

Toward National and Regional Distributions of Personal Income

By Dennis J. Fixler, David S. Johnson, Andrew Craig, and Kevin J. Furlong

THE RELATIONSHIP between macroeconomic growth and income inequality has been the focus of several recent studies (see Organisation for Economic Co-operation and Development (OECD) 2011; Boushey and Hersh 2012; Boushey and Price 2014; Cingano 2014). These studies build on a long research legacy. Almost 80 years ago, Kuznets (1934) in his original report on the national accounts suggested that growth in gross domestic product (GDP) was not sufficient to evaluate the performance of the economy—examining the income distribution was important as well. This view was echoed anew in the recent Economic Report of the President (Council of Economic Advisors 2015) and is the theme of the *Report by the Commission on the Measurement of Economic Performance and Social Progress* (Stiglitz, Sen, and Fitoussi 2009).

Appropriate statistics to measure income distribution are thus vital. In conjunction with Kuznet's article, in the 1950s, the Office of Business Economics, the predecessor to the Bureau of Economic Analysis (BEA), began producing measures of the distribution of income in the United States. These first estimates were released in 1953 and began with estimates for 1947 (Office of Business Economics 1953). Similar to a method proposed in this article, these estimates used the current population survey (CPS) from the Census Bureau to account for distribution and allocated the measure of personal income to quintiles. These estimates were regularly released in the SURVEY OF CURRENT BUSINESS from 1950 to 1962 (Fitzwilliams 1964), and the last estimates were produced for 1971 (Radner and Hinrichs 1974). The estimates were discontinued because of resource constraints.

In this article, BEA is exploring how to best reissue estimates of the distribution of income. The remainder

of this article discusses the following:

- The BEA personal income measure and how it can be reconciled with the Census Bureau money income measure to estimate appropriate personal income distributions. Gini coefficients are computed.
- An analysis and comparison of personal income and money income inequality measures.
- An analysis of regional income distribution for the four Census Bureau regions, using Theil index estimates to measure regional inequality.
- Conclusions drawn from this research effort and a look at future research initiatives.

Constructing Personal Income Data Sources and Methods

Since BEA's published personal income estimate is an aggregate without a distribution, we use the Census Bureau money income measure from the CPS as the basis for the distribution. To do so, it is necessary to bring the Census Bureau money income concept to the BEA national income and product accounts (NIPAs) personal income concept. Briefly, we first reconcile components by adding items that are in personal income but not in money income and removing items in money income that are not in personal income. Because the CPS suffers from underreporting, we adjust the values upward to achieve the national totals.

This paper follows up on the work of two previous papers: Fixler and Johnson (2014) and Fixler, Johnson, Craig, and Furlong (Forthcoming in 2017). These papers construct distributional estimates that are fully consistent with the NIPA personal income concept, closely following the work of McCully (2014), Furlong (2012), and the OECD Expert Group on measuring disparities in national accounts. Previous work by Fixler and Johnson (2014) focused on creating a NIPA-adjusted measure of Census Bureau money income, which kept the definition of money income—and then added other NIPA-specific income components, such as health spending and imputed interest. However, because the Census Bureau definition includes income components that are not included in personal income, such as retirement disbursements, the Fixler Johnson

Dennis J. Fixler is Chief Economist of the Bureau of Economic Analysis (BEA). David S. Johnson is Research Professor and PSID Director at the University of Michigan. Andrew Craig is an economist in the Office of the Chief Economist, and Kevin J. Furlong is an economist in the Office of the Director.

NIPA-adjusted income measure is still conceptually different from personal income, though not as different as the commonly used money income concept. For example, personal income includes, but the NIPA-adjusted money income concept excludes, the following: rental income from owner-occupied housing, employer fringe benefits (for example, retirement contributions and health insurance premiums), and imputed interest on insurance policy reserve funds. Fixler and others (2015) further mapped to NIPA personal income.

To estimate the distribution, we use the annual social and economic supplement of the CPS integrated with the consumer expenditure (CE) survey. The CPS collects data on income while the CE collects data on both income and expenditures. The CPS and the CE surveys are nationwide household surveys designed to represent the U.S. civilian noninstitutional population. There are differences between the surveys in the unit of measure and significant differences in frequency and design (see McCully 2014 for more information on the surveys). The sources used for the NIPA estimates of personal income and outlays are many and diverse; they include sample surveys conducted by the Census Bureau, administrative data from the Social Security Administration, and governmental benefits from other agencies.

To construct distributional estimates, personal income is first decomposed into its underlying detail level, consisting of more than 65 components ranging from wages and salaries and social security disbursements to less obvious components, such as imputed interest on life insurance and pension reserve funds. Each of these components is then matched to corresponding microdata to obtain distributional information. Both CPS and CE surveys are necessary because neither one contains all the information required to fully define personal income. For example, only the CE contains information on the rental value of owner-occupied housing, mortgage interest, and homeowner's insurance, all of which are needed to construct the rental income of owner-occupied housing.

Although both surveys are comprehensive, covering a wide range of income and consumption variables, it is not always possible to find an exact match in the microdata. In these instances, indicator variables are constructed from the microdata and are used to distribute the NIPA aggregates across each household. For example, neither survey contains a variable for employer contributions to pension plans. However, the CPS includes a variable indicating if the person participates in a pension plan or not. This variable is used in combi-

nation with a person's wage, which is assumed to be proportional to the employer contribution. Therefore, a person with a higher wage would receive a larger share of the NIPA aggregate than a person with a lower wage, given that they participated in a pension plan. Similarly, the imputed interest received from depository institutions is assumed to be proportional to a household's saving and checking accounts, two variables obtained from the CE.

