Bureau of Economic Analysis
U.S. DEPARTMENT OF COMMERCE

Technical Document: A Methodology for Distributing Personal Income

Marina Gindelsky
DistributionofPersonalIncome@bea.gov

The distribution of personal income is a topic of widespread interest, including for policymakers, economists, and statisticians. Though literature on the topic has dramatically increased over the past two decades, led particularly by Piketty and Saez (2003), income inequality has long been a topic of interest (Kuznets 1941, 1953). Indeed, the predecessor of the Bureau of Economic Analysis (BEA), the Office of Business Economics, published estimates of the "Size Distribution of Income" covering select years between the mid-1940s and the 1970s (Office of Business Economics 1953).

Recently, interest has particularly grown in the distributional national accounts (Piketty, Saez, and Zucman 2018). As each household member has some claim on the economic resources of the household and affects decisions regarding economic activities, the household is regarded by the Census Bureau and in the *System of National Accounts 2008* as the institutional unit for compiling distributional results (Organisation for Economic Co-operation and Development 2018). When considering households as a unit of measurement and national accounts as the aggregate, personal income represents a natural and appropriate income concept for decision-making (Fixler, Gindelsky, and Johnson 2019). Personal income is the income received by, or on behalf of, all persons from all sources—from participation as laborers in production, from owning a home or business, from the ownership of financial assets, and from government and business in the form of transfers. It includes income from domestic sources as well as from the rest of the world; it does not include realized or unrealized capital gains or losses.

There are many metrics that can be used to measure the distribution of personal income and quantify income inequality. They are summarized in data files on the BEA website for each year from 2007 to 2016 in tables 1-3. Table 1, "Major Components of Personal Income by Decile," distributes the household income portion of National Income and Product Accounts (NIPA) table 2.9, "Personal Income and Its Disposition by Households

and by Nonprofit Institutions Serving Households" by decile. Table 1 also includes the totals for each line item in the first column as well as a summary line for personal income.[1] By providing shares of income by decile, we are able to glean additional insights from the income distribution as compared to the commonly provided quintiles (Census Bureau (tables A3 and A4)). Table 2, "Decomposition of Personal Income for Households," shows the decomposition of personal income as compared with Census money income (discussed below in section 3). Table 3, "Inequality Metrics," reports inequality statistics for equivalized personal income and equivalized household income (described below in section 6). Additionally, charts 1 and 2 in the paper "Measuring Inequality in the National Accounts" show the levels of personal income and growth in this measure from 2007 to 2016.

This is a broad set of statistics that allows for a complete picture of the distribution of income, including the share in each quintile, the top 1 percent share, the top 5 percent share, the Gini index, and the ratio of the 90th percentile to the 10th percentile. These statistics are featured in the literature (including Piketty, Saez, and Zucman 2018; Atkinson, Piketty, and Saez 2011; Auten and Splinter 2019) as well as the World Inequality Database, Federal Reserve Board accounts (Distributional Financial Accounts), and Census Bureau (aforementioned tables A3 and A4) and Bureau of Labor Statistics (BLS) publications (Cunningham 2015).

Below is a description of the methodology used to estimate the distribution of personal income, as found in the NIPA tables, specifically table 2.1 ("Personal Income and Its Disposition"), line 1. Section 1 describes the overall strategy, including how each external dataset is used to calculate the statistics. Sections 2 through 6 outline the major components of personal income as found in table 2 on the BEA website. Section 7 summarizes the data sources.

1. Overall Strategy

The Annual Social and Economic Supplement of the Current Population Survey (a base microdata set, hereafter referred to as "CPS") was used for the analysis. There are many national accounts components that must be allocated to households (that is, households are given a value for each component). These components can be categorized in four broad categories, as found in table 2, "Decomposition of Personal Income for Households," on the BEA website: (1) adjusted money income (AMI), (2) financial (F), (3) health (H), and (4) other transfers (net) (T). Some of these components were allocated to households using additional data that does not come directly from the CPS in conjunction with CPS variables. In order to maximize transparency and utility to data users, all data used was public-use data and may be accessed using the links provided in section 7, "Data Sources."

The CPS, jointly sponsored by the Census Bureau and BLS, is a nationally representative annual household survey of the civilian noninstitutionalized population (approximately 98,000 households sampled, representing

1. NIPA totals for personal income can be found in BEA tables 2.1 and 2.9. Table 2.9 disaggregates personal income into household income and income from nonprofit institutions serving households. By utilizing table 2.9, we were able to distribute the aggregate totals for households to the households they represent.

