</tr>
<tr>
<td>2020</td>
<td>15,699</td>
</tr>
<tr>
<td>2019</td>
<td>15,226</td>
</tr>
<tr>
<td>2018</td>
<td>15,086</td>
</tr>
<tr>
<td>2017</td>
<td>11,181</td>
</tr>
</tbody>
</table>

- Income estimates back to 1967
- Gini, shares by quintile, percentiles, percentile ratios,...
Income and Poverty Statistics at Census

- Poverty estimates back to 1959
Income and Poverty Statistics at Census

- Many, many historical income and poverty tables updated annually
Income and Poverty Statistics at Census

- Historical data from ACS and decennial long form censuses for smaller geographies
- Small Area Income and Poverty Estimates (SAIPE) for improved income and poverty estimates for small areas and school districts
National Experimental Well-Being Statistics Project (NEWS)

• Goal - rethink how we can produce income and resource statistics
  • What is the best possible estimate given all the data currently available at Census for a given income/resource statistic?
  • Expand the set of income and resource statistics we produce
Why Does This Matter?
Survey Underreporting

Source: Rothbaum (2015)
* ACS Transfers includes both Transfers, Pension, and Retirement Income due to the lower level of detail in the questionnaire.
Why Does This Matter?
Misreporting Example – Income for Age 65+ Households

Other Goals

• Experimental
  • Updated regularly with additional data and better methods
  • Expand income/resource concepts being measured
  • Longer term – move to regular production

• Transparent and replicable
  • Decisions about how to use survey and administrative income are well-documented, supported, and apolitical
  • Create linked microdata and code database that is accessible through the RDC system
  • Long term – create a set of synthetic data sets (akin to the SIPP Synthetic Beta) for public release?

• Timeline
  • 2022 – 1st set of statistics for a year or small number of years
  • 2023- Additional statistics, additional years, improved methods,...
Which Statistics?

   • Same general statistics we produce in existing official reports
     (simple moments, distributional/inequality statistics, poverty, etc.)
Which Statistics?

2. Longitudinal Income, Resource, and Poverty Statistics – MOVS project (Mobility, Opportunity, and Volatility Statistics)
   • Income and earnings dynamics
Source Data – Survey and Census Data

• Information not available in administrative data
  • Demographics and socioeconomic characteristics (Race, education, etc.)
  • Income and benefits – address linkage and income coverage issues
  • Survey frames potentially provide sampling information needed for estimates
    (random sampling + vacancy assessments)

• Including:
  • CPS ASEC
  • ACS
  • Decennial Census
Source Data – Administrative Data

• IRS and SSA income data
  • 1040, W-2, 1099-R, 1099-IRMF, DER, social security and SSI payment data, etc.
• LEHD
• Numident
• Master Address File
• State and federal program data
  • SNAP, TANF, WIC, HUD, VA, Medicare/Medicaid data, etc.
• Firm Data
  • Business Register, Longitudinal Business Database, Form 5500 filings
• Third-Party Data
  • Black Knight data on home values
General Approach

Master Address File
- Survey samples (with vacant units identified)
- Administrative data (address linked)

Census data
- Administrative data (person and address linked)

Survey data

Imputation for survey nonresponse
- Address incomplete adrecs
- Imputation of survey information to adrec universe?

Frame (Housing-unit level)

Individual/Household Level

Processing and Estimation
Challenges

• Measurement error in administrative data – earnings in particular
• Linkage challenges – incomplete linkage and errors in linkage
• Coverage and representativeness
• Incomplete geographic coverage of administrative data
• Conceptual misalignment or incomplete income coverage in administrative data
• Timeliness/availability of administrative and survey data
• Changes in administrative data that may be unrelated to changes in the underlying income or resources

Challenge
Measurement Error in Administrative Data

• Earnings - 80% of income
  • Wage and salary earnings is probably the best reported of any income category in surveys (70% of income)
    • Particularly for aggregates and extensive margin agreement
    • Still, error in earnings matters more than in any other income type
Challenge
Measurement Error in Administrative Data

• Wage and salary – under-the-table earnings
  • Detailed occupation level differences in administrative and survey earnings largely match expectations about workers that are likely to be paid under the table (construction, food service/bartending, etc., from Bollinger et al., 2015 and our work with linked ACS data)