Because information is used from two surveys, personal income could not directly be estimated for each household in each survey. To overcome this problem, a synthetic data set was constructed using a statistical matching procedure that links housing units in the CPS to units in the CE through the use of 20 common variables in both surveys. A CPS household and a CE household are assumed to be statistically identical if a distance function between the two is minimized for all possible housing combinations.

Another issue with using the survey data for the NIPA personal income distribution is that the underlying populations covered differ. The CPS and the CE survey cover only the civilian noninstitutional population, while NIPA personal income estimates cover the income (and expenditures) of those defined as U.S. residents in the national accounts, which includes non-profit institutions serving households (NPISHs), the institutionalized population, federal civilian and military personnel stationed abroad, and persons whose usual place of residence is the United States and who are private employees working abroad for a period of less than 1 year. Excluded from the NIPA definition of residents are foreign nationals who work and reside in the United States for part of the year and foreign nationals studying in the United States. In addition, NIPA estimates include the income of those who died during the preceding year and who are not captured in the CPS. Excluding NPISHs' income and accounting for transfers between households and NPISHs gives a measure of household income, which will be used for the integration of the microestimates and macroestimates.

In order to align the NIPA population with that of the household surveys, we adjust the NIPA aggregates to align with the population covered in the household surveys. In most cases, this means removing certain population groups from the estimates, though in a couple of instances, it means adding population groups.

The next step is to construct the totals of each income component defined by the NIPA definition using the synthetic data and calculate scaling factors using

the actual NIPA totals. We then apply these factors to the underlying microdata, hence ratio adjusting each income component for each household using the component specific scaling factors.

Specifically, consider household i , with income,

$$y_i = \sum_j \alpha_j y_{ji}$$

where the scaling factors, α_j , depend on the source, j , of income (for example, wages or dividends) and the term y_{ji} is defined as household i 's income from source j in the integrated data set. The α_j are given by the ratio of aggregate personal income to aggregate income in the surveys (either CPS or CE in the integrated data set); specifically, $\alpha_j = Y_j / X_j$, where Y_j is the aggregate for source j in the personal income measure (in the NIPAs) and X_j is the aggregate for source j in the integrated data. This procedure increases each household's income by source and the new scaled household data is then used to obtain distribution measures.

To illustrate, consider only one source of income, such as wages. Then the scaled income for household i is equal to

$$y_i = \frac{NIPAwages}{CPSwages} \times wages_{CPS_i}$$

Additional sources of income would be similarly calculated and added to the total. This procedure generates a NIPA-based scaled income series for households in the CPS and thereby yields a NIPA-based income distribution.

One limitation of the above approach is that we assume that the levels of underreporting (and the difference between survey reports and NIPA measures) are the same for all households. Hence, every household receives the same scaling factor for each source of income. However, it is likely that different households have different levels of underreporting. Research has shown that there is large underreporting at the top of the distribution (see Sabelhaus and others 2015) and that there is underreporting of government transfer benefits at the bottom of the distribution (Meyer, Mok, and Sullivan 2015).

The main motivation for providing a distribution of

personal income is to measure income inequality. A metric that is often used to capture the inequality in an income distribution is the Gini coefficient. A Gini coefficient is based on the Lorenz curve that illustrates how an actual distribution of income differs from one that has an equal distribution (chart 1). The Gini coefficient is the area of A divided by A+B. If the Gini coefficient equals 0, then the Lorenz curve aligns with the line of equality; if it equals 1, then all of the income accumulates to a single person.

The formula used to compute the Gini coefficients that is used in the tables below is given by

$$Gini\ coefficient = \frac{1}{2n^2\bar{x}} \sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|$$

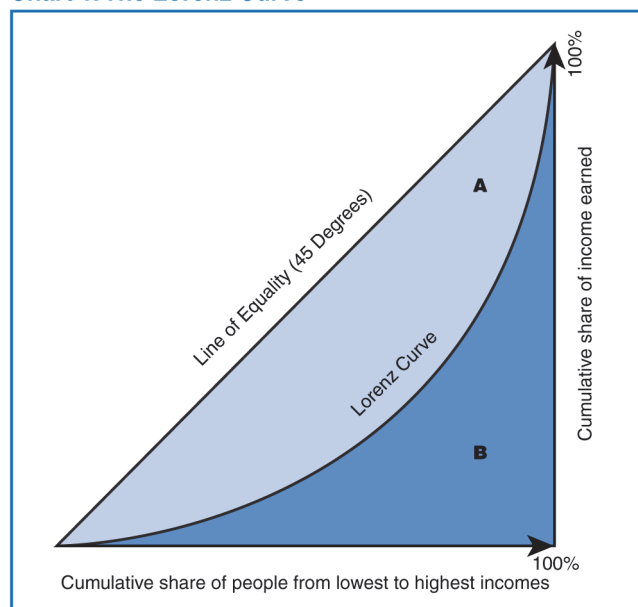
where x_i, x_j are observation pairs, and \bar{x} is the mean of observations.¹

Comparison of Personal Income and Money Income

Since the distributional information mostly comes from the CPS, which is usually used to construct the measure of money income, it is useful to compare

1. See Elbers and others (2005) for formula

Chart 1. The Lorenz Curve



them. Chart 2 illustrates the trend in the mean and median for personal income and money income for 2000–2012. The chart also illustrates the level of the differences between the two income concepts: personal income is larger than money income because it includes items such as employer contributions for retirement and health insurance and government transfers in-kind, such as Medicare, Medicaid, and food stamps. Because households differ in size and composition, income levels are adjusted to reflect the attending impact on the use of income. There are several ways of “equalizing” income, and we use the square root approach that divides each household income by the square root of the number of members in the household; this approach is commonly used in studies of income inequality and poverty measurement. Personal income has a higher mean and median because it is a broader concept of income. In addition, the mean and median for personal income increase over the period; the mean grows 41 percent, and the median grows 39 percent. In contrast, for money income, the mean grows 26 percent, and the median grows 29 percent.

