

Technical Document: A Methodology for Distributing Personal Income

Marina Gindelsky

[DistributionofPersonalIncome@bea.gov](https://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 U.S. 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. 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 U.S. Census Bureau (Census) 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. A full set of metrics for 2022 is available on the [BEA website](#) under “National Data & Documentation” in an excel file (see sheets “Table 1” and “Table 2”), with a partial set of provisional metrics for 2023 in a separate file. Table 1, “Major Components of Personal Income and Disposable Personal Income by Decile,” distributes the household income portion of National Income and Product Accounts (NIPAs) table 2.9, “Personal Income and Its Disposition by Households and by Nonprofit Institutions Serving Households,” by decile in two ways. First, households are ranked by equivalized (household size-adjusted) personal income and assigned to deciles (the first 10 numerical columns). Next, households are re-ranked by equivalized disposable personal income (personal income less taxes) and assigned to deciles. Table 1 also includes the totals for each line item in the first column and summary lines for personal income and disposable personal income.¹ The deciles of personal income and disposable personal income are published as iTable 2.10.

By providing shares of income by both personal income decile and disposable personal income decile, we are able to glean additional distributional insights, both as compared to the commonly provided quintiles (Census tables A3 and A4) and relative to solely pretax income. Table 2, “Inequality Metrics,” reports inequality statistics for household income, personal income, and disposable personal income.

A summary file containing all available metrics for all available years has been uploaded for ease of comparison of concepts and quantiles across the time series. 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.¹ 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), Census (aforementioned tables A3 and A4) and U.S. Bureau of Labor Statistics (BLS) publications (Cunningham 2015).

Below is a description of the methodology used to estimate the distribution of personal

¹ An appendix table with inequality statistics calculated based on equivalized income is available upon request.

income, as found in the NIPA tables, specifically table 2.1, “Personal Income and Its Disposition” line 1. This methodology also underlies the construction of the internationally comparable OECD research estimates (see accompanying OECD technical document). Section 1 describes the overall strategy, including how external datasets are used to calculate the statistics. Sections 2 through 7 outline the major components of personal income. Section 8 explains the extrapolation which underlies the provisional results and Section 9 describes technical updates to the previous release. Section 10 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 proportional amount for each component). These components can be categorized in four broad categories under [“Decomposition of Personal Income for Households,”](#) on the BEA website. These include: (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 in conjunction with CPS variables. To maximize transparency and utility to data users, most data used for 2000–2022 were public-use data and may be accessed using the links provided in section 10, “Data Sources.”

The CPS, jointly sponsored by Census and BLS, is a nationally representative annual household survey of the civilian noninstitutionalized population (approximately 92,000 households sampled, representing 130 million households for 2020) that includes many demographic and income questions and is administered in February, March, and April of each year. According to [Census](#), the survey is the primary source of labor force statistics for the U.S.² It is important to note the income data in the CPS are 1 year behind the data collection year (that is, the survey year). Our analysis starts in 2000 due to the annual availability and comprehensiveness of some component data sources (see section 10).

² During the period of analysis here, there has been a significant redesign. We incorporated the first portion of the redesign in 2014 and the second in 2018. This redesign significantly affects the results, raising inequality.

Overall, the estimation strategy comprises the following four steps: (1) identify a NIPA total to be distributed, (2) identify CPS variable(s) and external variables that could be used to allocate this total, (3) sum all component NIPA totals to personal income and subtotals of interest, and (4) construct inequality statistics. This process yields a dataset in which each CPS household has a value for each component of personal income. After all components have been summed 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 rankings of income are used for all income inequality metrics.

2. Allocations based on external datasets

Several additional datasets are used to allocate NIPA totals to households.