126 million households in 2016) that includes many demographic and income questions and is mainly administered in March of each year. According to the Census Bureau, the survey is the primary source of labor force statistics for the United States.[2] It is important to note the income data in the CPS is 1 year behind the data collection year (that is, the survey year). Our analysis starts in 2007 due to the availability of our component data sources, particularly Medicare data (see section 7).

Overall, the strategy used in this exercise was the following four-step process: (1) identify a NIPA total to be distributed, (2) identify CPS variable(s) and external variables that can be used to allocate this total, (3) sum all component NIPA totals to personal income and subtotals of interest, and (4) construct inequality statistics. After all components have been added together to compute personal income, equivalized personal income is calculated by dividing personal income by the square root of the number of household members. For example, if household income is $10,000 and there are four members of the household, equivalized household income is $5,000 (half of $10,000). By equivalizing, we are able to arrive at comparable figures for all households (that is, it is an adjustment for household size). Equivalized personal income is used for all income inequality metrics.

2. Allocations Based on External Datasets

There are five datasets used in addition to the CPS to allocate NIPA totals to households.

A. Statistics of Income: adjustment for very high incomes

Data from the Internal Revenue Service (IRS) Statistics of Income (SOI) program is used to adjust the top (that is, highest incomes) of the income distribution (or "tail") in order to more accurately reflect true inequality, which is thought to be underestimated in the CPS. There are three main reasons we adjust CPS data. First, the CPS is believed to unsuccessfully survey those with very high incomes, which introduces nonresponse bias in inequality estimates (Bollinger et al. 2018). Second, there is a perception there is underreporting by those with top incomes. Third, the CPS has top codes, which vary by year, for those with top incomes so as not to risk identification of those individuals. For example, if an individual reports $10 million annually, he/she may be given a value of $1 million. For these reasons, it becomes prudent to adjust CPS incomes (Armour et al. 2016). The adjustment process is described below and is different from previous exercises (Fixler et al. 2017; Fixler, Gindelsky, and Johnson 2018, 2019).

We aggregate the SOI data on tax units (see section 7) into the proportion of income with adjusted gross income (AGI) less than $500,000 versus AGI greater than or equal to $500,000 for six components: (1) wages, (2) business income, (3) ordinary dividends, (4) taxable and nontaxable interest income, (5) farm income, and (6) rents and royalties. These shares sum to 1 for each income category in each tax year. For example, a 2018

2. It is conveniently processed for ease of use by the National Bureau of Economic Research and can be downloaded from its website (see section 7).

CPS that asks respondents about 2017 income would correspond to tax year 2017 SOI data. For each of the six components listed above (that is, *Source X*), the procedure is as follows:

1. Group CPS households by money income (comparable to AGI) into two groups:
   - Incomes greater than or equal to $500,000
   - Incomes less than $500,000
2. Sum *Source X* (weighted) in the CPS and subtract from the NIPA total.
3. A portion (the percent SOI share for *Source X*) of this aggregate difference is then allocated to households in their respective income group proportional to their share of *Source X* in the CPS total.
4. The original CPS value is added to this "extra" imputed value by household such that when aggregated across all households for *Source X*, the total will add up to the NIPA total.

For example, in 2012, 53 percent of AGI for ordinary dividends in the SOI data is at least $500,000. The NIPA total for dividend income is $808 billion, while total weighted CPS dividend income is $123 billion. That leaves $808 billion–$123 billion = $685 billion to be allocated to CPS households as follows: 53 percent of $685 billion = $363 billion to households with incomes that are at least $500,000 and 47 percent of $685 billion = $322 billion to households with incomes that are less than $500,000. Each household then receives extra dividend income proportional to its share of dividend income in its group such that aggregate weighted household dividend income (original income plus extra income) will sum to $808 billion.

B. Congressional Budget Office: adjustment for underreporting of certain government programs

The Congressional Budget Office (CBO) (see section 7) provides a CPS crosswalk that uses an algorithm to assign probabilities of receipt of Supplemental Security Income (SSI), Medicaid, and Supplemental Nutrition Assistance Program (SNAP) benefits to each individual in the CPS for a given survey year and then correspondingly imputes values. We use these imputed values rather than reported CPS values for these variables in order to correct for underreporting in these important categories as articulated in the CBO methodology paper.