It is also of interest to compare the Gini coefficients published by the Census Bureau, using money income from the CPS, with the Gini coefficients constructed from our household-based personal income estimates.

Table 1 provides the Gini coefficients, and chart 3 illustrates the differences. The Gini coefficient grows 3 percent, indicating an increase in inequality over the period. Observe that the Gini coefficient for personal income is always below that of money income. It is not surprising that the levels of the personal income Gini coefficients are lower as personal income captures more of the impact of the safety net (essentially government transfer payments) on income. Note that the trends are similar. The table also includes disposable personal income, which will be discussed below, and shows that the Gini coefficients are even lower after personal taxes are deducted.

Table 2, which is modeled on NIPA table 2.1, provides a distribution of personal income by component. The first two columns show the difference between the published numbers and what we call the household-based numbers. The difference between these two columns derives from the inclusion of values, such as the income of NPISHs, which are excluded from our household-based estimates. (See the earlier discussion about limiting the analysis to the CPS, which is household based.)

Table 2 is divided into three panels—one for each year. In addition to the published estimates, we provide our household-based estimates, the distribution



Chart 2. Mean and Median Money Income and Personal Income

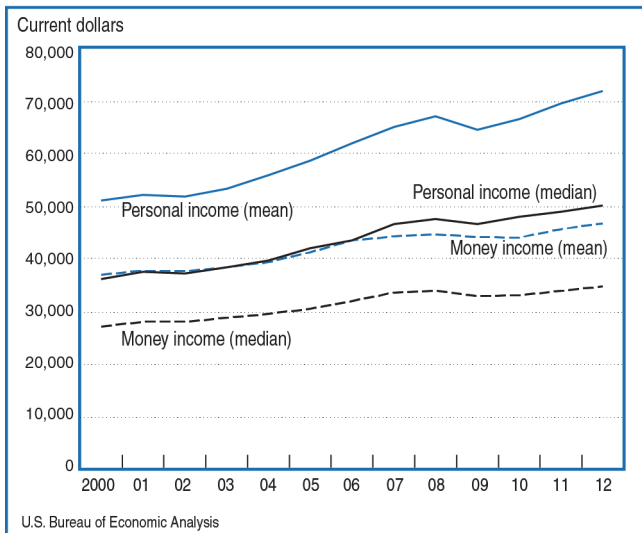


Chart 3. National Gini Coefficients for Money Income, Personal Income, and Disposable Personal Income

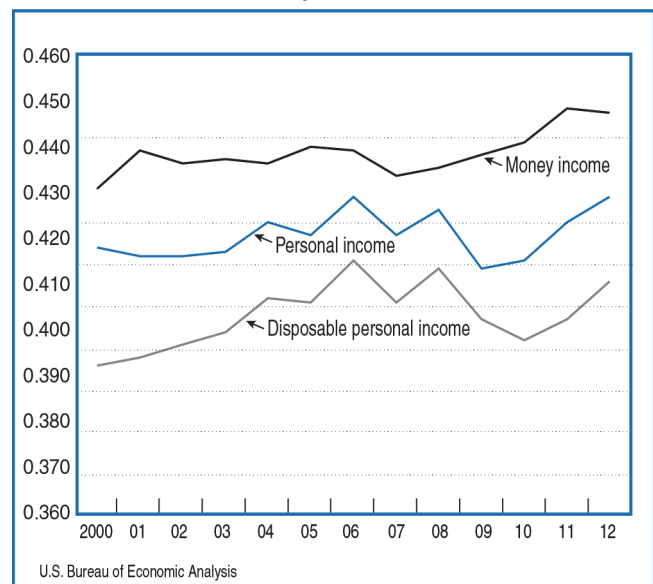


Table 1. National Gini Coefficients for Money Income, Personal Income, and Disposable Personal Income, 2000–2012

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Money income	0.438	0.447	0.444	0.445	0.444	0.448	0.447	0.441	0.443	0.446	0.449	0.457	0.456
Personal income	0.424	0.422	0.422	0.423	0.430	0.427	0.436	0.427	0.433	0.419	0.421	0.430	0.436
Disposable personal income.....	0.396	0.398	0.401	0.404	0.412	0.411	0.421	0.411	0.419	0.407	0.402	0.407	0.416

NOTE. The Gini calculations use equalized income using the square root of family size.

of the components by quintiles and the quintile shares. As explained above, the household-based estimates re-

move from personal income the income from NPISHs and the institutionalized population. Observe that the