A. Adjustment for very high incomes

Data from the Internal Revenue Service (IRS) Statistics of Income (SOI) program are used to adjust the top (that is, highest incomes) of the income distribution for three main reasons. First, the CPS is believed to unsuccessfully survey those with very high incomes, introducing nonresponse bias in inequality estimates (Bollinger et al. 2018). Second, there is a perception that 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 an income of \$10 million annually, they may be assigned a value of \$1 million by Census. For these reasons, it becomes prudent to adjust CPS incomes (Armour et al. 2021). The adjustment process is described below.³

First, we account for misreporting in the reported SOI data. As discussed in detail by

³ Previous exercises had different adjustment strategies. See: Fixler et al. 2017; Fixler, Gindelsky, and Johnson 2018, 2019, 2020; Gindelsky 2020, 2023.

DeBacker et al. (2020), Johns and Slemrod (2010), Auten and Splinter (2019), and others, audit data from the IRS National Research Program and Taxpayer Compliance Measurement Program have shown that administrative data are significantly underreported (i.e., the existence of a “tax gap”) due to tax noncompliance. BEA does make an adjustment for aggregate misreporting of proprietors’ income and partnership income in the macrodata, seen in NIPA table 7.14, line 2. However, the extent of misreporting differs substantially by income source (for example, self-employment vs. wage income) and varies significantly across the distribution.⁴ Thus, we first multiply the SOI data for wages, self-employment, dividends, interest and rents and royalties by each adjustment factor to be consistent with the audited totals.⁵

Next, we aggregate the SOI data on tax units (for 2000–2022, see section 7; for 2023, we use internal tax data—see section 8) into the relative proportions of income with adjusted gross income (AGI) in categories as follows: for 2000–2010, less than \$200,000, and \$200,000+; for 2011 forward, less than \$200,000, \$200,000–\$500,000, and AGI greater than or equal to \$500,000.⁶ These categorizations are done for five components:⁷ (1) wages, (2) self-employment income (including farm),⁸ (3) dividends, (4) interest income, and (5) rents and royalties. These shares sum to 1 for each income

⁴ DeBacker et al. (2020) provide estimates (2006–2014) of misreporting by income category for wages, dividends, Schedule C, Schedule D, Schedule E, and “All Other” in table A4, “Percentage Changes in Income as a Result of Audit by Income Source and AGI Group.”

⁵ Although we are able to roughly adjust the SOI data to reflect audited totals for these categories, we are not able to do a detailed adjustment of the SOI data to reflect the higher “true” totals that would include underreported income that examiners were unable to detect (Johns and Slemrod 2010). IRS researchers employ an econometric technique called “detection-controlled estimation” (DCE) to model the true amounts. BEA received a report from IRS reflecting this “true” misreporting for proprietors’ income and partnership income, which is used to adjust business income in NIPA table 7.14. However, though Johns and Slemrod (2010) do provide some distributional estimates of the misreporting impact (reported versus “true”) in their tables, they do so for percentiles and for tax year 2001. Thus, we are unable to use these for our current exercise.

⁶ In previous years, we have used three categories for all years. They have now been collapsed to 2 for 2000–2010 to limit noise caused by a thin tail in the CPS (i.e., very few observations above \$500,000 in years before 2011).

⁷ We run TAXSIM35 (a National Bureau of Economic Research (NBER) program that calculates tax liability from survey data) for all the years to impute tax variables based on the scaled NIPA values for wages, self-employment, and unemployment. One of the model outputs, AGI, is very highly correlated with the CPS AGI variable for the years after 2010 due to improvements in Census modelling. We do not have information on capital gains in the CPS for 2008–2017, which is used in TAXSIM, thus our tax liability would be slightly understated for those years.

