Distribution of Personal Income

Marina Gindelsky

March 24, 2022

Objective

- Distribution of National Accounts
  - Distribute NIPA (macro) totals to households (micro)
  - Link income inequality to growth

"As income inequality has widened, and has attracted more attention and more anger, the demands for distributionally disaggregated national accounts have become more urgent…" (Deaton 2020)

Personal Income

- BEA distributes personal income
- Personal income is the income received by persons from
  - participation in production
  - government and business transfers
  - service flows from homeownership
  - holding interest-bearing securities and corporate stock
- Closest to the measure of economic resources available to households to purchase goods and services
- Requires least amount of assumptions to distribute national totals

BEA Release

- Distribution of Personal Income (PI) and Disposable Personal Income (DPI)
  - March 2020 release: 2007-2016
    - Levels of PI, real growth, and distributional statistics for households
  - December 2020 release: 2007-2018
    - Updated methodology and documentation based on feedback received from experts and data users. All items from previous release and DPI
  - December 2021 release: 2000-2019
    - Updated methodology and documentation based on feedback received from experts and data users. Additional years (2000-2006)
- Inequality metrics
  - Income deciles of PI and DPI overall and their components
  - Top 1%, Top 5%, Bottom 5%
  - Gini, 90/10 ratio, 80/20 ratio

BEA Methodology Summary

- Construct a distribution of personal income from publicly available data to provide insight to the income distribution
- Strategy (see technical document and working paper for more detail)
  1. Identify a NIPA total to be distributed (60 components of PI)
  2. Identify CPS variable (s) (+ outside data) to allocate component
  3. Sum all household components to subtotals of interest, PI, and DPI.
     - Adjusted Money Income (AMI) (e.g., wages and business income allocated using CPS var with SOI tail adj.)
     - Financial (F) (e.g., imputed interest allocated using information from SCF)
     - Health (H) (e.g., Medicare using CPS var; Medicaid using machine learning algorithm)
     - Other Transfers (net) (T) (e.g., WIC using CPS; SNAP using machine learning algorithm)

     Household Income = AMI + F + H + T

     Personal Income = Household Income - *Household Current Transfer Receipts from Nonprofits - Nonprofit Institution Transfer Receipts from Households + Nonprofit Institution Income*

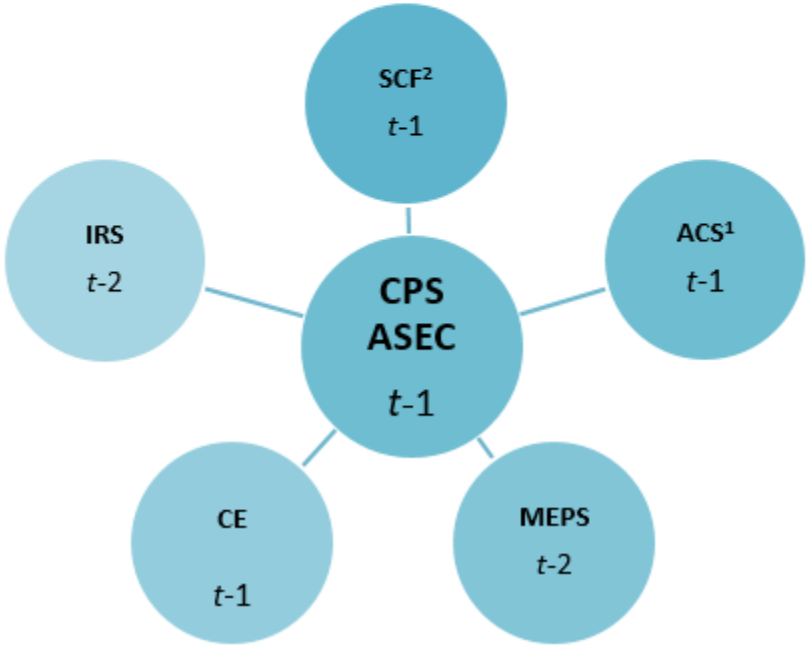
  4. Equivalize (divide by $\sqrt{household\ size}$) and rank households to compare households of different sizes to each other