Table 2. Personal Income and Its Disposition

	2000											
	Billions of dollars					Percent						
	Pub-lished	House-hold based	Quintiles (levels)					Quintiles (percent of total)				
			1	2	3	4	5	1	2	3	4	5
Personal income.....	8,633	8,395	359	761	1,206	1,853	4,215	4	9	14	22	50
Compensation of employees.....	5,857	5,798	151	481	897	1,466	2,804	3	8	15	25	48
Wages and salaries.....	4,826	4,785	126	395	734	1,197	2,333	3	8	15	25	4
Supplements to wages and salaries.....	1,031	1,014	25	86	163	269	470	2	8	16	27	46
Employer contributions for employee pension and insurance funds.....	686	672	15	57	109	181	311	2	8	16	27	46
Employer contributions for government social insurance.....	345	342	10	30	55	88	159	3	9	16	26	47
Proprietors' income with inventory valuation and capital consumption adjustments.....	758	751	1	14	40	95	600	0	2	5	13	80
Farm.....	32	31	1	2	3	4	22	2	6	10	13	70
Nonfarm.....	726	720	0	12	37	91	579	0	2	5	13	80
Rental income of persons with capital consumption adjustment.....	188	180	11	19	25	35	90	6	10	14	19	50
Personal income receipts on assets.....	1,454	1,425	34	92	162	273	865	2	6	11	19	61
Personal interest income.....	1,070	1,050	30	81	137	228	573	3	8	13	22	55
Personal dividend income.....	383	375	4	11	24	45	291	1	3	7	12	78
Personal current transfer receipts.....	1,083	935	187	220	197	165	165	20	24	21	18	18
Government social benefits to persons.....	1,041	893	186	217	192	154	145	21	24	21	17	16
Social security.....	401	381	89	91	80	61	60	23	24	21	16	16
Medicare.....	219	155	41	38	31	23	21	27	25	20	15	14
Medicaid.....	200	140	24	41	34	27	14	17	29	24	20	10
Unemployment insurance.....	21	20	2	4	5	5	5	10	21	24	22	23
Veterans' benefits.....	25	25	2	2	5	6	9	7	10	21	26	37
Other.....	175	173	29	41	36	31	36	17	24	21	18	21
Other current transfer receipts, from business (net).....	42	42	1	3	5	12	20	3	7	12	28	49
Less: Contributions for government social insurance, domestic.....	706	699	26	68	117	183	305	4	10	17	26	44
Less: Personal current taxes.....	1,232	1,189	13	45	106	228	797	1	4	9	19	67
Equals: Disposable personal income.....	7,401	7,205	346	716	1,101	1,626	3,418	5	10	15	23	47
	2006											
	Billions of dollars					Percent						
	Pub-lished	House-hold based	Quintiles (levels)					Quintiles (percent of total)				
			1	2	3	4	5	1	2	3	4	5
Personal income.....	11,389	11,014	466	975	1,529	2,380	5,664	4	9	14	22	51
Compensation of employees.....	7,502	7,416	207	586	1,107	1,868	3,648	3	8	15	25	49
Wages and salaries.....	6,057	6,006	172	475	887	1,491	2,982	3	8	15	25	50
Supplements to wages and salaries.....	1,445	1,410	35	111	221	377	666	2	8	16	27	47
Employer contributions for employee pension and insurance funds.....	998	967	22	75	152	261	456	2	8	16	27	47
Employer contributions for government social insurance.....	447	443	13	36	68	116	210	3	8	15	26	47
Proprietors' income with inventory valuation and capital consumption adjustments.....	1,054	1,044	4	21	56	148	817	0	2	5	14	78
Farm.....	36	36	0	2	3	9	22	1	4	8	24	63
Nonfarm.....	1,018	1,009	3	19	53	139	794	0	2	5	14	79
Rental income of persons with capital consumption adjustment.....	208	198	12	19	27	41	99	6	9	13	21	50
Personal income receipts on assets.....	1,938	1,901	34	96	176	316	1,278	2	5	9	17	67
Personal interest income.....	1,215	1,193	31	85	151	254	671	3	7	13	21	56
Personal dividend income.....	724	708	3	11	25	62	607	0	2	4	9	86
Personal current transfer receipts.....	1,610	1,370	243	337	314	249	226	18	25	23	18	16
Government social benefits to persons.....	1,588	1,349	242	335	311	244	217	18	25	23	18	16
Social security.....	544	516	103	130	114	87	82	20	25	22	17	16
Medicare.....	399	282	63	74	62	44	39	22	26	22	15	14
Medicaid.....	299	210	35	62	55	38	20	17	29	26	18	10
Unemployment insurance.....	30	30	3	6	7	8	7	10	19	24	25	22
Veterans' benefits.....	39	38	2	5	10	10	11	6	14	25	26	29
Other.....	277	273	36	58	63	58	58	13	21	23	21	21
Other current transfer receipts, from business (net).....	22	21	1	2	3	5	9	5	11	16	25	43
Less: Contributions for government social insurance, domestic.....	923	914	35	86	152	241	399	4	9	17	26	44
Less: Personal current taxes.....	1,352	1,305	28	65	131	265	816	2	5	10	20	63
Equals: Disposable personal income.....	10,037	9,709	438	910	1,398	2,116	4,848	5	9	14	22	50
	2012											
	Billions of dollars					Percent						
	Pub-lished	House-hold based	Quintiles (levels)					Quintiles (percent of total)				
			1	2	3	4	5	1	2	3	4	5
Personal income.....	13,888	13,394	554	1,177	1,838	2,875	6,951	4	9	14	21	52
Compensation of employees.....	8,607	8,509	210	627	1,234	2,128	4,311	2	7	15	25	51
Wages and salaries.....	6,932	6,873	175	508	984	1,690	3,516	3	7	14	25	51
Supplements to wages and salaries.....	1,674	1,636	34	119	250	438	795	2	7	15	27	49
Employer contributions for employee pension and insurance funds.....	1,161	1,128	21	81	175	308	542	2	7	16	27	48
Employer contributions for government social insurance.....	514	509	13	38	75	129	253	3	8	15	25	50
Proprietors' income with inventory valuation and capital consumption adjustments.....	1,260	1,249	3	19	49	146	1,032	0	2	4	12	83
Farm.....	72	72	(0)	1	2	11	58	-1	1	3	15	81
Nonfarm.....	1,188	1,177	3	18	47	135	975	0	2	4	11	83
Rental income of persons with capital consumption adjustment.....	533	512	28	50	68	106	261	5	10	13	21	51
Personal income receipts on assets.....	2,089	2,049	28	88	176	346	1,412	1	4	9	17	69
Personal interest income.....	1,256	1,235	25	75	146	274	715	2	6	12	22	58
Personal dividend income.....	833	815	4	13	30	72	697	0	2	4	9	86
Personal current transfer receipts.....	2,351	2,015	320	476	464	398	357	16	24	20	20	18
Government social benefits to persons.....	2,308	1,972	319	474	457	388	334	16	24	20	20	17
Social security.....	762	723	128	175	163	135	121	18	23	19	17	17
Medicare.....	555	392	79	98	85	69	60	20	25	22	18	15
Medicaid.....	417	292	44	85	78	56	29	15	2	27	19	10
Unemployment insurance.....	84	83	11	16	20	19	17	13	19	25	23	20
Veterans' benefits.....	70	69	3	5	15	21	25	4	8	21	30	36
Other.....	419	414	54	94	95	87	82	13	23	23	21	20
Other current transfer receipts, from business (net).....	43	42	1	3	6	10	23	2	7	15	23	53
Less: Contributions for government social insurance, domestic.....	951	942	34	86	154	249	418	4	9	16	26	44
Less: Personal current taxes.....	1,504	1,451	22	58	130	276	966	2	4	9	19	67
Equals: Disposable personal income.....	12,384	11,943	531	1,119	1,708	2,600	5,985	4	9	14	22	50