⁸ Here, by self-employment, we mean self-employment income present on IRS Schedule C, partnership income present on IRS Schedule K, and farm income from Schedule F. We have disaggregated data on S-corps and partnerships from SOI and are thus able to use the partnership data directly, in proportion to the partnership data in the NIPA total for non-farm proprietors’ income. Though it is not possible to distinguish between sole proprietorships and partnerships in the self-employment income question in the CPS, we believe survey respondents include both types of income. Farm has been combined with self-employment due to the small number of CPS observations and volatility. To be consistent with BEA macro totals, we are subtracting the net loss column from the net income column in the SOI data for the totals used in the next step for self-employment, and rents and royalties.

category in each tax year. For example, a 2022 CPS that asks respondents about 2021 income would correspond to tax year 2021 SOI data. For each of the five components listed above (simply referred to as Source X below), the procedure is as follows:

1. Create tax units⁹ in the CPS and sort by AGI into two or three groups as above:
 - AGI greater than or equal to \$500,000¹⁰ for 2011+; AGI greater than or equal to \$200,000 for 2000–2010
 - AGI of \$200,000–\$500,000 for 2011+
 - AGI less than \$200,000
2. Allocate these totals to CPS tax units (which we construct) corresponding to the above categories such that the resulting distribution resembles the SOI data, respectively for each Source X.

For example, if the SOI distribution of dividends was calculated such that tax units with AGI under \$200,000, \$200,000–\$500,000, and \$500,000+ had shares of total dividends of 50%, 30%, and 20% respectively, tax units in the CPS are allocated shares of the NIPA dividend total such that those with AGI under \$200,000, \$200,000–\$500,000, and \$500,000+ end up with (weighted) shares of total NIPA dividends of 50%, 30%, and 20% respectively. This is a methodological update from previous versions which resulted in slightly increased inequality for all years by increasing top shares (see section 9 for details).

⁹ As defined by the [Urban Institute](#), “A tax unit is an individual, or a married couple, that files a tax return or would file a tax return if their income were high enough, along with all dependents of that individual or married couple. A tax unit is therefore different than a family or a household in certain situations. For example, a cohabiting couple constitutes one household but if the individuals are not legally married, they would file separate tax returns and thus be considered two tax units. A family could consist of a married couple and the wife's elderly mother who lives with them. That family would be considered two tax units since, if the elderly mother had a large enough income, she would be required to file a federal income tax return on her own. In general, the number of tax units tends to be larger than the number of families or households reported elsewhere.”

¹⁰ Households are sorted into tax units, using information on household structure and filing status. If the proportion of tax units with AGI of at least \$500,000 is slightly too low to represent the proportion in the SOI tables, a subset of high earnings households is iteratively added to the group until this proportion is roughly equivalent to that of the SOI for the given year. We do the same with AGI of \$200,000–\$500,000. We do not do this for households with lower incomes, given that it may well be that households at the lower end do not file taxes but still report earned incomes to the CPS. Our approach thus becomes a weighted average of the CPS and SOI data. Tax units are re-aggregated into households afterward such that when aggregated across all households for *Source X*, the total will add up to the NIPA total.

B. Adjustment for underreporting of certain government programs

In the current exercise, we model the reciprocity of Supplemental Security Income (SSI), Medicaid, and Supplemental Nutrition Assistance Program (SNAP) benefits to each individual in the CPS for a given survey year. From administrative data on enrollments of these programs available from the [Social Security Administration](#), [Kaiser Family Foundation](#), and [U.S. Department of Agriculture](#), respectively, we first identify the gap in underreporting. We then use the general approach of the [Congressional Budget Office \(CBO\)](#) (Habib 2018) to determine who in the CPS is most likely to underreport program participation and use a machine-learning method to improve the identification of those individuals. More specifically, we use a Random Forest algorithm, which can model varied nonlinearities in the data. Program participation probabilities are made using out-of-sample predictions to remove biases from overfitting. Finally, we impute program income for these individuals, such that it corresponds to state-level NIPA totals.¹¹

We also adjust for underreporting of unemployment income for 2000–2021 using a crosswalk from Larrimore, Mortenson, and Splinter (2023). This adjustment imputes reciprocity to CPS respondents and assigns values, which we then scale to the NIPA total.

