Improving the Measure of the Distribution of Personal Income

Dennis Fixler, Marina Gindelsky and David Johnson

ESCoE Discussion Paper 2019-14

August 2019

Improving the Measure of the Distribution of Personal Income
Dennis Fixler, Marina Gindelsky and David Johnson
ESCoE Discussion Paper No. 2019-14
August 2019

Abstract

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. Along with 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 Zwijneneburg (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.

Dennis Fixler, Bureau of Economic Analysis, dennis.fixler@bea.gov, Marina Gindelsky, Bureau of Economic Analysis, marina.gindelsky@bea.gov and David Johnson, University of Michigan, johnsods@umich.edu.

Improving the Measure of the Distribution of Personal Income

By Dennis Fixler, Marina Gindelsky, and David Johnson[1]

Improving the measure of the distribution of Personal Income

By Dennis Fixler, Marina Gindelsky, and David Johnson[2]

Introduction

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. Along with 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 Zwijneneburg (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.[3] 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 comprehensive, including all forms of income that eventually accrue to individuals." (p. 561). Though international comparability of NI is a benefit, it must be balanced against the facts that relative to

---

[3] Rothbaum (2015) recently provides a detailed comparison for each income source.

PI, NI includes such aggregates as Business Current Transfer Payments (net) and Net Interest and Miscellaneous payments on assets, whose proper allocation to households is arguable and that PI and its companion disposable PI are crucial to the macroeconomic analysis of consumption. [4] It should be noted that PI and NI are fairly close in aggregate and trend.

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.

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.[5,6] The Census Bureau defines money income 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." The threshold of $500,000 represents approximately the top 1% of the

---

[4] PI=NI –[corp. profits + taxes on production + contributions for gov. soc. ins. + net interest + bus. current transfer + current surplus of gov. enterp.] + [personal income receipts on assets + personal current transfer receipts].

[5] A Generalized Beta distribution was also considered but found to fit poorly. Alphas for alternative thresholds were calculated and were of a similar magnitude, varying little with time. For more information on the Pareto methodology, see Appendix.

[6] 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.

4

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= $842k) 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.

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).[7] 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.[8] Having set the components of income, the CPS money values are scaled up to the PI level (e.g., NIPA totals for each PI category) by factors based on the ratio of the CPS weighted total to the PI value. [9]

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.

---

[7] FISIM is included in this category.

[8] This does not include Other Social Insurance Benefits

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

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).[10]

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**

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.

---

[10] Formally, PI=Household Income – transfers from NPISH +NPISH Income – transfers from Households. For 2007 results, see Appendix Table B.

FIGURE 1. QUINTILE DISTRIBUTION OF HOUSEHOLD INCOME BY COMPONENT: 2012
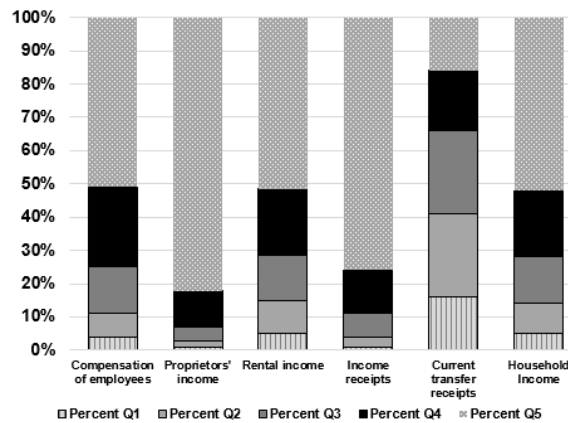
Figure 1 shows that the fourth quintile is 20% of PI while the fifth quintile is 52% in 2012. Not surprisingly 76% of interest and dividend income is received by the top quintile. Also note that the household current transfer receipts are greater than compensation in the lowest quintile and nearly equal in the second quintile.[11]

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 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%

---

[11] For numerical results for 2012 and 2007, see Appendix Table C.