household-based estimates are around 97 percent of the published value of personal income for the 3 years.

To simplify the discussion of the table, the focus will be on three primary components of income that are frequently discussed in the context of examining the income distribution: labor income, capital income, and transfer payments. The first will be represented by wages and salaries, the second by personal income receipts on assets, and the third by personal current transfer receipts.

Looking at the estimates for 2000, the share of wages and salaries earned by the top quintile is nearly twice the next quintile and over 15 times the share earned by the first quintile. The pattern is the same for income receipts on assets, but the dominance of the top quintile is stronger, approximately three times the share of the next quintile. For personal current transfer receipts, the first three quintiles have the largest shares, but the distribution of transfer receipts is relatively flat. This result is likely due to the counterbalance between the lower quintiles receiving higher shares of Medicare and Medicaid and the higher quintiles receiving unemployment insurance, veterans benefits and refundable tax credits—the last is embedded in the “other” category.

For 2006, the pattern of wages and salaries is the same as it is for personal income receipts on assets. For personal current transfer receipts, there is a reduction in the share of the first quintile and a reduction in the share of the top quintile—Medicare and Medicaid take a higher share. Note again that the “other” category is significant.²

For 2012, the patterns described above remain, though there is some variation, there is no noteworthy departure from trend.

The estimates in table 2 are consistent with the trend for the Gini coefficients for personal income il-

2. It should be noted that the ratio of the household-based estimates of Medicare and Medicaid to the published estimates are around 70 percent, much lower than the ratio of the other components because of the “recent death discrepancy”; while the NIPAs capture the medical expenses for the recently deceased, the survey-based CPS does not. It should also be noted that this is also a consequence of controlling to the published personal income estimates. Because government transfers are poorly captured in the CPS, they have large scale factors (low coverage ratios). By scaling the microdata, the monetary value of government transfers assigned to each household will be much higher than what the household actually received. For example, if household “i” in the CPS reported Medicaid of \$20,000 and the coverage ratio is 50 percent (scale factor = 2), the adjusted value will be \$40,000. As a result, this household will appear to have more personal income than a household that did not receive Medicaid, all else being equal.

lustrated above. Observe that the shares of the top quintile are stable and that the magnitudes increase slightly. Though we don’t provide the details in the table, it is clear from the chart of the Gini coefficients that we would expect the shares of the top quintile to have declined in 2009 and 2010 and that there was a subsequent increase in inequality. Note that the Gini coefficients for 2006 and 2012 are the same.

Regional Analysis

The focus of income distribution analysis is usually at the national level. However, it is also of interest to see how inequality is distributed across the country. Because the decomposability of the Gini index is not straightforward, we use the Theil index to measure regional inequality. The Theil index is based on information theory and was developed in Theil (1967), which also provided an alternative derivation for the Tornqvist index formula. The Theil inequality measure is given by

$$\text{Theil index} = \sum_i f_i \left(\frac{y_i}{\mu} \right) \log \left(\frac{y_i}{\mu} \right)$$

where y_i is observation i (income of household), f_i is the population share of observation i , and μ is mean of observations.³

One of the analytically useful features of the Theil index is that it can be decomposed into between-group and within-group effects:

$$\text{Theil index} = \underbrace{\sum_j g_j \left(\frac{\mu_j}{\mu} \right) \log \left(\frac{\mu_j}{\mu} \right)}_{\text{Between}} + \underbrace{\sum_j GE_j g_j \left(\frac{\mu_j}{\mu} \right)}_{\text{Within}}$$

where j refers to the subgroup, g_j refers to the population share of subgroup j , and GE_j refers to the inequality (Theil) in subgroup j , μ_j refers to the mean in subgroup j , and μ refers to the total population mean.