C. 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](#).¹² We use 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.¹³ We match

¹¹ In preliminary versions of our exercise (Gindelsky 2020), we used a crosswalk, provided directly by CBO, that uses an algorithm to assign probabilities of receipt of SSI, Medicaid, and SNAP to each individual in the CPS for a given survey year and then correspondingly imputes values. As the CBO estimates are not available on a regular schedule (and when published, with a 2-year lag), we modeled the benefits to match as closely as possible. Our resulting overall reciprocity and program income means are 99 percent correlated with those of CBO for 2000–2018. On a household level, program income reciprocity and means are correlated in the microdata at more than 70 percent for all programs. However, we have greater household-level variance in our imputations due to our process of controlling to state-level NIPA program totals, available in NIPA table SAINC35.

¹² Additionally, we use imputations of defined benefit pensions available [here](#) in an SCF crosswalk, provided by the Fed.

¹³ 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

the SCF survey year to the CPS survey year (both income variables are lagged 1 year, while asset variables are contemporaneous). First, households are binned by income (not including capital gains). These eight bins are constructed analogously in the CPS. In order to allocate imputed interest received from banks, credit agencies, and investment companies, we first calculate the total amount of “banking” (checking plus savings balance) and “insurance” (cash value of life insurance). Then, the relative share of each income bin to total income is imputed for each bin and matched to households in the CPS by corresponding income bin. Thus, the NIPA totals for imputed interest for these two categories are then allocated proportionally by income.

To allocate imputed interest¹⁴ and dividends from pensions, we first calculate the shares of retirees with defined contribution (DC) and defined benefit (DB) plans, respectively. Next, we calculate the share of DC participants who also have DB plans and the share of disbursements that comes from DC plans. Prior to the CPS survey redesign, the relevant information is (1) whether an individual is participating in a retirement plan or (2) whether an individual receives disbursements from a retirement plan. It is not reported in the CPS whether an individual is participating in a DC or DB plan explicitly. Thus, individuals reporting participation in a retirement plan are randomly allocated to DC alone, DB alone, or DB and DC, until the share of those participating in these plans matches that of the SCF. Once individuals have been allocated to DC, DB, or both, they are given a share of the imputed interest and dividends. Similarly, a portion of those reporting “Regular payments from IRA, KEOGH, or 401(k) accounts” is assigned to the group of those receiving disbursements from DC plans, consistent with the share in the SCF. It is possible to reasonably identify those receiving disbursements from DB plans in the CPS in that same question. Recipients are given a share of the imputed interest and dividends according to their reported disbursement amount.

After the CPS survey redesign (implemented in 2018), it is possible to identify those contributing to retirement plans and those receiving disbursements from DC plans. It is, however, not possible to identify those participating in both DB and DC plans. Thus,

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.

¹⁴ This imputed interest is deducted from the interest variable in the CPS when scaling monetary interest to the NIPA total.

participants in DB plans are allocated a share of imputed interest and dividends according to their wage, while participants in DC plans are allocated a share according to their reported contribution. Those receiving disbursements are allocated a proportional share.

D. 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 for 2000–2004. 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 annually), 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 (ranked by money income) in the CPS for households that own their home to impute a value of rental income for owner-occupied housing based on their income.

We use the American Community Survey (ACS) from Census in a similar way as the CE to impute rental income for owner-occupied housing for 2005–2023. The ACS is preferred given the sample size of the survey. Though the ACS does not contain a rental equivalence question, we use the BEA strategy of Rassier et al. (2021) to impute a value for each household. We use stacked 5-year Public Use Microdata Sample (PUMS) files where possible to have the largest possible sample size and 1-year files where not. ACS households are ranked on money income (as closely as possible to the CPS definition) and then sorted into before-tax deciles. Then, the share of rental equivalence is calculated as above for the CE.