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 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 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 Extension to the States

The U.S. is a large economy comprised of diverse State economies. Accordingly, it is interesting to examine how the States fair with the distribution of Household Income by State.[12]

Table 3 shows the 2012 median and mean household income, equivalized, for each state in 2012[13]. In addition, each median and mean is deflated by the BEA Regional Price Parity (RPP). The RPP is a multilateral (spatial) price index in which the U.S. average is defined as 100. The RPPs range from 117.8 in Hawaii to 86.5 in Mississippi. Table 3 shows that there are different state rankings going from high to low. For example, using nominal median income it is the District of Columbia with the highest medium income, Michigan in the middle and Arkansas the lowest. However, with adjusted median income the ranking is: Minnesota, Indiana and then Hawaii. The last follows from Hawaii having prices on average almost 18% higher than the US

---

[12] The Household Income estimates generated here are different from the State Personal Income estimates produced by BEA in its Regional Economics Program. However, the correlation between these two estimates is 0.99. Two possible reasons for the difference: (1) the allocation of NIPSH, see table 1 and (2) the the CPS sample design and methods, which are geared towards producing national estimates (Current Population Survey, 2013 ASEC Technical Documentation). Accordingly, caution should be exercised when interpreting state-level estimates.

[13] Equivalization is determined by dividing income by the square root of the number of family members in the household.

average. With nominal mean income the ranking is: District of Columbia, Iowa and Arkansas. The ranking with the adjusted mean income is: District of Columbia, Mississippi and Hawaii.

Table 3: Summary Statistics, 2012[14]

	Household Income*					Household Income*			
	Median		Mean			Median		Mean	
	Nominal	RPP Adj.	Nominal	RPP Adj.		Nominal	RPP Adj.	Nominal	RPP Adj.
Alabama	42893	48687	62321	70739	Missouri	47983	53732	78642	88064
Alaska	60384	57291	83773	79481	Montana	43818	46864	64465	68947
Arizona	44065	45381	67833	69859	Nebraska	53988	59590	73454	81076
Arkansas	39479	44964	54540	62118	Nevada	43232	43802	65111	65969
California	53821	47671	89482	79257	New Hampshire	63291	59935	85677	81134
Colorado	54790	54194	80486	79610	New Jersey	58948	51528	85630	74851
Connecticut	62090	56859	96992	88820	New Mexico	41819	44020	81211	85486
Delaware	49027	48446	69910	69081	New York	49429	42870	73306	63579
District of Columbia	66643	56621	113773	96664	North Carolina	42312	46142	60324	65784
Florida	46312	46733	66919	67527	North Dakota	53666	58974	85242	93672
Georgia	45423	49265	64955	70450	Ohio	44208	49450	61560	68860
Hawaii	50018	42460	72068	61179	Oklahoma	46417	51631	65276	72609
Idaho	45066	48250	64181	68716	Oregon	49163	49811	68880	69788
Illinois	50001	49653	79657	79103	Pennsylvania	50864	51691	70901	72053
Indiana	46387	50752	62179	68030	Rhode Island	53991	54647	82758	83764
Iowa	52630	58348	71822	79625	South Carolina	43416	47815	58684	64630
Kansas	49078	54170	68561	75675	South Dakota	50459	56759	69943	78676
Kentucky	43181	48627	56093	63168	Tennessee	44567	49082	63643	70091
Louisiana	44137	48289	58758	64287	Texas	48451	50364	74106	77033
Maine	52292	53142	73140	74329	Utah	47748	49174	70366	72468
Maryland	63310	57502	104432	94852	Vermont	56737	56287	81264	80619
Massachusetts	59813	56110	94786	88918	Virginia	56464	54820	84495	82034
Michigan	49341	52213	71279	75427	Washington	56460	54550	81812	79045
Minnesota	60443	61930	83414	85465	West Virginia	41941	47338	58248	65743
Mississippi	40030	46278	64944	75080	Wisconsin	51378	54949	79046	84541
					Wyoming	51240	53487	73929	77170

*Equivalized Income

Table 4 consists of 3 panels; the quintile distribution of household income and its major
components for the three ranked states in nominal median income. Observe that the distributions

[14] It is important to note that the sample design and methods of weighting CPS data are geared towards producing national estimates (Current Population Survey, 2013 ASEC Technical Documentation). Accordingly, caution should be exercised when interpreting state-level estimate.

of household income and compensation of employees gets much flatter going from the District to Arkansas.