Intuitively, the between-group component of inequality measures the level of inequality that would occur if everyone within each group j had income level μ_j ; that is, everyone in the group had the mean level of income. We also provide a raw Theil index; this is the

3. See Elbers and others (2005) for formula.

Theil index for a member of a subgroup in isolation. The Theil index measures of inequality are similar in magnitude to the Gini coefficients as indicated in table 3 and chart 4.

Chart 4 suggests that there is essentially a level difference between the inequality measures, while the trends are similar.⁴ Using the decomposition above, table 4 shows how the national Theil index is decomposed within and between four Census Bureau regions of the United States.

At the national level, the Theil index increases 8.3 percent between 2000 and 2006, while only increasing

4. Since the Theil index is more sensitive to changes at the top of the distribution, this could contribute to the larger increases between 2010 and 2012. It should be noted that similar increases in the Theil index occur with the Census Bureau data.

Chart 4. National Theil Indexes and Gini Coefficients

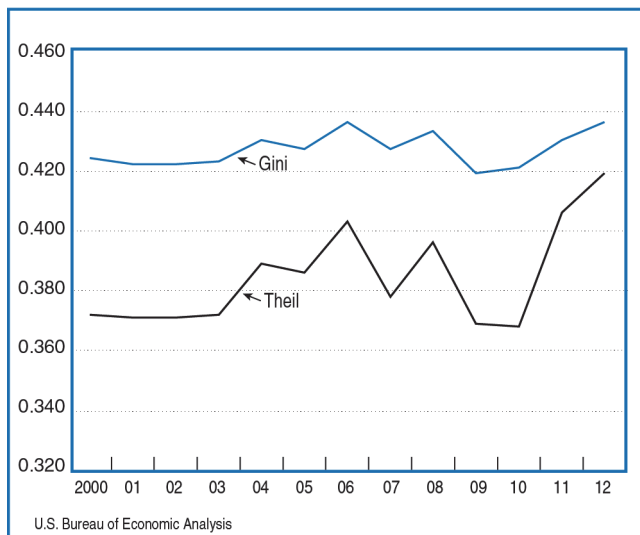


Table 4. National Theil Decomposition

	2000	2006	2012
National decomposition	0.372	0.403	0.419
Regional decomposition			
Between	0.002	0.004	0.003
Within	0.369	0.400	0.416
Northeast	0.076	0.081	0.086
Midwest	0.076	0.083	0.083
South	0.129	0.138	0.145
West	0.088	0.097	0.103

4.0 percent between 2006 and 2012; this is likely due to the impact of the recession. Note that the sum of the between-group and within-group components is equal to the national Theil index and that the sum of the regional Theil indexes is equal to the within-group component. The very small value of the between-group component of the Theil index indicates that geography plays almost no role in determining the level of inequality; it is almost entirely coming from the within-group region measure. Note that the South has the highest level of inequality for each year.

We next consider each region separately and recompute a Theil index for each state in the region. In tables 5–8, for each state we compute a “raw Theil” index value, which is simply the inequality in the state not considering its membership in the region. To compute the contribution of each state to the region’s index, the raw Theil index value is basically weighted by the state’s share of population in the region to obtain the contribution to the within-group calculation. Again, the overall Theil index is the sum of the between-group and within-group component.

Table 5 provides the specifics for the Northeast region. The raw Theil index gives a measure of inequality in the state alone, and note that New York is near the top in the level of inequality in 2006 and 2012, but in 2000, the level is higher in Maine. Maine and Vermont start out with high raw Theil measures in 2000, but

Table 5. Theil Decomposition for the Northeast Region

	2000		2006		2012	
	Raw Theil	Contribution to within	Raw Theil	Contribution to within	Raw Theil	Contribution to within
Maine	0.397	0.009	0.275	0.009	0.359	0.009
New Hampshire	0.317	0.008	0.324	0.008	0.374	0.010
Vermont	0.397	0.004	0.376	0.004	0.364	0.004
Massachusetts	0.364	0.045	0.436	0.056	0.424	0.057
Rhode Island	0.335	0.006	0.402	0.008	0.463	0.009
Connecticut	0.303	0.020	0.438	0.034	0.437	0.033
New York	0.391	0.137	0.428	0.144	0.450	0.150
New Jersey	0.373	0.061	0.376	0.066	0.413	0.069
Pennsylvania	0.346	0.075	0.361	0.072	0.353	0.074
Mean	0.358		0.379		0.404	
Standard Deviation	0.033		0.051		0.040	
Within		0.366		0.400		0.414
Between		0.001		0.007		0.006
Theil		0.368		0.407		0.420

Table 3. National Theil Index and Gini Coefficients, 2000–2012

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Theil index	0.372	0.371	0.371	0.372	0.389	0.386	0.403	0.378	0.396	0.369	0.368	0.406	0.419
Gini coefficients	0.424	0.422	0.422	0.423	0.430	0.427	0.436	0.427	0.433	0.419	0.421	0.430	0.436

these steadily decline by 2012. Connecticut and Massachusetts on the other hand have higher values in 2006 and 2012 than they do in 2000. Once population share is considered, New York has the greatest contribution to the within-group measure. Again, the between-group measure is extremely small and indicates that geography plays almost no role; it is all about the inequality in each state.

The Midwest region has a more equal distribution of income, as indicated by the last line of table 6. For some states, the increase in the raw Theil is large. Notably, the raw Theil for North Dakota increases 46 percent between 2000 and 2012—likely because of the growth in mining and fracking. A large growth in North Dakota’s personal income was also found in BEA’s published state personal income estimates. It should be pointed out that the correlation between BEA’s published state personal income estimates and the state personal income estimates computed in this paper is 0.99 for the 3 years examined.