C. Survey of Consumer Finances: imputation of financial income

We use the Summary Extract Public Dataset of the Survey of Consumer Finances (SCF), as obtained from the Federal Reserve Board of Governors Economic Research website. We use three asset variables from this dataset to distribute the three imputed components of personal income, which will be described for each relevant component. Because the SCF is triennial, we first interpolate the SCF variables for the years in which the SCF is not observed.[3] Next, we place households into before-tax bins by income level. The share of the total asset variable

3. We use the Fernandez procedure in the R package "tempdisagg." The Fernandez procedure extends the Denton and Chow-Lin approaches by obtaining its solution by minimizing a quadratic loss function in the differences between the series to be created and a linear combination of the high-frequency series. The results of this method very closely match the results of the interpolation used in the Federal Reserve Board Distributional Financial Accounts.

held by all households in the given bin is calculated. The CPS households are placed into the same income bins by adjusted money income. Finally, the NIPA totals are then allocated by the distribution of each of these asset variables (for example, share of asset variable by income bin) to the CPS households in the respective bins.

D. Consumer Expenditure Survey: imputation of rental income for owner-occupied housing

We use the Consumer Expenditure Survey (CE) from BLS to impute rental income for owner-occupied housing. Using this data source, we first rank "consumer units" (roughly the same as households) by before-tax income, creating deciles. We next construct a share of rental equivalence to before-tax family income. For example, if a household's income is $100,000 and it reports the expected rental value of its home is $4,000 monthly ($48,000 annual), the rent-to-income share would be 48 percent. The median share is calculated for each income decile. This share is then applied to income deciles in the CPS for households that own their home to impute a value of rental income for owner-occupied housing based on household income.

E. Centers for Medicare & Medicaid Services: Medicare

We use average annual expenditure on Medicare data (by state) from the Centers for Medicare & Medicaid Services (CMS) in order to allocate Medicare expenditure to households. This allows us to capture the significant variation in Medicare expenditure by state when distributing the national totals.

As noted above, the following sections will describe the income components as enumerated in table 2. We start with adjusted money income.

3. Adjusted Money Income (table 2, line 2)

In these estimates, we adjust money income (as defined by Census) in order to be consistent with the concepts used in the NIPA estimates. It is helpful to compare the definitions of Census money income and personal income:

"Census money income is defined as income received on a regular basis (exclusive of certain money receipts such as capital gains) before payments for personal income taxes, social security, union dues, Medicare deductions, etc. Therefore, money income does not reflect the fact that some families receive part of their income in the form of noncash benefits, such as food stamps, health benefits, subsidized housing, and goods produced and consumed on the farm." (From "Income & Poverty" on the Census Website)

"Personal income is the income that persons receive in return for their provision of labor, land, and capital used in current production, plus current transfer receipts less contributions for government social insurance (domestic). Personal income is equal to national income minus corporate profits with IVA and CCAdj, taxes on production and imports less subsidies, contributions for government social insurance, net interest and miscellaneous

payments on assets, business current transfer payments (net), and current surplus of government enterprises, plus personal income receipts on assets and personal current transfer receipts." (From Chapter 2 of Concepts and Methods of the U.S. National Income and Product Accounts)

Though Census money income in many ways is a narrower definition of income, it does include variables that are not in personal income, such as retirement disbursements. Accordingly, we add up the components of Census money income that are in personal income, excluding variables such as retirement disbursements (*ret-val*) or certain sources of disability income (*dis-val*) and survivor income (*sur-val*).[4] We call this approximation adjusted money income.[5] It is primarily comprised of income from wages and salaries, self-employment (farm and nonfarm), interest, dividends, and social security income, which together sum to 95 percent of adjusted money income in 2016. The remaining 5 percent is comprised of income from additional sources such as rents and royalties, unemployment insurance, and disability income, among others.

We distribute the components that make up AMI individually. Wages and salaries, farm income, nonfarm income, rental income of persons (other private business), interest income, and dividend income are distributed by the relevant variables, with the SOI adjustment referred to in section 2A. Federal benefits including social security, unemployment insurance, railroad retirement, black lung benefits, pension benefit guaranty, veterans benefits, and workers' compensation are distributed by the relevant variables in the CPS. State and local benefits including temporary disability insurance, assistance, employment and training, education, and others are distributed to CPS households by their relevant variables, respectively. SSI is distributed using the CBO crosswalk referred to in section 2B. Household current transfer receipts from nonprofit institutions are distributed to CPS households by the value of their educational assistance in the CPS.