Source Data and Flow Timeline

To produce an income distribution, we have the following data sources available:

CPS ASEC: Current Population Survey, Annual Supplement conducted every March (Census Bureau/BLS)
IRS SOI: Public use distributional summary tables by the Statistics of Income program (IRS)
SCF: Survey of Consumer Finances conducted every 3 years (Federal Reserve)
ACS: American Community Survey conducted every year (Census Bureau)
CE: Consumer Expenditure Survey conducted every year (BLS)
MEPS: Medical Expenditure Panel Survey, household sample conducted in rounds annually (AHRQ)

The graphic on the left shows how each source feeds into the CPS ASEC, while the table on the right summarizes the available data for each dataset in December of year *t*.



	Availability	Reference Period
CPS ASEC	Fall *t*	*t-1*
IRS	Fall *t*	*t-2*
ACS[1]	Fall *t*	*t-1*
MEPS	Fall *t*	*t-2*
SCF[2]	Fall *t*	*t-1*
CE	Fall *t*	*t-1*

[1]Release rescheduled due to delays with COVID data collection.
[2]Conducted every 3 years.

Calculation Example: Annual

- For income items that use only CPS data such as unemployment insurance
  1. Sum unemployment for all households in the CPS, weighted by their pop weight
  2. Calculate each household's share of the total
  3. Multiply each share by the NIPA total for unemployment

Example

Household	Original CPS value	Population Weight	Share of Total = Original CPS value/CPS total	Imputed Value = Share * NIPA Total
1	100	20	100/6500 = 0.0154	0.0154* 10,000 = 154
2	0	10	0/6500 = 0	0* 10,000 = 0
3	500	5	500/6500 = 0.0769	0.0769 * 10,000 = 769
4	400	5	400/6500 = 0.0615	0.0615 * 10,000 = 615