Table 4: State Decompositions

District of Columbia, 2012 (Shares of State Totals)

	Q1	Q2	Q3	Q4	Q5
Household Income	2.9%	4.4%	5.8%	10.5%	76.5%
Compensation of Employees	1.5%	2.9%	4.9%	12.1%	78.6%
Proprietors' income with inventory valuation and capital consumption adjustments	0.2%	0.6%	1.8%	4.3%	93.2%
Rental income of households with capital consumption adjustment	2.0%	2.4%	3.9%	11.0%	80.7%
Household income receipts on assets	0.8%	1.0%	1.8%	4.9%	91.6%
Household current transfer receipts	17.3%	22.2%	21.7%	17.9%	20.9%
Less: Contributions for government social insurance, domestic	2.8%	4.5%	6.2%	15.6%	70.8%

Michigan, 2012 (Shares of State Totals)

	Q1	Q2	Q3	Q4	Q5
Household Income	5.1%	9.6%	15.9%	20.9%	48.6%
Compensation of Employees	3.4%	7.3%	14.2%	26.2%	48.9%
Proprietors' income with inventory valuation and capital consumption adjustments	0.3%	1.8%	3.5%	6.7%	87.8%
Rental income of households with capital consumption adjustment	5.9%	12.7%	18.0%	23.2%	40.3%
Household income receipts on assets	1.3%	2.5%	6.8%	15.8%	73.6%
Household current transfer receipts	13.5%	23.3%	31.7%	17.1%	14.4%
Less: Contributions for government social insurance, domestic	4.6%	10.1%	19.1%	27.2%	39.1%

	Q1	Q2	Q3	Q4	Q5
Household Income	10.2%	15.6%	17.6%	21.0%	35.6%
Compensation of Employees	6.9%	10.6%	17.7%	25.8%	39.0%
Proprietors' income with inventory valuation and capital consumption adjustments	0.6%	4.7%	6.0%	23.1%	65.6%
Rental income of households with capital consumption adjustment	10.8%	18.7%	21.4%	21.8%	27.2%
Household income receipts on assets	3.5%	5.4%	11.1%	14.0%	66.0%
Household current transfer receipts	22.8%	33.8%	22.8%	14.0%	6.5%
Less: Contributions for government social insurance, domestic	8.5%	15.4%	18.7%	25.3%	32.1%

*Equivalized income and categories

Table 5 groups the States by Region and shows the Gini coefficient and the 90/10 ratio—the latter is the ratio of threshold levels—10% of households below and 90% of households below. The Gini coefficients and the 90/10 ratios are not affected by the RPP adjustment because each household in a state is divided by the same price index. The national Gini in 2012 is 0.46 and the 90/10 ratio is 6.33. The average Gini coefficients for each region are in order (the spread between the high and low values are in parentheses): 0.45 (0.13) in West, 0.44 in Northeast (0.08) and Midwest (0.11) and 0.43 (0.16) in South. The average 90/10 ratio for each region are in order (spread between high and low in parentheses): 6.47 (2.5) in West, 6.12 (6.17) in South, 6.03 (2.97) in Northeast, and 5.62 (1.05) in Midwest. The large spread in the South is due to the large 90/10 ratio for the District of Columbia. If instead the District had a 90/10 ratio that was the average of those of its neighboring states, Maryland and Virginia, then it would be 6.5, and the South would then have a mean 90/10 ratio of 5.82 with a spread of 2.19, thereby bringing the South in line with the other regions. It is interesting to note that the 90/10 ratio is lowest in the Midwest and the corresponding spread is less than one-half of those in the other regions.