4. Financial Items (table 2, line3)

The objective of this group is to impute financial items in the NIPAs to CPS households. This category is the sum of allocations for pensions and profit sharing, life insurance, rental income from owner-occupied housing, and imputed interest.

Employer contributions to pension plans and group life insurance are distributed by wage. Imputed interest is distributed using an imputation derived from the SCF for banking, insurance, and employee pension plans, using the Chow-Lin method for nonsurvey years, for each relevant item, as referred to in section 2C. Owner-occupied housing rental income is distributed using an imputation derived from the Consumer Expenditure Survey by income bracket referred to in section 2D.

---

4. The exclusion of retirement disbursements constitutes approximately 75 percent of the money income excluded.

5. Adjusted money income also excludes potential sources of intrasectoral transfers, which would net out in the sector and are not associated with current period production, such as other financial assistance (*fin-val*), other income (*oi-val*), alimony (*alm-val*), child support (*csp-val*), and other noninstitutional educational assistance (*ed-val*). We do include incomes from these sources tied to railroad retirement, other retirement, workers' compensation, black lung benefits, and state and local government disability.

5. Health Items (table 2, line 4)

The objective of this group is to impute health items in the NIPAs to CPS households. This category is the sum of allocations for employer contributions for health insurance, Medicare, Medicaid, military medical insurance, and other medical care payment assistance.

Employer contributions to health insurance are distributed by the corresponding CPS variable, except for military medical insurance, which is distributed to active military members. Medicare is distributed by assigning the state average expenditure to CPS individuals who report receiving Medicare. Medicaid is distributed using the CBO crosswalk referred to in section 2B. Medical assistance is distributed to those who report receiving SNAP and Women, Infants, and Children (WIC) benefits or other welfare assistance, except the Children's Health Insurance Program (CHIP), which provides assistance to households reporting children covered by CHIP.

6. Net Other Transfers (table 2, line 5)

The objective of this group is to impute items related to transfers in the NIPAs that are not contained in adjusted money income or health to CPS households. These items include employer and employee contributions for government social insurance (net): SNAP, WIC, refundable tax credits, energy assistance, educational assistance, and other transfers.

Employer/employee contributions to old-age, survivors, disability, and hospital insurance are distributed by the imputed value of FICA (Federal Insurance Contributions Act) in the CPS. Employer contributions to workers' compensation and supplemental unemployment are distributed respectively by wage. Other employer and employee/self-employed contributions to government social insurance are also distributed by wage. Military medical insurance (federal benefits and employee/self-employed contributions to government social insurance) are distributed to CPS households with active military members. SNAP is distributed using the CBO crosswalk referenced in section 1B. Refundable tax credits are distributed to CPS households by the total value of reported earned income tax credit, child tax credit, and additional child tax credit benefits. Energy assistance and WIC are distributed to CPS households by the relevant indicators. Household current transfer receipts from business (net) are distributed equally to all households. All other transfers are distributed to households reporting receipt of WIC, SNAP, or other welfare assistance.

Finally, to calculate personal income from household income, *household current transfer receipts from nonprofits* and *nonprofit institution transfer receipts from households* are deducted, and *nonprofit institution income* is added. This residual is distributed equally to all individuals in the CPS.

7. Data Sources

Description of the data. There were many data sources (listed below with links) used for the compilation of these estimates. All are publicly available and made available with varying time lags after the data is collected. Information on incomes from the CPS and Consumer Expenditure Survey is regularly available with a 1-year time lag (for example, 2016 values became available in 2017) since the surveys ask respondents about income received in the previous calendar year. Key data from the IRS SOI program, which is used to adjust top incomes, is regularly available with a 2-year lag (for example, 2016 values became available in 2018). The SCF is a triennial survey that is conducted every 3 years and asks about income received in the previous year, similar to the CPS and the Consumer Expenditure Survey. The CBO constructs a crosswalk from CPS values for certain transfer programs in order to correct for underreporting in these categories. There is no regular publication schedule for these estimates, and the most recent year available is 2016. Finally, Centers for Medicare & Medicaid Services (CMS) data on Medicare expenditure by state is currently available with a 2-year time lag.