CPS Total Unemployment = 100*20 + 0*10 + 500*5 + 400*5 = 6,500

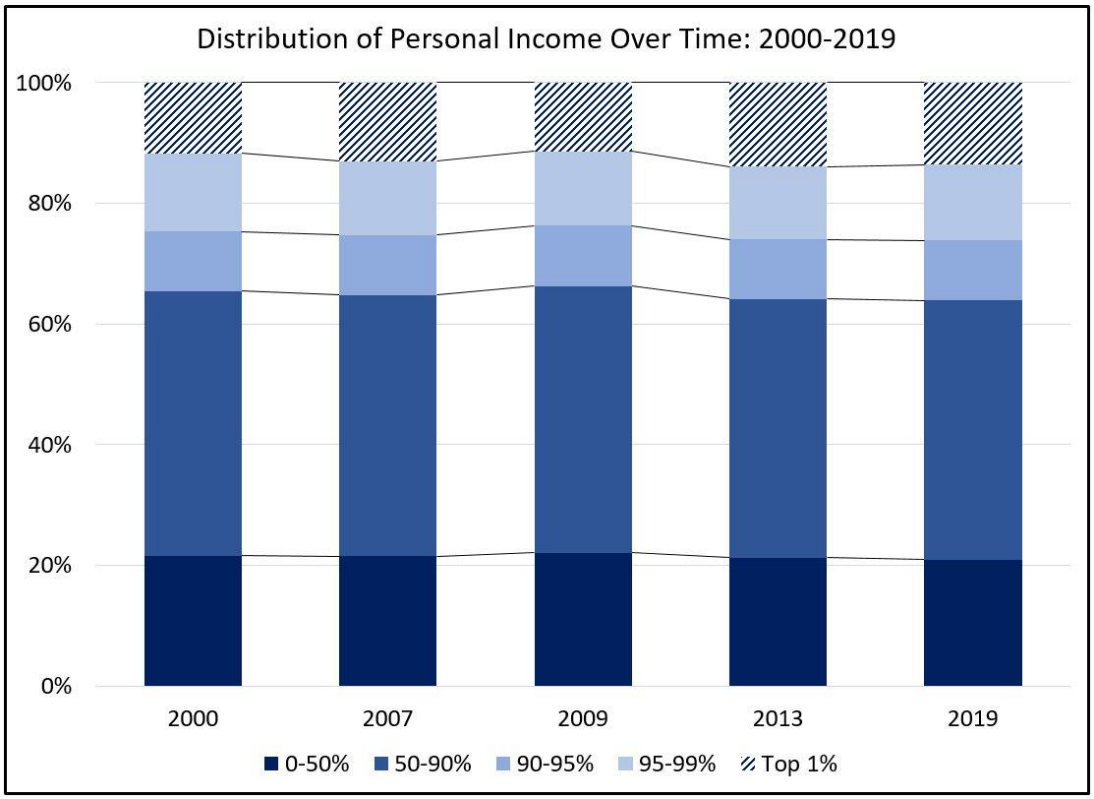
Suppose NIPA Annual Total = 10,000

When imputed values are multiplied by population weight, they sum to the NIPA total, the same way as the original CPS values sum to the CPS total, when multiplied by the weight: 154*20+0*10+769*5+615*5 =10,000

- For income items that use additional data sources, first adjustment original CPS value, then proceed the same way

