

Improving the measure of the distribution of Personal Income

By Dennis Fixler, Marina Gindelsky, and David Johnson *

* Bureau of Economic Analysis (dennis.fixler@bea.gov); Bureau of Economic Analysis (marina.gindelsky@bea.gov); University of Michigan (johnsods@umich.edu). Marina Gindelsky is the corresponding author. We greatly thank Andrew Craig for his extensive assistance in creating and evaluating the CPS and NIPA data. We also thank Arjan Bruil for his helpful comments and discussion. The views expressed in this research, including those related to statistical, methodological, technical, or operational issues, are solely those of the authors and do not necessarily reflect the official positions or policies of the Bureau of Economic Analysis or the University of Michigan, or the views of other staff members. The authors accept responsibility for all errors.

With releases of GDP in the U.S., there are typically stories about the impact on inequality and the distribution of growth. The Financial Times stated: “What’s the matter with GDP?” suggesting that GDP is missing information about who gets the increase (Smith, July 2018). Interest has grown regarding the relationship between the distribution of aggregate growth and increase in inequality. This disconnect has been amplified during the past few years, fueled by the Great Recession.

The recent rise in inequality, especially at the top of the distribution, has reinvigorated the effort to produce distributional measures. Led by the creation of the World Inequality Database and Piketty, Saez and Zucman (PSZ) (2018), new consistent measures of the distribution of the national accounts have been

developed (see also Auten and Splinter (2018) and Zwijneburg (2019)).

As Kuznets (1955) stressed, a distribution of the national accounts is necessary to completely examine how economic growth, whose measures rely on national account statistics, is distributed. In earlier work at the Bureau of Economic Analysis (Fixler and Johnson (2014) and Fixler et al. (2017)), tried to develop a distribution of personal income using survey data.

This paper uses survey data, tax records, and administrative data for 2007 and 2012 to improve the measures of the distribution. Supplementary data sources are particularly important for measuring the top income categories and accordingly, we adjust the Current Population Survey (CPS) data to reflect higher income households and estimate alternative measures of inequality. Though reducing the 90/10 ratio, the tail adjustment and inclusion of incomes from supplementary sources significantly raises top income shares and mean income compared to measures calculated using the internal CPS data alone.

I. Measuring Income

The first step in the methodology is to evaluate the source of the gap between the micro and macro data. Fixler and Johnson (2014) demonstrated that the aggregate level of CPS income is much less than the comparable income in the NIPA.¹ Once the definition of income is controlled for, some of the remaining differences could be due to under-reporting in the CPS or high-income individuals could be “missing” from the CPS. If the source of the gap were entirely due to under-reporting, we could close the gap by substituting tax data for the income components of the CPS. Many researchers have attempted to match household survey data to tax or earnings records (see Burkhauser et al. (2017), Bollinger et al. (*forthcoming*), Rothbaum (2015)).

In Fixler, Gindelsky and Johnson (FGJ) 2018, we examined the usefulness of matching the CPS to the tax data and compared the universe in each. Following the method of Fixler and Johnson (2014) and FJFC (2017), we showed that the substitution of income tax variables for the CPS income variables is not a panacea for mis-reporting problems. Moreover, we showed that there is little to gain in terms of differences between matched and unmatched

files. Accordingly, in this paper we use the public use file of the CPS and an alternative strategy for adjusting the top tail of the distribution using tax data, described in the next section.

Our goal, as described in earlier research, is to create a distribution for the US National Account concept of *Personal Income (PI)*, which is the income received by persons from participation in production, from government and business transfers, and from holding interest-bearing securities and corporate stocks. In addition, we eventually hope to develop a table comparable to the decomposition growth table that shows the annual growth rates of GDP and the distribution of these changes across the distribution of households according to personal income.