• Self-employment – tax avoidance
  • Confirmed by audit studies and consumption/income relationship for the self-employed
  • Nearly ½ of self-employment income in the National Income and Product Accounts is imputed due to under-reporting to the IRS
Combining Survey and Adrec 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 adrec or survey earnings

4. Final “best” estimate of earnings for each individual/household
Response Propensity by W-2 Earnings at Address
(Controlling for linked demographics, such as race and age)

Source: Rothbaum and Bee, 2020. “Coronavirus Infects Surveys, Too: Nonresponse Bias During the Pandemic in the CPS ASEC”
Challenge
Linkage Issues

• Addressing misreporting
  • ~10% of individuals in a survey cannot be linked to their SSN

• Representativeness/Weighting
  • Administrative records may come from nonrepresentative samples
  • Surveys have random samples but nonrandom selection into response

• Linkage error – understudied
Challenge
Incomplete geographic coverage of administrative data

• Some data is only available for some locations (and in some years)
• Examples
  • SNAP
  • LEHD (in some years)
  • TANF
  • WIC
• Missing information problem - Impute
Addressing Incomplete Geographic Coverage
Imputing to States without Adreces

Changes in Administrative Data

- Can change over time due to statutory/regulatory changes that affect programs and agencies

Graph: Percent of PIKed Households with a 1040 Filed

Economic Stimulus Act of 2008 provided a tax credit for those with low income → many more 1040s filed for one year

Source: Rothbaum, 2018
2006-2013 ACS Linked to IRS records.
Changes in Administrative Data

• Can change over time due to statutory/regulatory changes that affect programs and agencies
  • Auten and Splinter (2018) argue that much of the inequality increase in tax data from 1960 to present in work by Piketty, Saez, and Zucman is due to changes in the tax code and the nature of tax reporting, not in actual underlying income changes
Outline

• Enhance Income Measurement and Statistics Using Expanded IRS Data
• Intergenerational Mobility
• Intragenerational Mobility
• Income Distributions
Enhance Income Measurement and Statistics Using Expanded IRS Data
Evaluation of Potential Benefits of Expanded IRS Data

• Virtually all Title 26 data at Census falls under the current limits of 6103(j)
  • Limited individual tax forms such as 1040s, W-2s, and 1099-Rs with only select fields provided
  • No clear way to assess data completeness, presence of duplicates, or amended returns

• Goal of this project is to illustrate benefits of expanded tax data obtained under 6103(n) to the American Community Survey (ACS)
  • Validate survey income responses with new administrative data analogs
  • Gain a better understanding of current tax data obtained under 6103(j)
  • Produce enhanced repeated cross-sectional statistics measuring the distribution of income
Preliminary Findings on Self-Employment Income
2011 ACS-IRS Data (TY 2010)

- IRS self-employment income could be from a variety of sources
  - Sole proprietorships (includes independent contractors)
  - Partnerships
  - Corporations (S vs. C)

- ACS Question 43b
Misalignment Between ACS and IRS Self-employment
By Type of IRS Self-Employment Income
Preliminary Findings on “Other Income”
2011 ACS-IRS Data (TY 2010)

• Final component of income question intended to capture remaining elements of money income (regular income)
  • Explicitly excludes one-time payments such as realized capital gains
  • Difficult to fully validate given its breadth and inclusion of some non-taxable components

• ACS Question 43h
Extensive Margin Misreporting, Other Income vs. Capital Gains

ACS 2011 compared to TY2010 IRS 1040

P(Any IRS 1040 Capital Gains) vs. Percentile of the ACS Other Income Distribution
In Progress and Moving Forward

• Comprehensively validate income questions for 2011-2019 ACS
  • What factors explain discrepancies?
  • Are discrepancies between ACS and IRS data changing over time?