Table 5: Household Income, 2012

	Gini	90/10		Gini	90/10
West			**South**		
Alaska	0.42	5.39	Alabama	0.44	5.61
Arizona	0.47	7.00	Arkansas	0.42	5.90
California	0.51	7.53	Delaware	0.42	4.85
Colorado	0.46	7.30	District of Columbia	0.54	11.59
Hawaii	0.44	7.09	Florida	0.44	6.02
Idaho	0.42	5.96	Georgia	0.44	6.71
Montana	0.44	5.14	Kentucky	0.38	5.12
Nevada	0.46	6.40	Louisiana	0.40	6.00
New Mexico	0.55	7.64	Maryland	0.53	5.95
Oregon	0.42	6.21	Mississippi	0.49	4.94
Utah	0.44	5.71	North Carolina	0.43	5.33
Washington	0.44	6.62	Oklahoma	0.43	5.78
Wyoming	0.43	6.15	South Carolina	0.40	5.49
			Tennessee	0.43	5.45
			Texas	0.48	6.89
			Virginia	0.48	7.04
			West Virginia	0.41	5.43

*Equivalized Income

Household Income, 2012

	Gini	90/10		Gini	90/10
Northeast			**Midwest**		
Connecticut	0.48	6.67	Illinois	0.48	6.21
Maine	0.40	4.46	Indiana	0.40	5.35
Massachusetts	0.48	6.26	Iowa	0.41	5.35
New Hampshire	0.42	5.65	Kansas	0.42	5.72
New Jersey	0.45	7.09	Michigan	0.44	5.66
New York	0.45	6.09	Minnesota	0.44	5.69
Pennsylvania	0.42	5.33	Missouri	0.51	5.16
Rhode Island	0.47	7.43	Nebraska	0.43	5.57
Vermont	0.42	5.31	North Dakota	0.49	5.97
			Ohio	0.41	5.77
			South Dakota	0.43	5.44
			Wisconsin	0.47	5.56

*Equivalized Income

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.[15]

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 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

---

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

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.

Zwijneneburg, J. 2019. "Unequal distributions: EG DNA versus DINA approach." *American Economic Association Papers and Proceedings,* forthcoming.

Online Appendix

"Improving the Measure of the Distribution of Personal Income"
Dennis Fixler
Marina Gindelsky
David Johnson

I.   NIPA totals by category

We first decompose each item in Table 2.9 into components that can then be distributed to the relevant households.

Online Appendix Table A: NIPA subtotals in billions from Table 2.9*

[P] denotes items that make up Pseudo Income; [F] denotes items that make up Financial Income;

[H] denotes items that make up Health Income; [T] denotes items that make up Net Transfers

	2007	2012
Household income	***11,989.26***	***14,001.55***
Compensation of employees	***7,878.86***	***8,566.73***
Wage and salary Disbursements[P]	6396.8	6927.5
Pension and profit sharing[F]	407.5	439.7
Group Life Insurance[F]	17.8	12.4
Group Health Insurance[H]	517.1	597.2
Old-age, survivors, disability, and hospital insurance[H]	395.5	427.9
Military medical insurance (received)[H]	4.0	5.5
Publicly administered government employee insurance funds[H]	11.5	10.5
Workers' Compensation[T]	65.9	65.7
Supplemental Unemployment[T]	0.5	0.5
Other[T]	62.1	79.9
Proprietors' income with inventory valuation and capital consumption adjustments	***994.0***	***1347.3***
Farm[P]	40.3	60.9
Nonfarm[P]	953.8	1286.4
Rental income of households with capital consumption adjustment	***179.1***	***508.8***
Rental income from owner occupied housing[F]	97.6	353.5
Other private business rental income[P]	81.5	155.3
Household income receipts on assets	***2126.7***	***2118.8***
Household interest income	*1335.3*	*1310.5*
Monetary interest – Pensions[F]	193.9	166.7