E. Employer contribution to health insurance and Medicare

In the latest CPS redesign, Census no longer provides an imputation for the employer contribution to health insurance. Accordingly, we impute this variable ourselves (from 2010 forward). We use the MEPSnet Query Tool to create tables of average premium

per enrolled employee and percent of total premiums contributed by employees by family size (single or family), state, year, and firm size at private sector establishments. If CPS respondents report that the employer pays the full premium, the average employer premium is imputed by state, year, family size, and firm size for the respondent. If CPS respondents report that the employer pays part of the premium and the amount paid out of pocket, the employer contribution is then imputed using the average percent of the premium that employees pay by state, year, family size, and firm size.

Additionally, we use the Medicare 5% microdata to calculate the average individual Medicare expenditure by year, age category, and disability status. This is then imputed to each CPS individual reporting Medicare reciprocity, prior to being scaled to match the NIPA totals (by state). Although these nonpublic data are provided to BEA by the Centers for Medicare & Medicaid Services (CMS), we have compiled a state-level crosswalk, available upon request. We have switched to this data source to extend the series back to 2000 and create a more nuanced imputation. The previously used publicly available CMS average expenditure data files were unavailable before 2007.

As noted above, the following sections will describe the core aggregated income components. A numeric breakout by relative size is available for the December 2020 release in [Fixler, Gindelsky, and Johnson 2020](#).

3. Adjusted Money Income (AMI)

In these estimates, we adjust money income (as defined by Census) 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 “About [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 (*dbtn-val*) or certain sources of disability income (*dis-val*) and survivor income (*sur-val*).¹⁵ We call this approximation adjusted money income.¹⁶ 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 2018. 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 and 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.

Unemployment insurance has been imputed for 2000–2021 using a crosswalk from

¹⁵ Prior to the 2014–2018 redesign, the relevant variable was “ret-val.” The exclusion of retirement disbursements constitutes approximately 75 percent of the money income excluded.

¹⁶ 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.

Larrimore, Mortenson, and Splinter (2023), which corrects for the underreporting of such payments in the CPS; the crosswalk is based on administrative data. This is particularly influential for 2020–2021 estimates.¹⁷

State and local benefits including temporary disability insurance, public assistance, employment and training, education, and others are distributed to CPS households by their relevant variables, respectively. 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

This section describes the methods used 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 and dividends.

Employer contributions to pension plans and group life insurance are distributed to those who participate in employer-sponsored pension plans and health insurance, respectively, by wage. Imputed interest and dividends are distributed using an imputation derived from the SCF for banking, insurance, and retirement assets for each relevant item, as referred to in section 2C. Owner-occupied housing rental income is distributed using an imputation derived from the CE (2000–2004) and the ACS (2005 forward) by income bracket referred to in section 2D.

5. Health Items

This section describes the methods used to impute health items in the NIPAs to CPS households. This category is the sum of allocations for employer contributions for health insurance, Medicare,¹⁸ Medicaid, military medical insurance, and other medical care

¹⁷ Unemployment insurance from 2022 on is imputed according to the cps variable.

¹⁸ The totals for social security, Medicare, and Medicaid have been adjusted slightly down for the proportion of likely nursing

payment assistance.

Employer contributions to health insurance are distributed by the corresponding CPS variable until 2010. After 2010, employer contributions to health insurance are distributed using data from MEPS, as described in section 2E. The exception is military medical insurance, which is distributed to active military members.

As our objective is to distribute the components of personal income, we must distribute the expenditure on Medicare and Medicaid to participants. Medicare is distributed by assigning average individual expenditures by year, age category, and disability status using Medicare 5% data, as described in section 2E. Medicaid is distributed using a machine-learning approach to assign reciprocity and NIPA state-level totals to impute values, as described in section 2B. Medical assistance is distributed to those who are assigned SNAP and report Women, Infants, and Children (WIC) benefits or other assistance, except the Children's Health Insurance Program (CHIP), which provides assistance to households reporting children covered by CHIP.