Table 7 gives the state decomposition for the South region. The District of Columbia has high levels of inequality for each year; however, its contribution to the within-group measure is very small because of its low population share. Mississippi also has substantial increases in the raw Theil in 2006 and 2012 relative to 2000.

Table 6. Theil Decomposition for the Midwest Region

	2000		2006		2012	
	Raw Theil	Contribution to within	Raw Theil	Contribution to within	Raw Theil	Contribution to within
Ohio.....	0.321	0.056	0.331	0.054	0.351	0.054
Indiana.....	0.303	0.026	0.346	0.031	0.311	0.026
Illinois.....	0.345	0.067	0.390	0.077	0.477	0.098
Michigan.....	0.340	0.055	0.358	0.053	0.396	0.058
Wisconsin.....	0.308	0.025	0.391	0.035	0.424	0.038
Minnesota.....	0.343	0.031	0.343	0.030	0.352	0.032
Iowa.....	0.323	0.014	0.353	0.015	0.354	0.016
Missouri.....	0.327	0.030	0.414	0.038	0.397	0.035
North Dakota.....	0.319	0.002	0.386	0.003	0.466	0.006
South Dakota.....	0.335	0.003	0.362	0.004	0.430	0.005
Nebraska.....	0.339	0.008	0.333	0.009	0.366	0.011
Kansas.....	0.310	0.012	0.453	0.020	0.397	0.017
Mean.....	0.326		0.371		0.393	
Standard deviation.....	0.014		0.035		0.047	
Within.....		0.329		0.369		0.396
Between.....		0.003		0.002		0.004
Theil.....		0.332		0.371		0.400

Finally, table 8 presents the estimates for the states in the West region. The increase in the raw Theil index for New Mexico is eye catching. Also note that the contribution of California to the within-group component is large because of their high population share.

Table 7. Theil Decomposition for the South Region

	2000		2006		2012	
	Raw Theil	Contribution to within	Raw Theil	Contribution to within	Raw Theil	Contribution to within
Delaware.....	0.317	0.009	0.272	0.009	0.314	0.009
Maryland.....	0.373	0.025	0.349	0.023	0.361	0.024
District of Columbia.....	0.470	0.003	0.495	0.004	0.500	0.005
Virginia.....	0.402	0.036	0.421	0.036	0.454	0.040
West Virginia.....	0.303	0.005	0.376	0.006	0.384	0.006
North Carolina.....	0.337	0.025	0.384	0.030	0.400	0.031
South Carolina.....	0.316	0.012	0.399	0.015	0.393	0.014
Georgia.....	0.366	0.030	0.315	0.026	0.408	0.033
Florida.....	0.333	0.052	0.416	0.075	0.412	0.068
Kentucky.....	0.409	0.017	0.430	0.015	0.310	0.010
Tennessee.....	0.497	0.029	0.375	0.019	0.399	0.022
Alabama.....	0.357	0.014	0.436	0.017	0.464	0.018
Mississippi.....	0.365	0.009	0.490	0.011	0.466	0.009
Arkansas.....	0.334	0.007	0.366	0.008	0.382	0.008
Louisiana.....	0.343	0.012	0.386	0.013	0.383	0.013
Oklahoma.....	0.440	0.014	0.457	0.014	0.371	0.012
Texas.....	0.421	0.088	0.436	0.089	0.426	0.095
Mean.....	0.375		0.375		0.402	
Standard deviation.....	0.055		0.055		0.049	
Within.....		0.381		0.403		0.409
Between.....		0.008		0.007		0.009
Theil.....		0.389		0.411		0.418

Table 8. Theil Decomposition for the West Region

	2000		2006		2012	
	Raw Theil	Contribution to within	Raw Theil	Contribution to within	Raw Theil	Contribution to within
Montana.....	0.344	0.009	0.323	0.009	0.357	0.009
Idaho.....	0.346	0.006	0.333	0.006	0.374	0.007
Wyoming.....	0.326	0.002	0.353	0.002	0.352	0.003
Colorado.....	0.338	0.025	0.399	0.031	0.396	0.030
New Mexico.....	0.291	0.006	0.474	0.011	0.608	0.018
Arizona.....	0.371	0.028	0.393	0.032	0.388	0.031
Utah.....	0.315	0.009	0.293	0.009	0.349	0.011
Nevada.....	0.392	0.011	0.380	0.013	0.368	0.012
Washington.....	0.446	0.044	0.355	0.035	0.346	0.034
Oregon.....	0.394	0.022	0.433	0.023	0.341	0.017
California.....	0.379	0.211	0.417	0.224	0.454	0.242
Alaska.....	0.265	0.003	0.301	0.003	0.329	0.003
Hawaii.....	0.297	0.005	0.355	0.006	0.382	0.007
Mean.....	0.346		0.346		0.388	
Standard deviation.....	0.048		0.048		0.070	
Within.....		0.376		0.400		0.420
Between.....		0.003		0.005		0.005
Theil.....		0.379		0.404		0.426

Conclusions

This article has provided distributional estimates of personal income along with measures of inequality both nationally and regionally. It shows that the level of inequality has increased in recent years and that while the level of inequality has increased in the four regions, the rate of increase is not the same across the country.