It is natural to look at the PI income concept for decision making, especially for consumption even though it includes income received by nonprofit institutions serving households, by private non-insured welfare funds, and by private trust funds. PSZ, however, use National Income (NI) claiming: “[it is] in our view a more meaningful starting point, because it is internationally comparable, it is the aggregate used to compute macroeconomic growth, and it is compre-

¹ Rothbaum (2015) recently provides a detailed comparison for each income source.

hensive, including all forms of income that eventually accrue to individuals.” (p. 561) PI and NI are fairly close in aggregate and trend.²

II. Data and Methods

The main data used in our analysis are individual-level data from the publicly available CPS Annual Social and Economic Supplement (CPS ASEC) for survey years 2008 and 2013 (earnings years 2007 and 2012). We supplement those data with other sources of data. First, to account for the well-known deficiency in the number of households in upper income brackets of the CPS, we use Federal tax data to model the tail for incomes above \$500,000 to enhance the distribution of households. Second, to incorporate the components of Personal Income that are not in money income, such as imputed interest and health transfer payments we use the Survey of Consumer Finances (SCF) and data from CMS among other sources. In previous work, the allocation of such NIPA categories was largely confined to a matching algorithm between the Bureau of Labor Statistics (BLS) Consumer Expenditure (CE) Survey and the CPS.

² $PI=NI -[\text{corp. profits} + \text{taxes on production} + \text{contributions for gov. soc. ins.} + \text{net interest} + \text{bus. current transfer} + \text{current surplus of gov. enterp.}] + [\text{personal income receipts on assets} + \text{personal current transfer receipts}]$.

³ A Generalized Beta distribution was also considered but found to fit poorly. Alphas for alternative thresholds were calculated and were

A. CPS upper tail enhancement

An important consideration when utilizing CPS data for distributional measures stems from the underrepresentation of top incomes. To overcome this limitation, we construct an alternative distribution in the following way. Using nonpublic microdata from the Form 1040 Federal income tax data housed at the Census Bureau, we fit a Pareto distribution (estimating alpha by maximum likelihood) for tax units with money incomes of at least \$500,000.^{3,4} The threshold of \$500,000 represents approximately the top 1% of the distribution of tax units and top 0.5% of the distribution of CPS households in 2012. Using the Pareto alpha, we imputed a corresponding distribution for CPS households with pseudo income (calculation described below) of at least \$500,000. CPS incomes for households above this threshold (original mean income = \$847k) were then replaced with imputed values (subsequent mean income = \$1.28m). The components of pseudo income (e.g., wage, business income, transfers, etc.) were then proportionally adjusted to sum to pseudo income.

of a similar magnitude, varying little with time. For more information on the Pareto methodology, see Online Appendix.

⁴ Jenkins (2016) provides a thorough discussion of issues concerning the modeling of the upper income distribution using a Pareto distribution. Nevertheless, we believe our estimate of alpha is “fit for purpose” because of the robust approach to estimation.

B. Addition & Allocation of NIPA Categories

The analysis begins with the concept of pseudo income developed in FJFC. Pseudo income is an intermediate concept that includes categories common to both Census money income and NIPA Personal income and excludes categories that are in the former and not the latter—such as retirement income.

In moving from pseudo income to Personal Income, three groups of variables are added: financial, health, and net transfers. In the financial group, the single largest component to add is imputed interest from financial institutions, insurance companies and pensions (See FJFC, Table 2). Other items include rental income from owner occupied housing, pensions, and life insurance. In the health group, the largest components are group health insurance, Medicare and Medicaid. The net transfers group contains many sources of transfer income including workers compensation, refundable tax credits, and SNAP, while subtracting out employer and employee social contributions. Having set the components of income, the CPS money values are scaled up to the PI level by factors based on the ratio of the CPS weighted total to the PI value.⁵

⁵ For full list of items included in each category and more detailed description of methodology, see Online Appendix Table A.

Essentially, our approach is to use publicly available data to distribute NIPA income to households. For example, the information from the SCF is used to allocate imputed interest. Medicare data from CMS was used for some health categories and CBO imputations were used for Medicaid and SNAP.