6. Net Other Transfers

This section describes the methods used 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 (net) employer and employee contributions for government social insurance,¹⁹ 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) as calculated by TAXSIM35, the NBER program for calculating tax liability from survey data. Employer contributions to workers' compensation and supplemental unemployment are distributed respectively by wage. Other employer and

home residents, not surveyed by CPS, but receiving these benefits. The residual is then distributed equally to all CPS individuals.

¹⁹ In table 2.1, these items are included as part of compensation (line 2) and subsequently deducted in line 25.

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 machine-learning approach described in section 2B.

Refundable tax credits are distributed to CPS households separately by the criteria for the relevant credit where possible. For example, the CPS values for the earned income tax credit (EITC), child tax credit (CTC), and additional child tax credit benefits are used, as is participation in the marketplace for the health premium tax credit.²⁰ Together, these categories represent 95 percent of the refundable tax credits in 2018. One complication is that the TAXSIM calculations impute the credits for the previous calendar year, regardless of when they are received by the households. Thus, there can be an issue of timing for some credits and households.

Energy assistance and WIC are distributed to CPS households by the corresponding CPS variables. All other transfers are distributed to households reporting receipt of WIC, SNAP, or other assistance. Finally, “household current transfer receipts from business (net)” are distributed according to the distribution of all other (summed) components of personal income. As there is no other information available to distribute this income component, its distribution is such that the distribution of personal income will not be meaningfully impacted by its inclusion.²¹

There are some items in the NIPA accounts for which no distributional information is available. These items are distributed such that their inclusion does not change the overall distribution. For example, if income consisted of wages, self-employment, and other, and the Gini of the “wages and self-employment” total were 0.5, the “other” category is now distributed according to the total of “wages and self-employment” and thus the Gini remains at 0.5. One such imputed total is “Household current transfer receipts from business” in table 1 (line 9) for each year.

²⁰ Prior to 2005, TAXSIM imputations are used for the child care tax credit.

²¹ See section 7 for a more detailed explanation of this strategy.

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. As with “household current transfer receipts from business,” this residual is distributed such that it does not meaningfully change the income distribution.

7. Taxes

The calculation of tax liability (and associated components) was done using the NBER TAXSIM35 program (Feenberg and Coutts 1993). CPS households were organized into tax units, whose income inputs were scaled to NIPA totals. Federal taxes before credits (*fiitax*), state taxes before credits (*siitax*), FICA, EITC, CTC, and alternative minimum tax liability were merged from TAXSIM back into the CPS for the calculations to proceed. In the latest CPS redesign, the Census imputation for property taxes is no longer available. Thus, we impute property taxes for TAXSIM by calculating the property tax rate using the ratio of the tax variable to the property value variable ACS by public use microdata area by county. Where a local rate is not available, due to disclosure restrictions, we impute a state-level property tax rate and assign it to those observations. Disposable personal income is calculated by subtracting federal taxes and state taxes from personal income.

8. Provisional Estimates

As with the December 2022 release, the current release includes provisional estimates, now for calendar year 2023. Where possible, calendar year 2023 data was used to calculate these estimates. In cases where these data were unavailable (tax data and some health care expenditure data) an extrapolation was performed using available 2022 data (or 2021 data in the case of Medicare which was unavailable this year).

The most notable extrapolation was for the tax data. Inequality trends for key income sources were evaluated and forecast with the help of 1040 microdata provided to Census. Specifically, we compared the distributions for AGI, wages, interest, and

dividends for weeks 1–43 of the tax year across multiple years. Although these weeks do not constitute a full tax year, they are nevertheless beneficial in assessing the inequality trend year over year (as they do include returns filed by the late filer deadline).