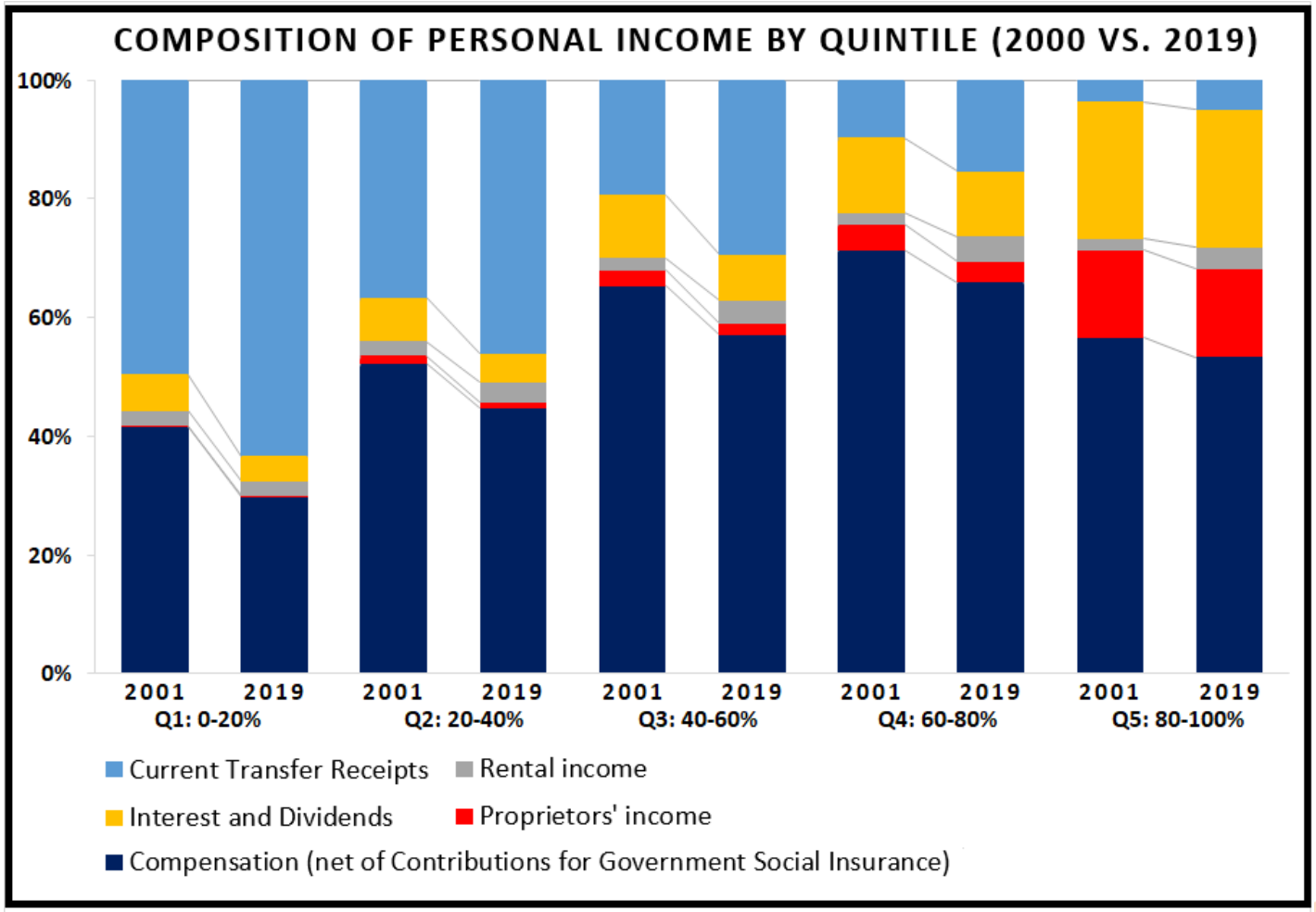
BEA: Results

- Total PI and DPI grew 26% from 2000-2019
  - Equivalized median DPI grew (25.9%) vs. median PI (24.0%)
  - Top 1% share of PI (DPI) increased 11.7%→13.6% (9.9%→11.8%)
- Growth was unequal throughout distribution
  - 53% of growth in PI and DPI went to top 20%* (cannot follow individuals over time, but group is relatively sticky in this time period)
  - Share of top quintile of PI (DPI) went up 1.3pp (2pp)
  - Share of bottom quintile went down 0.4pp (0.7pp)



Composition of PI

- Compensation has fallen as a share of income for all quintiles, while transfers have risen
- Income for the top quintile is driven much more by assets and proprietors' income