• Develop new imputation models to address item non-response

• Produce enhanced median income, poverty, and inequality series based on linked data

• Future research may include validating other ACS characteristics (e.g., health insurance, school enrollment)
Intergenerational Mobility
Opportunity Atlas

• Comprehensive census tract-level atlas of children’s outcomes in adulthood

• Children’s income distributions, incarcerations rates, and other outcomes in adulthood by parental income, race, and gender

• Users can view data for every census tract in America, overlay their own data, export data for their own analysis, download full datasets, and save images of maps

• Users can also utilize filters to observe a combination of characteristics

• Places with high incarceration, low employment, low income by race groups and sex or places with high income, high college graduation rates by race groups

DRB #: CBDRB-FY18-319.
<table>
<thead>
<tr>
<th>22 Outcomes By Parental Income, Race, and Sex</th>
<th>13 Neighborhood Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household Income Frac in Top 20% Based on Individual Income</td>
<td>Median Rent 2012-2016</td>
</tr>
<tr>
<td>Incarceration Rate Frac in Top 1% Based on Individual Income</td>
<td>Job Growth Rate 2004-2013</td>
</tr>
<tr>
<td>Teenage Birth Rate % Staying in Same CZ as Adults</td>
<td>Median Household Income 2012-2016</td>
</tr>
<tr>
<td>Individual Income % Staying in Same Tracts as Adults</td>
<td>Median Household Income 1990</td>
</tr>
<tr>
<td>Fraction Married Household Income Stayed in CZ</td>
<td>Poverty Rate in 2012-2016</td>
</tr>
<tr>
<td>Spouses Income Individual Income Stayed in CZ</td>
<td>Fraction College Grad. in 2012-2016</td>
</tr>
<tr>
<td>Employment Rate Household Income for U.S. Natives</td>
<td>Fraction Non-White in 2010</td>
</tr>
<tr>
<td>High School Graduation Rate Household Income for Immigrants</td>
<td>Foreign-Born Share in 2012-2016</td>
</tr>
<tr>
<td>College Graduation Rate Number of Children</td>
<td>Frac. Single Parents in 2012-2016</td>
</tr>
<tr>
<td>Hours Worked Per Week</td>
<td>Population Density 2010</td>
</tr>
<tr>
<td>Hourly Wages</td>
<td>Density in Jobs in 2013</td>
</tr>
<tr>
<td>Fraction in Top 20% Based on HH Income</td>
<td>Fraction with Short Commutes 2012-2016</td>
</tr>
<tr>
<td>Fraction in Top 1% Based on HH Income</td>
<td>2010 Census Response Rate</td>
</tr>
</tbody>
</table>
Data Sources and Sample Definitions


• Target sample: Children in 1978-83 birth cohorts who were born in the U.S. or are authorized immigrants who came to the U.S. in childhood

• Analysis sample: 20.5 million children
The Geography of Upward Mobility in the United States

Average Household Income for Children with Parents Earning $27,000 (25th percentile)

Note: Blue = More Upward Mobility, Red = Less Upward Mobility
Source: Chetty, Hendren, Jones, Porter 2018

DRB #: CBDRB-FY18-319.
Incarceration Rates for Black Men in Los Angeles with Parents Earning < $2,200 (1st percentile)

WATTS, Black Men:
Share Incarcerated on April 1, 2010
= 44.1% (9.3%)

DRB #: CBDRB-FY18-319.
Incarceration Rates for Black Men in Los Angeles with Parents Earning < $2,200 (1st percentile)

WATTS, Black Men:
- Share Incarcerated on April 1, 2010 = 44.1% (9.3%)

COMPTON, Black Men:
- Share Incarcerated on April 1, 2010 = 6.2% (5.0%)

DRB #: CBDRB-FY18-319.
Incarceration Rates for Hispanic Men in Los Angeles with Parents Earning < $2,200 (1st percentile)

WATTS, Hispanic Men:
Share Incarcerated on April 1, 2010 = 4.5% (2.8%)

COMPTON, Hispanic Men:
Share Incarcerated on April 1, 2010 = 1.4% (0.8%)

DRB #: CBDRB-FY18-319.
Mean Household Income for Black Men in Los Angeles with Parents Earning $27,000 (25th percentile)

WATTS, Black Men:
Mean Household Income = $7,286 ($2,576)

COMPTON, Black Men:
Mean Household Income = $19,141 ($2,149)