The uncertainty accompanying the provisional estimates for 2023 is largely a function of this extrapolation, as tax data are the key piece of information used for adjusting the CPS for underreporting (as described in section 2A). Thus, significant deviations in the shares of top incomes in the tax data lead to movements in the top shares of the BEA distributional estimates. Accordingly, we have provided a range for the top 1 percent and top 5 percent shares, rather than exact point estimates. The ranges provided reflects analysis conducted with historical data.

This strategy was first tested in the December 2022 release. We found that incorporation of the published SOI data did lead to top share estimates falling within the published ranges. Therefore, this strategy was also utilized for the present release, with the understanding that though changes in top shares may lead to changes in decile shares as well, analysis has shown most of the expected variation to be within the top decile. At present, we do not expect large revisions to these shares (independent of significant updates to NIPA totals or methodology) which fall outside of these intervals.

9. Technical Updates

There has been one major update since the previous release (December 2023). The methodology for adjusting top incomes has been updated this year. In prior releases, total income by category (wages, self-employment, interest, dividends, and rental income) was first summed in the CPS and then compared to the NIPA totals for each item. The “missing” portion (i.e., where the CPS underrepresents the NIPA totals) differs by component. For example, while 5% is missing for wages, 73% is missing for proprietor’s income. Previously, this missing portion was allocated by income source according to the aggregated distribution in the SOI data, as previously described. However, there is significant annual variation in CPS coverage (and totals) for some of these categories (see “methodological update” note for more details). Moreover, item nonresponse has been increasing over time.

Accordingly, the method was adjusted to scale all CPS households to represent the SOI distribution, rather than just the missing portion. This change makes the scaled CPS distribution more similar to the SOI distribution. Though this change was made for all survey years, the SOI adjustment differs prior to 2011. From 2011 forward, tax units in the CPS are grouped by AGI <\$200,000, \$200,000–\$500,000, and \$500,000+ as in prior releases. However, as households in the 2000s had lower incomes, and thus thinner “tails” (fewer observations with AGI above \$500,000), households with AGI \$200,000–\$500,000 were pooled together with households with AGI above \$500,000 pre-2011. This adjustment was made to limit the impact of noise on income shares, due to small sample sizes. In the absence of this grouping, top shares would be further increased for these years by approximately 2pp.

Overall, for 2000–2022, this change resulted in a level-like shift without significant changes in trend. More specifically it raised inequality by increasing the top 1% share by 2.3pp on average throughout the series and the Gini coefficient by 1.4pp. Individual component series were also impacted with a higher share of income for sources such as interest and dividends allocated to the top deciles.

10. Data Sources

There were many data sources (listed below with links) used for the compilation of these estimates. Most are publicly available and made available with varying time lags after the data are collected. Information on incomes from the CPS and CE is regularly available with a 1-year time lag (for example, 2023 values became available in fall of 2024) 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, are regularly available with a 2-year lag (for example, 2021 values became available in November 2023). The SCF is a triennial survey and asks about income received in the previous year, similar to the CPS and the CE. Medical Expenditure Panel Survey (MEPS) published estimates are available with a 1-year lag, while the microdata are available with a 2-year lag.