16

Imputed interest received by households from depository institutions[F]	145.1	154.4
Life Insurance Carriers[F]	255.4	249.0
Imputed interest received from property and casualty insurance companies[F]	11.8	10.7
From employee pension plans[F]	366.3	450.3
Monetary interest - IRA, KEOGH, Mutual Funds[P]	362.9	279.4
Household dividend income	*791.4*	*808.4*
Household dividend income[P]	791.4	808.4
Household current transfer receipts	**1771.9**	**2410.4**
Government social benefits	*1679.7*	*2300.1*
Medicare[H]	428.2	554.7
Medicaid[H]	324.2	417.5
Other state & local medical care[H]	12.6	13.9
Social security[P]	575.7	762.1
Unemployment insurance[P]	32.7	83.6
Railroad retirement[P]	9.8	11.4
Pension benefit guaranty[P]	2.5	2.4
Veterans Benefits[P]	41.6	70.0
Workers' compensation (received)[P]	14.4	13.2
Temporary disability insurance[P]	4.8	5.5
Black lung benefits[P]	0.6	0.4
Supplemental security income[P]	42.2	53.0
Other public assistance and income maintenance[P]	33.1	37.3
Education assistance[P]	20.8	28.6
State & local employment & training[P]	1.0	1.0
Alaska dividend payments[P]	1.0	0.6
SNAP[T]	30.9	74.9
Refundable tax credits[T]	60.1	98.7
Energy Assistance[T]	2.7	4.1
WIC Food[T]	5.4	6.8
Retired military personnel and dependents at nonmilitary facilities[T]	2.1	1.8
All other government social benefits[P]	33.2	58.5
From business (net)	*18.2*	*23.9*
Other current transfer receipts, from business (net)[T]	18.2	23.9

From nonprofit institutions	*74.0*	*86.4*
Household current transfer receipts from NPISHs[P]	74.0	86.4
Less: Contributions for government social insurance, domestic	*961.4*	*950.5*
Supplementary medical insurance[H]	47.9	60.4
Employer's actual social contributions[T]	461.7	513.3
Employee's actual social contributions[T]	402.3	330.0
Self-employed[T]	49.5	46.8

*Items in bold and italics are calculated subtotals. Table 2.9 last revised on November 20, 2018.

II.   Imputation Strategies

The general strategy for assigning incomes to each individual proceeded as follows: the NIPA total (see disaggregation of Table 2.9 above) was allocated proportionally to individuals (then aggregated up to households) or households which satisfied conditions pertaining to the category in the CPS. For example, "Household Dividend Income" was distributed to individuals by the weighted amount of dividends for each person as a share of the total amount of dividend income in the CPS. If total dividend income in the CPS is $123b, an individual reporting $100 of dividend income with a weight of 656.04 would receive a share of 8.129e-10, and correspondingly would receive a value of $657.09 of NIPA dividend income ($808.36b). In this way, we preserve the distribution of dividend income while scaling the values up such that the total is equal to the NIPA total. Many of the NIPA items were able to be allocated solely on the basis of information available in corresponding CPS categories. Items which were allocated using outside sources of information are described below.

a. Imputed Interest: In moving from pseudo income to Personal Income, the single largest component to add is imputed interest (See FJFC, Table 2). The category contains the imputed interest from financial institutions, insurance companies and pensions. To allocate the PI imputed interest total, we use information from the Survey of Consumer Finances (SCF) to determine the shares of income that come from interest. Those shares are based on nominal values of checking and savings accounts balances, cash value of life insurance policies, and retirement accounts balances as reported in the SCF for each income band. For example, for households with incomes between $100,000, and $200,000, the total cash value of life insurance policies is $1.94b in 2012. The total cash value of life insurance policies received by households in the SCF is $11.7b. Therefore, households in

the CPS with incomes between $100,000, and $200,000 are allocated a "share" of 1.94/11.7=0.166. Once this share has been imputed into the CPS, it is allocated proportionally (by household weight) for each relevant household, as applicable, such that all households together in that income category receive 0.166 of the NIPA value, in this case $10.7b for Imputed Interest Received from Property and Casualty Insurance Companies. The same methodology is used to impute other imputed interest components

b. <u>Medicare:</u> In order to impute the value of Medicare for each household, we allocated actual per capita costs by state for those ages 65+ and under 65 to those who reported receiving Medicare benefits in the CPS in each state. These costs are derived from state-level summary files reported by the Centers for Medicare & Medicaid Services (CMS).