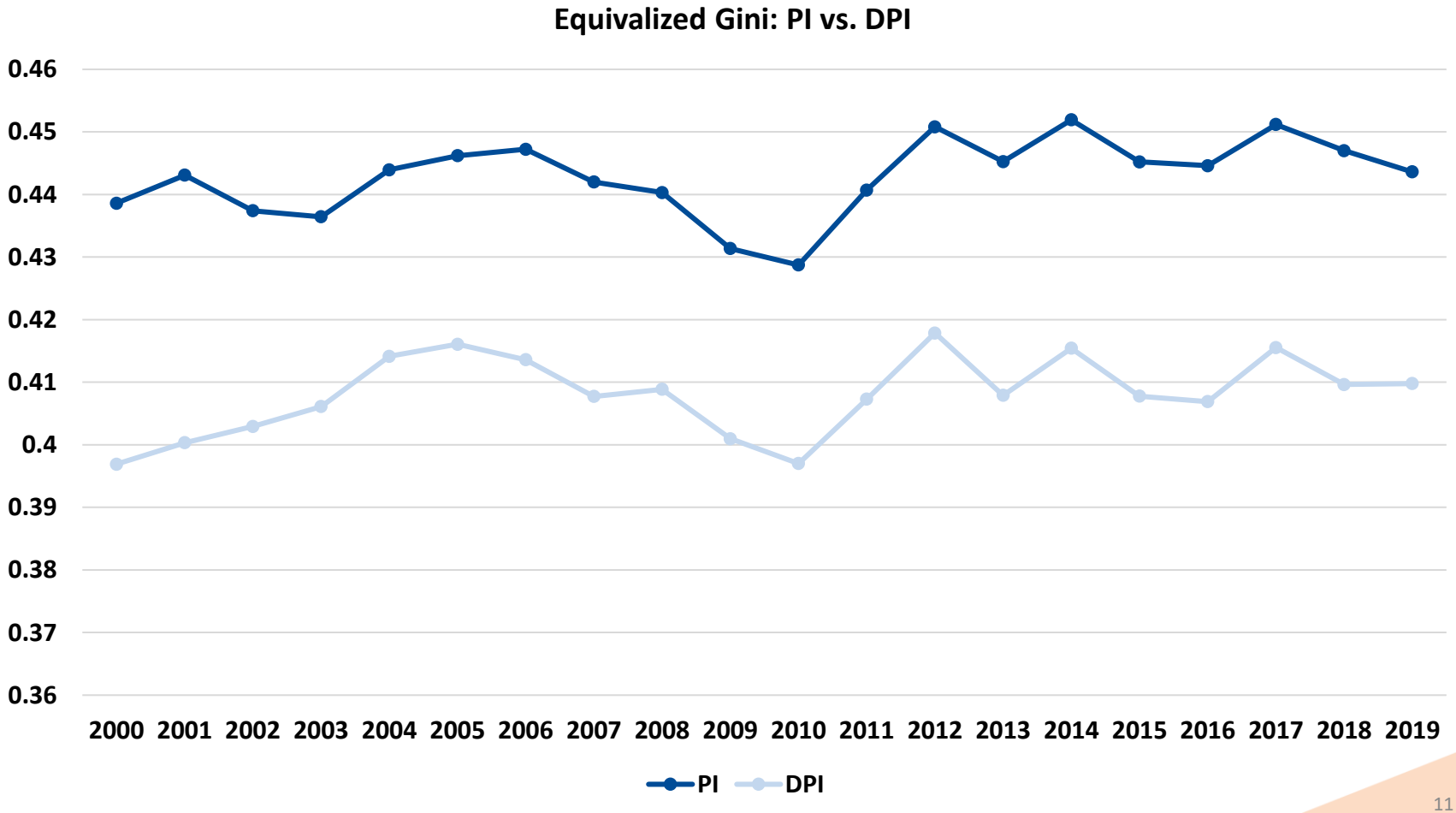


Inequality Trends

- Overall inequality of PI (Eq. Gini) is largely unchanged since 2000, despite some volatility during Great Recession
  - Share of top 1% (and top 5%) has risen, but share of bottom has not
  - Share of top quintile has risen a little bit but not very much – driven by top 5%
- Inequality of DPI shows the same patterns
- Slightly fall in inequality from 2018-2019 driven by fall in top shares
- A lower increase in inequality than Census (but the survey change effects have been mitigated by the SOI adjustment)
- Similar trend in top 1% and top 5% as PSZ for both PI & DPI

Equivalized Gini of PI & DPI

- Very similar trend for PI & DPI over time
- Falling inequality during Great Recession as transfers increased and returns to assets decreased; fairly stable afterwards



Additional Prominent Inequality Estimates for U.S.

Estimates	Measure	Frequency	Availability
BEA	Personal Income and Disposable Personal Income	Annual	2000-2019
Auten and Splinter (2019)	National Income (Pre-tax and post-tax)	Annual	1960-2015
Piketty, Saez, and Zucman DINA tables	National Income (Pre-tax and post-tax)	Annual	1913-2019
CBO	Household Income Pre-and-post-tax	Annual	1979-2018
Census Bureau	Household Money Income	Annual	1967-2019
World Inequality Database	National Income and Wealth Quantiles	Annual	1913-2019
Federal Reserve Distributional Financial Accounts	Household Wealth Quantiles	Quarterly	1989Q3-2021Q3 (2019SCF is latest available microdata, so results extrapolated post 2018)

Challenges

- Data challenges (and our approach)
  - No access to micro tax data: use IRS (SOI) public tables to approximate top tail
  - Availability of source data – time lag: exploring extrapolation and machine learning techniques to reduce this
  - No corresponding microdata for some macro concepts (e.g., imputed interest): match on income and observables to impute
- Measurement challenges (and our approach)
  - Designing allocation algorithm for each source: best possible approximation with help from experts
  - Extent of underreporting in CPS differs by income source and income group – intensive and extensive margins: exploring extrapolation and machine learning techniques to impute values
  - Consistency of variable definitions over time: do our best to indicate breaks and explain their impacts

- Solicit feedback and explore new data sources

Lessons Learned

- No one-size-fits-all solution
  - Every imputation strategy has its advantages and disadvantages
  - No single dataset captures the whole distribution fully
- Results can be sensitive to imputations for major components of PI, such as earnings
  - Should devote most resources to accurately measuring labor income, including imputing zeros
  - Health care spending is large and important, but less sensitive
- A suite of inequality statistics is best
  - Focusing on one metric is less useful to data users
  - Data users are very interested in the bottom of the distribution
- Would be very useful to tie together concepts of income, consumption, and wealth
  - Fed is attempting to tie PI to Distributional Financial Accounts
  - Collaboration with Bureau of Labor Statistics to distribute Personal Consumption Expenditures (PCE)