DRB #: CBDRB-FY18-319.
Mean Individual Income for Black Women in Los Angeles with Parents Earning $27,000 (25th percentile)

WATTS, Black Women:
Mean Household Income = $19,489 ($1,985)

COMPTON, Black Women:
Mean Household Income = $21,509 ($1,850)

Mean Individual Income for Black Women in Los Angeles with Parents Earning $27,000 (25th percentile)

DRB #: CBDRB-FY18-319.
Moving Forward

• Current Results and Additional Information can be found at: https://www.opportunityatlas.org/

• Coming soon – The Radius of Economic Opportunity: Evidence from Migration and Local Labor Markets – paper, dataset, and data tool

• Future planned work includes:
  • Opportunity Atlas - integrating additional years of data
  • Studying the relationship between social mobility and factors such as placed based policies, labor markets, gentrification, and characteristics at birth, plus additional research on race and ethnicity
  • Studying the relationship between income, race, and mortality
Intragenerational Mobility
Local Area Earnings Inequality and Mobility Statistics

• Goal: Extend measures of inequality and mobility to local areas.
• LEHD administrative earnings data will be used to create an interactive web application, showing both national and MSA level estimates of:
  • Inequality – dispersion of worker earnings at a point in time (typically annual)
  • Volatility – dispersion of the change in worker earnings (short duration, year-to-year)
  • Mobility – movement of a worker from one part of the earnings distribution to another (long duration, multiple years of earnings)
• Estimates will be non-parametric when possible, decomposable, and allow for comparisons across MSA’s, time, and demographic characteristics
Example: Measuring Inequality

• Inequality measures are typically based on the distribution of employment and total earnings across earnings bins (histogram).

• There are various single number measures of inequality that summarize these distributions:
  • Gini coefficients, Percentile Ratios, Earnings Share Ratios

• Each of the above measures places a particular weight/importance on different parts of the earnings/total earnings distribution.

• In addition to single number measures, we plan to show the earnings and total earnings distributions for each MSA, allowing the user to better understand why inequality is changing.
Total Earnings: San Francisco 1998 and 2017
Gini 1998 – 0.512    Gini 2017 – 0.525

Bin Height$_b$ $\propto$ sum of workers$_b$ * average earnings$_b$
Mobility, Opportunity, and Volatility (MOVS)

- Goal: integrated, regular release of household and individual income, income growth, and income persistence statistics
- Leverage administrative record and Census Bureau collected data
- Focus on three core concepts:
  - mobility in terms of individual’s simple change in position over time
  - changes in the concentration of income
  - average direction of position change for demographic groups within the distribution
- A suite of statistics on income mobility, income volatility, and related topics
  - Income growth curves, rank-rank profiles, transition matrices
  - Concentration of affluence or poverty
MOVES First Phase: Longitudinal Market Income

• Define the working-age population for 2005 using combined demographic data, a variety of administrative records, and Census Bureau collected data

• Link year-to-year IRS 1040 and W2 income and earnings
  • Develop annual household unit identifiers and equivalence scales, calculate equivalized income for individuals in age range

• Intended product is a public-use data tool of statistics by demographic group and geography
Intragenerational Mobility: Moving Forward

• Goal is to provide high value mobility, opportunity, volatility, and inequality statistics
• Plan to develop and disseminate a suite of experimental statistics for households, workers, and individuals for varying levels of geography and subgroups
• Aiming to release an initial set of measures in Fiscal Year 2023
Income Distributions
Research to Better Understand Income Distributions

• **Demographic Measures** (example research below)
  
  
  
  

• **Migration** (example research below)
  
  
  
Research to Better Understand Income Distributions

• **Firm Inequality and Labor Market Fluidity** (example research below)

• **Global Income Dynamics** (example research below)
  - McKinney, Kevin, John M Abowd, and Hubert P. Janicki. 2022. “U.S. Long-Term Earnings Outcomes by Sex, Race, Ethnicity, and Place of Birth”
  - [https://mebdi.org/global-repository-income-dynamics](https://mebdi.org/global-repository-income-dynamics)
Research to Better Understand Income Distributions

• **Item Non-Response**  (example research below)
  
  
  
  

• **Unit Non-Response**  (example research below)
  
  
  
Thank you!