c. <u>Medicaid & SNAP:</u> Two of the variables believed to be underreported in the CPS are Medicaid recipiency and the Supplemental Nutrition Assistance Program (SNAP). To compensate the Congressional Budget Office (CBO) has created an adjustment method for allocating these values to CPS individuals. We use the CBO imputations, as described in Habib (2018) for Medicaid and SNAP. For more information on the CBO imputation procedure, please see Habib (2018) CBO working paper.

d. <u>Supplementary Medical Insurance:</u> This category is based on the distribution of Medicare Part D enrollment. The Medicare Part D enrollment data comes from the CMS Statistics Reference Booklet, specifically tables I.3 and I.4. We used enrollment by age group for Parts A and/or B and Part D to compute a ratio of Part D to Parts A/B enrollment for age groups <65, 65-74, 75-84, and 85+. For example, in calendar year 2012 there were 23.396 million enrollees in Part A/B and 13.712 million Part D enrollees for the age group 65-74. Accordingly, 58.61% (13.712/23.396) of CPS individuals ages 65-74 who reported receiving Medicare benefits were randomly allocated a portion of Supplementary Medical Insurance by household weight. Individuals belonging to the other age groups were similarly allocated.

e. <u>Rental Income from Owner Occupied Housing:</u> Rental values were calculated for each household using information derived from the Consumer Expenditure Survey (CE) produced by the Bureau of Labor Statistics. For each consumer unit in the CE, the reported value of the "Estimated Rental Value of Owned Home" was taken as a share of the reported "Income Before Taxes". For example, if a household reported a monthly estimated rent of $1,000 (therefore $12,000 annually) and an income of $30,000, it would receive a share of 12000/30000=40%. Outliers were removed (i.e., the top 1% of resulting shares). Income deciles were constructed from ranked household income in the CE and the median share per decile was then used to create a crosswalk. Income deciles were

constructed for CPS households ranked on money income. Each owner-occupied CPS household then received the corresponding median rental income share for its income decile.

f. <u>WIC:</u> Program Costs for WIC came from their annual report, WIC Program Participation and Costs found here: https://www.fns.usda.gov/pd/wic-program. Program costs were distributed to CPS individuals reporting WIC recipiency.

g. <u>Alaska Dividend:</u> Dividend payments by the State of Alaska to residents can be found here: https://pfd.alaska.gov/Division-Info/Summary-of-Applications-and-Payments. Each CPS individual residing in Alaska was allocated a share of the total.
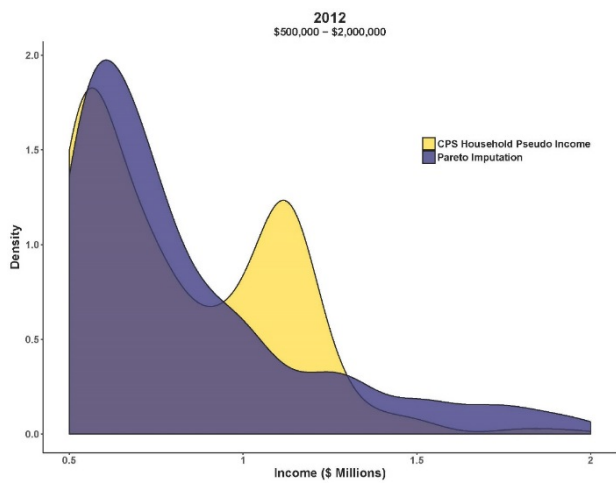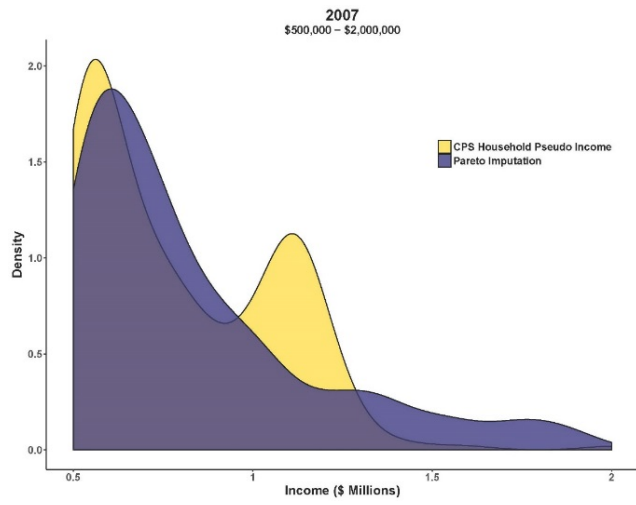
III.   CPS Pareto tail enhancement