III. Results

Table 1 below shows the transition from pseudo income to Personal Income. Because the focus of the analysis is on households, we use the household income value from NIPA Table 2.9 to get to Personal Income—the difference is in the income for the Nonprofit Institutions Serving Households (NPISH).⁶

TABLE 1: COMPONENTS OF PERSONAL INCOME WITH TOTALS AND AVERAGES FOR 2012 IN 2012 DOLLARS

	HH AVG.	TOTAL (\$B)
Pseudo Income	\$87,636	\$10,732
Plus		
Financial	\$14,998	\$1,837
Health	\$16,062	\$1,967
Net Transfers	-\$4,359	-\$534
Equals		
HH Income	\$114,336	\$14,002
+NPISH	\$70	\$9
Personal Income	\$114,406	\$14,010

⁶ Formally, $PI = \text{Household Income} - \text{transfers from NPISH} + \text{NPISH Income} - \text{transfers from Households}$. For 2007 results, see Online Appendix Table B.

As per NIPA table 2.9, Household Income is composed of multiple components: Compensation of Employees, Proprietors' Income with Inventory Valuation and Capital Consumption Adjustment, Rental Income of Households with Capital Consumption Adjustment, Household Income Receipts, and Household Current Transfer Receipts, less contributions to government social insurance. When the data is ranked by equivalized household income, we can calculate the contribution of each income quintile to overall household income as in Figure 1 below for 2012.

FIGURE 1. QUINTILE DISTRIBUTION OF HOUSEHOLD INCOME BY COMPONENT: 2012

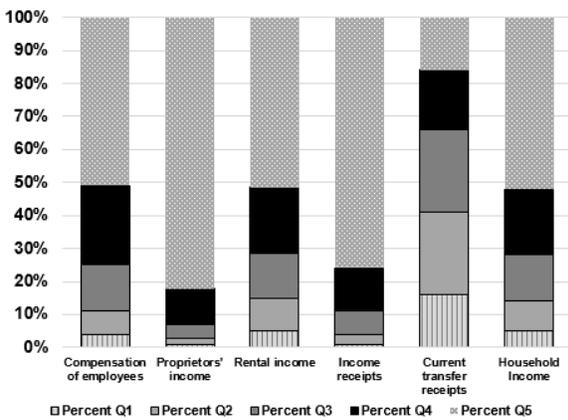


Figure 1 shows that the fourth quintile is 20% of PI while the fifth quintile is 52% in 2012. Not surprisingly 76% of interest and dividend income is received by the top quintile. Also

note that the household current transfer receipts are greater than compensation in the lowest quintile and nearly equal in the second quintile.⁷

Table 2 shows some inequality measures for 2012 (top panel) and 2007 (bottom panel). These measures are based on equivalized income that is computed by dividing the income value by the square root of the number of household members. For more inequality measures, see Online Appendix Table D.

TABLE 2— INEQUALITY STATISTICS FOR EQUIVALIZED HOUSEHOLD INCOME 2007 AND 2012

Income Definition	Mean	Gini	90/10	Top 1% Share
2012				
Money Income	\$46,587	0.46	9.54	8.8%
Pseudo Income	\$57,204	0.53	10.90	14.3%
Household Income	\$74,452	0.46	6.33	13.3%
2007 (in 2012 dollars)				
Money Income	\$48,279	0.44	9.05	7.4%
Pseudo Income	\$58,066	0.50	9.92	12.9%
Household Income	\$73,022	0.45	6.25	12.5%

Note: Pseudo Income includes the tail enhancement

Note that the tail enhancement, along with conversion from money income to pseudo

⁷ For numerical results for 2012 and 2007, see Online Appendix Table C.

income, substantially increases the Gini relative to Census money income (MI). Observe that while the Gini for pseudo income is higher than it is for the adjusted Census MI, there is little difference in the Gini between MI and household income (and therefore personal income). However, the 90/10 ratio and top 1% share fall moving from pseudo income to household, such that they are lower than for MI. This result is indicative of lower income quintiles receiving substantial income from transfers such that the 10th percentile of household income is double that of MI.