1. **National Income and Product Accounts.** The primary source for our national macrodata for 2000 to 2023 is the BEA [NIPA data](#). These data are subject to frequent revision. The current set of estimates uses the October 2024 release.
2. **Current Population Survey, March Supplement.** This is the core data series that is the base of the analysis for all noninstitutional households for survey years 2001 to 2024. These data were downloaded from the [NBER website](#) for 2000 to 2017. For 2018–2024, they were used directly from the Census website.
3. **Consumer Expenditure Survey.** This survey was used to estimate rental equivalence for owner-occupied housing from 2000 to 2004. The data were downloaded from [BLS](#).
4. **Survey of Consumer Finances.** This survey was used to distribute imputed interest income from the NIPAs for 2007, 2010, 2013, 2016, 2019, and 2022. For the non-survey years, values were imputed with the methodology described in section 2C. The data were downloaded from the [Federal Reserve Board](#).
5. **Internal Revenue Service Statistics of Income data.** In order to adjust top incomes in the CPS, summary IRS Form 1040 data are aggregated by AGI for tax years 2000–2022 in table 1.4 found in “SOI Tax Stats” on the [IRS website](#).
6. **American Community Survey 2005–2019 (stacked 5-year PUMS files) and 1-year files for 2020 forward.** In order to impute property tax rates at the county level, stacked 5-year [files](#) are used where available, and 1-year files for the most recent year.
7. **Medical Expenditure Panel Survey data.** Tables II.C.1, II.C.3, II.D.1, and II.D.3 were downloaded (2010–2023) using the [MEPSnet Query](#) for the insurance component to impute the employer premium and percent employee contribution.
8. **Centers for Medicare & Medicaid Services.** Medicare 5% microdata were provided to BEA by CMS and are used to perform calculations for Medicare expenditure, at the state-level (by age and disability status).
9. **IRS 1040 Microdata.** 1040 microdata made available to the Census Bureau by the IRS are utilized to calculate provisional distributional results for 2023. These data include

all tax units who have filed returns in the first 43 weeks of the 2023 tax year. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product (Data Management System (DMS) number: P-7503840, Disclosure Review Board (DRB) approval number: CBDRB-FY25-POP001-0028).

10. Unemployment. To address underreporting of unemployment in the CPS, a crosswalk (2000–2021) from Larrimore, Mortenson, and Splinter (2023) is employed. This crosswalk provides imputed reciprocity and benefits to CPS individuals based on administrative data.

References

- Armour, Philip, Gerald Auten, Richard Burkhauser, and Jeff Larrimore. 2021. "Recent Trends in U.S. Top Income Shares in Tax Record Data Using More Comprehensive Measures of Income Including Accrued Capital Gains." *Journal of Political Economy* 129 (5): 1319–1360.
- 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.
- DeBacker, Jason, Bradley Heim, Anh Tran, and Alexander Yuskavage. 2020. "Tax Noncompliance and Measures of Income Inequality." *Tax Notes* 166 (7): 1103–1118.
- Feenberg, Daniel, and Elizabeth Coutts. 1993. "An Introduction to the TAXSIM Model." *Journal of Policy Analysis and Management* 12 (1): 189–194.
- 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. 2020. "Measuring Inequality in the National Accounts." BEA Working Papers, WP 2020–03.
- 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.
- Gindelsky, Marina. 2020. "Technical Document: An Updated Methodology for Distributing Personal Income." BEA website.
- Gindelsky, Marina. 2023. "Technical Document: An Updated Methodology for Distributing Personal Income." BEA website.
- Habib, Bilal. 2018. "How CBO Adjusts for Survey Underreporting of Transfer Income in

Its Distributional Analyses: WP 2018-07.” CBO Working Papers, July.

Johns, Andrew and Joel Slemrod. 2010. “The Distribution of Income Tax Noncompliance.” *National Tax Journal* 63 (3): 397–418.

Kornfeld, Robert. 2023. “Improved Measures of Regulated Investment Companies and Real Estate Investment Trusts for the U.S. National Economic Accounts.” *Survey of Current Business*. May 2023.

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.

Larrimore, Jeff, Jacob Mortenson, and David Splinter. 2023. “Unemployment Insurance in Survey and Administrative Data,” *Journal of Policy Analysis and Management* 42(2) pp. 571-579.

Mataloni, Lisa, Dorian Turner, and Stephanie McCulla. 2023. “Improved Estimates of the National Economic Accounts.” *Survey of Current Business*. November 2023.

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).

Rassier, Dylan G., Bettina H. Aten, Eric B. Figueroa, Solomon Kublashvili, Brian J. Smith, and Jack York. 2021. “Improved Measures of Housing Services for the U.S. Economic Account.” *Survey of Current Business* 101 (5).

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