Using nonpublic microdata from the Form 1040 Federal income tax data housed at the Census Bureau, we fit a Pareto distribution for tax units with money incomes of at least $500,000. The Pareto distribution used is
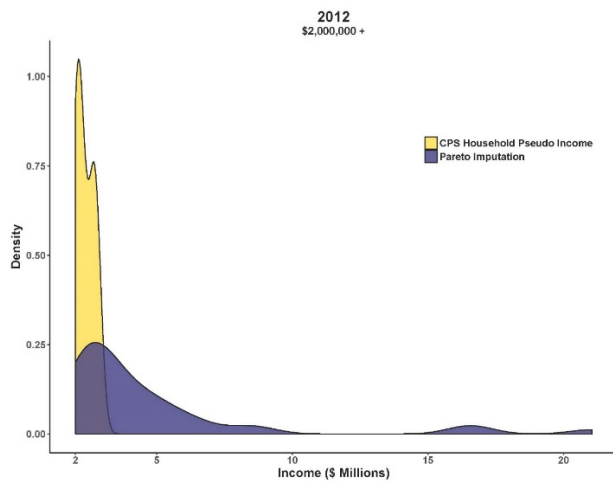
$$F(x) = 1 - \left(\frac{x}{x_0}\right)^{-\alpha}$$

where x is income and x_0 is $500,000 of income. Alpha is the estimated parameter by fitting a Pareto distribution by Maximum Likelihood to the aforementioned 1040 microdata. For more information, see STATA package *paretofit*.

A comparison of the CPS pseudo income distribution and the imputed Pareto-based distribution can be seen in Online Appendix Figure A below for 2007 and 2012. The top panel shows the income brackets from $500,000 to 2 million dollars; the bottom shows the distribution from 2 million dollars and above. The top panel shows that the Pareto distribution fits CPS incomes very well (note that the heaping of CPS data is caused by internal topcodes). The bottom panel shows that using the Pareto imputation replaces the incomes of some CPS households with values outside of the CPS possibilities to be more in line with the administrative data.

<u>Online Appendix Figure A: Pareto Imputation and CPS Pseudo Income 2007 and 2012</u>

2007
$500,000 – $2,000,000

CPS Household Pseudo Income
Pareto Imputation

Density

Income ($ Millions)

2012
$500,000 – $2,000,000

CPS Household Pseudo Income
Pareto Imputation

Density

Income ($ Millions)

2007
$2,000,000 +
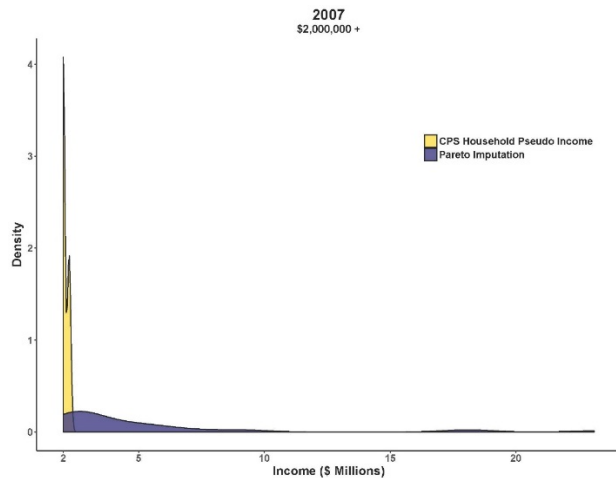
2012
$2,000,000 +

IV.   Supplementary Tables

Online Appendix Table B: Components of Personal Income with Totals and Averages for 2007 and 2012 in 2012 dollars*