With the tail enhancement, the share of the top 5 percent in 2012 is 27.2 percent, which is higher than our original estimate of 23.9 percent in FJFC without the tail enhancement (See Web Appendix Table D). The share of the top 1 percent, 13.3 percent, is equivalent to the post-transfer share in Auten and Splinter (2018), but lower than the post-tax and transfer share in PSZ.

IV. Conclusion

This paper is part of a project to create a distribution for the US national account concept of Personal Income. We focus on three main areas: enhancing the upper tail of the CPS

to add household detail and extending the top income possible beyond the CPS cap, improving the addition and allocation of PI income categories not included in CPS money income, and accomplishing these computations through the use of publicly available microdata. We think the latter is important for a statistical agency because it facilitates transparency and replicability.⁸

Though mean equivalized household money income decreased from 2007 to 2012, household income (and therefore personal income) increased slightly. While the Gini indexes shows little change, the 90/10 ratio shows a significant decline in inequality moving from money income to household.

These trends highlight the importance of distinguishing personal income from money income. The next steps are to develop a time series of estimates and, following the lead of the OECD working group, to develop distributional measures for personal consumption expenditures (PCE).

REFERENCES

Auten, G. and D. Splinter. 2018. "Income Inequality in the United States: Using Tax

⁸ Using the Pareto coefficient (alpha) calculated from the nonpublic 1040 microdata, all further calculations can be made with public use data.

- Data to Measure Long-term Trends.” Unpublished manuscript.
- Bollinger, C., B. Hirsch, B., C. Hokayem, and J. Ziliak. 2019. “Trouble in the Tails? What We Know about Earnings Nonresponse Thirty Years after Lillard, Smith, and Welch,” *Journal of Political Economy*, forthcoming.
- Boushey, H. and A. Clemens. 2018. “Disaggregating growth: Who prospers when the economy grows” Washington Center for Equitable Growth Equitable Growth, March 2018
- Burkhauser, R., Herault N., Jenkins, S. and R. Wilkins, “Top incomes and inequality in the UK: reconciling estimates from household survey and tax return data,” Oxford Economic Papers, Oxford Economic Papers, 70(2), 2018, 301–326, 2017
- Fixler, D. and D. Johnson. 2014. “Accounting for the Distribution of Income in the US National Accounts” in *Measuring Economic Stability and Progress*. D. Jorgenson, J. S. Landefeld, and P. Schreyer, editors, Chicago: University of Chicago Press.
- Fixler, D., D. Johnson, K. Furlong, and A. Craig, A. 2017. “A Consistent Data Series to Evaluate Growth and Inequality in the National Accounts.” *Review of Income and Wealth* 63:S2: S437-S459.
- Fixler, D., M. Gindelsky, and D. Johnson. 2018. “Improving the measure of the distribution of Personal Income,” Paper presented at IARIW 35th General Conference, Copenhagen.
- Jenkins, S. 2017. “Pareto Models, Top Incomes, and Recent Trends in UK Income Inequality”, *Economica*, 84, pp. 261-289.
- Kuznets, S. 1955. “Economic Growth and Income Inequality,” *The American Economic Review*, Vol. 45, No. 1, pp. 1-28.
- Piketty, T., E. Saez, and G. Zucman. 2018. “Distributional National Accounts: Methods and Estimates for the United States since 1913.” *Quarterly Journal of Economics*, 113(2), 553-609.
- Rothbaum, J. 2015. “Comparing Income Aggregates: How do the CPS and ACS Match the National Income and Product Accounts, 2007-2012.” SEHSD Working Paper 2015-01. U.S. Census Bureau.
- Smith, C. 2018. “What’s the matter with GDP?” *Financial Times*, 26 July 2018.
- Zwijnenburg, J. 2019. “Unequal distributions: EG DNA versus DINA approach.” *American Economic Association Papers and Proceedings*, forthcoming.