Microdata

1. **Current Population Survey, March Supplement.** This is the core data series that is the base of the analysis. We used all households (except those residing in group quarters) from the survey years 2008 to 2017. This data was downloaded from the National Bureau of Economic Research website.

2. **Consumer Expenditure Survey.** We used this survey in order to estimate rental equivalence for owner-occupied housing from 2008 to 2017. The data was downloaded from BLS.

3. **Survey of Consumer Finances.** We used this survey in order to distribute imputed interest income from the NIPAs for 2007, 2010, 2013, and 2016. For the nonsurvey years, we imputed values with the methodology described in section 2C. The data was downloaded from the Federal Reserve Board.

4. **Centers for Medicare & Medicaid Services.** We used data from 2007 to 2016 on the average annual expenditure for Medicare by state found on the CMS website.

5. **Congressional Budget Office.** We used the CPS-crosswalk values of Medicaid, SNAP, and SSI for each year to correct for underreporting in these categories. These crosswalks can be found on the CBO website.

6. **Internal Revenue Service Statistics of Income data.** In order to adjust top incomes in the CPS, we used summary IRS Form 1040 data aggregated by AGI for tax years 2007 to 2016 in table 1.4 found in "SOI Tax Stats" on the IRS website.

Macrodata

1. **National Income and Product Accounts.** The primary source for our national macrodata for 2007 to 2016 is the BEA NIPA data on the BEA website. Please note this data is subject to frequent revision. For the current set of estimates, we used the October 2019 release.

2. **Census Bureau.** We compare our estimates with official estimates for Census money income produced by the Census Bureau in our results. These estimates can be found in tables A-2, A-4, and A-5 on the Census website. Since these numbers have been deflated using CPI-U-RS data from BLS, they are reinflated using the personal consumption expenditures tables produced by BEA.

References

Armour, Philip, Gerald Auten, Richard Burkhauser, and Jeff Larrimore. 2017. "Recent Trends in U.S. Top Income Shares in Tax Record Data Using More Comprehensive Measures of Income Including Accrued Capital Gains." Working paper no. 23007. Cambridge, MA: National Bureau of Economic Research, June.

Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez. 2011. "Top Incomes in the Long Run of History." *Journal of Economic Literature* 49 (1): 3–71.

Auten, Gerald, and David Splinter. 2019. "Top 1 Percent Income Shares: Comparing Estimates Using Tax Data." *AEA Papers and Proceedings* 109 (1): 307–311.

Bollinger, Christopher R., Barry T. Hirsch, Charles M. Hokayem, and James P. Ziliak. 2019. "Trouble in the Tails? What We Know about Earnings Nonresponse 30 Years after Lillard, Smith, and Welch." *Journal of Political Economy* 127 (5).

Cunningham, J. Chris. 2015. "Measuring wage inequality within and across U.S. metropolitan areas, 2003–13." *Monthly Labor Review* 138.

Fixler, Dennis, David Johnson, Kevin Furlong, and Andrew Craig. 2017. "A Consistent Data Series to Evaluate Growth and Inequality in the National Accounts." *Review of Income and Wealth* 63 (S2): S437–S459.

Fixler, Dennis, Marina Gindelsky, and David Johnson. 2019. "Improving the Measure of the Distribution of Personal Income." *AEA Papers and Proceedings* 109: 302–306.

Fixler, Dennis, Marina Gindelsky, and David Johnson. 2018. "Towards a Distribution of Household Income: Linking Survey Data to Administrative Data." Paper presented at the International Association for Research in Income and Wealth 35th General Conference, Copenhagen, Denmark, August 20–25.

Kuznets, Simon. 1941. *National Income and Its Composition*, 1919–1938, Volume I. New York: National Bureau of Economic Research.

Kuznets, Simon. 1953. *Shares of Upper Income Groups in Income and Savings*. New York: National Bureau of Economic Research.

Organisation for Economic Co-operation and Development. 2018. *Handbook on compiling distributional results on household income, consumption and saving consistent with national accounts*.

Piketty, Thomas, and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913–1998." *Quarterly Journal of Economics* 118 (1).

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman. 2018. "Distributional National Accounts: Methods and Estimates for the United States." *Quarterly Journal of Economics* 13 (2).

U.S. Office of Business Economics. 1953. *Income Distribution in the United States, by Size*, 1944–1950. Washington, DC: U.S. Government Printing Office.