2012

	Household average	Totals (billions)
Pseudo Income	**$87,636**	**$10,732**
Plus		
Financial	$14,998	$1,837
Health	$16,062	$1,967
Net Transfers	-$4,359	-$534
Equals		

Household Income	**$114,336**	**$14,002**
+NPISH	$70	$8.6
Personal Income	**$114,406**	**$14,010**

2007

	Household average	Totals (millions)
Pseudo Income	**$89,095**	**$10,405**
Plus		
Financial	$14,003	$1,635
Health	$15,407	$1,799
Net Transfers	-$6,231	-$728
Equals		
Household Income	**$112,275**	**$13,112**
+NPISH	$173	$20
Personal Income	**$112,449**	**$13,132**

*Note: Totals may not add up due to rounding. See Table A for components that make up each category

Online Appendix Table C: Quintile distribution of household income by component: 2012 and 2007 in 2012 dollars

2012

Household income	Total ($B)	% Q1	% Q2	% Q3	% Q4	% Q5
Compensation of employees	8566.7	4%	7%	14%	24%	51%
Proprietors' income with inventory valuation and capital consumption adj.	1347.3	1%	2%	4%	11%	83%
Rental income of households with capital consumption adj.	508.8	5%	10%	14%	20%	52%

Household income receipts	2118.8	1%	3%	7%	13%	76%
Household interest income	1310.5	2%	4%	9%	17%	67%
Household dividend income	808.4	0%	1%	3%	7%	89%
Household current transfer receipts	2410.4	16%	25%	25%	18%	16%
Government social benefits	2300.1	16%	26%	26%	18%	14%
From business (net)	23.9	1%	4%	11%	24%	60%
From nonprofit institutions	86.4	5%	8%	14%	26%	47%
Less: Contrib. for government social insurance, domestic	950.5	4%	10%	17%	26%	43%
Household Income	**14001.6**	**5%**	**9%**	**14%**	**20%**	**52%**

2007

Household income	Total ($B)	% Q1	% Q2	% Q3	% Q4	% Q5
Compensation of employees	8616.7	4%	8%	15%	24%	50%
Proprietors' income with inventory valuation and capital consumption adj.	1087.1	1%	3%	7%	14%	75%
Rental income of households with capital consumption adj.	195.9	5%	9%	13%	19%	54%
Household income receipts	2325.8	2%	3%	7%	13%	75%

Household interest income	1460.3	2%	5%	9%	17%	67%
Household dividend income	865.5	0%	1%	3%	7%	88%
Household current transfer receipts	1937.8	17%	27%	23%	17%	15%
Government social benefits	1837.0	18%	28%	24%	17%	13%
From business (net)	19.9	1%	4%	12%	24%	59%
From nonprofit institutions	80.9	3%	7%	13%	24%	53%
Less: Contrib. for government social insurance, domestic	1051.4	5%	10%	17%	26%	43%
Household Income	**13111.9**	**5%**	**9%**	**14%**	**20%**	**52%**

Online Appendix Table D: Inequality statistics for equivalized household income for 2007 and 2012

	Mean	**Gini**	**90/50**	**90/10**	**Top 5% Share**	**Top 1% share**
2012						
Eq. HH Money Income	$46,587	0.4557	2.6430	9.5380	22.18%	8.79%
Eq. HH Pseudo Income (with tail adj.)	$57,204	0.5253	3.0404	10.904	29.82%	14.33%
Eq. HH Income	$74,452	0.4638	2.7161	6.3263	27.20%	13.25%
2007 (in 2012 dollars)						
Eq. HH Money Income	$48,279	0.4408	2.5853	9.0453	21.61%	7.40%
Eq. HH Pseudo Income (with tail adj.)	$58,066	0.5018	2.8558	9.9150	28.17%	12.91%

Eq. HH Income	$73,022	0.4531	2.6487	6.2485	26.49%	12.46%