Our derivation of a distribution of BEA's personal income from the Census Bureau's CPS is not without limitations. In addition to the mapping and scope issues listed above, the main limitation is using the same scaling factor for households to move from the Census Bureau money income concept to personal income. Clearly, the factor for the income receipts on assets is different in the upper quintile households than in the lower quintile households. One way to improve on the factors would be to use federal income tax data to inform the factors. Indeed, such a procedure was experimented with in Fixler and Johnson (2014). The use of tax data may improve other aspects of the mapping. There is a parallel research effort using such data to measure income inequality; see for example, Auten and Splinter (2016). Another limitation is the focus on nominal income. Future research will examine inflation-adjusted (real) personal income and explore the use of BEA's regional price parities to examine regional income inequality.

The distribution of personal income presented above is similar in spirit to the distribution of national income presented in Piketty, Saez, and Zucman (2016). National income is broader than personal income and conceptually includes such categories as corporate profits, taxes on production and imports less subsidies, contributions for government social insurance, business current transfer payments and the current surplus of government enterprises. To create a distribution of national income, they too use microlevel data—data from the CPS and income tax records. And to capture the additional categories listed, they must impute values to the corresponding values for households. Because our focus is on the distribution of household income and its ultimate relationship to personal consumption expenditures, a subject for future research, our use of personal income is appropriate.

Measuring the distribution of household income has received worldwide attention. BEA participates in an Organisation for Economic Co-operation and Development (OECD) Expert Group on Disparities in a National Accounts Framework. The goal of this group is to establish a methodology to construct distributional estimates of income, consumption, and saving

consistent with national accounting concepts using microdata. The results of the group's efforts are summarized in Fesseau and Mattionetti (2013) and Zwi-jnenburg, Bournot, and Giovannelli (2016). Generally, the participating countries were able to provide estimates in accordance with the proposed methodology. However, several shortcomings were identified, including the following: (1) the lack of microdata on several national account-specific income components and (2) substantial data gaps between microaggregates and national account totals. Going forward, this expert group will continue to refine the methodology to improve the shortcomings mentioned above, with the aim of eventually establishing a regular publication of distributional results on a per country basis.

References

- Auten, Gerald, and David Splinter. 2016. "Using Tax Data to Measure Long-Term Trends in U.S. Income Inequality." Unpublished paper.
- Boushey, Heather, and Adam Hersh. 2012. *The American Middle Class, Income Inequality, and the Strength of Our Economy: New Evidence in Economics.* Center for American Progress Report.
- Boushey, Heather, and Carter C. Price. 2014. "How are Economic Inequality and Growth Connected?" Washington Center for Equitable Growth.
- Cingano, Federico. 2014. *Trends in Income Inequality and Its Impact on Economic Growth.* OECD Social, Employment, and Migration Working Paper No. 163. Paris: OECD, December.
- Council of Economic Advisors. 2015. *Economic Report of the President, 2015.* Washington, DC: U.S. Government Printing Office.
- Elbers, Chris, Peter Lanjouw, Johan A. Mistiaen, and Berk Özler. 2005. "Re-Interpreting Sub-Group Inequality Decompositions." Policy Research Working Paper 3687. Washington, DC: World Bank, August.
- Fesseau, Maryse, and Maria Liviana Mattionetti. 2013. "Distributional Measures Across Household Groups in a National Accounts Framework: Results From an Experimental Cross-Country Exercise on Household Income, Consumption, and Saving." OECD Statistics Working Papers No. 2013/04. Paris: Organisation for Economic Co-operation and Development Publishing.
- Fitzwilliams, Jeannette M. 1964. "Size Distribution of Income in 1963." *SURVEY OF CURRENT BUSINESS* 44 (April 1964).
- Fixler, Dennis F. and David S. Johnson. 2014. "Accounting for the Distribution of Income in the U.S. National Accounts." In *Measuring Economic Stability*

and Progress, edited by Dale W. Jorgenson, J. Steven Landefeld, and Paul Schreyer. Chicago: University of Chicago Press, for the National Bureau of Economic Research.

Fixler, Dennis F., David S. Johnson, Andrew Craig, and Kevin Furlong. Forthcoming. "A Consistent Data Series to Evaluate Growth and Inequality." Forthcoming in 2017. *Review of Income and Wealth*.

Furlong, Kevin J. 2012. "Synthetic Data—Unconstrained Statistical Matching: Technical Appendix." Unpublished. Bureau of Economic Analysis.

Kuznets, Simon. 1934. *National Income, 1929–1932*. Senate Document No. 124, 73rd Congress, 2nd Session. Washington, DC: U.S. Government Printing Office.

McCully, Clinton P. 2014. "Integration of Micro and Macro Data on Consumer Income and Expenditures." In *Measuring Economic Stability and Progress*, edited by Dale W. Jorgenson, J. Steven Landefeld, and Paul Schreyer. Chicago: University of Chicago Press, for the National Bureau of Economic Research.

Meyer, Bruce D., Wallace K.C. Mok, and James X. Sullivan. 2015. "The Under-Reporting of Transfers in Household Surveys: Its Nature and Consequences." Unpublished Paper. University of Chicago, June.

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

Organisation for Economic Co-operation and Development (OECD). 2011. *How's Life? Measuring Well-*

Being. Paris: OECD Publishing.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman. 2016. "Distributional National Accounts: Methods and Estimates for the United States Since 1913." National Bureau of Economic Research Working Paper 22945.

Radner, Daniel B., and John C. Hinrichs. 1974. "Size Distribution of Income in 1964, 1970 and 1971." *SURVEY OF CURRENT BUSINESS* 54 (October).

Sabelhaus, John, David Johnson, Stephen Ash, David Swanson, Thesia I. Garner, John Greenlees, and Steve Henderson. 2015. "Is the Consumer Expenditure Survey Representative by Income?" In *Improving the Measurement of Consumer Expenditures*, edited by Christopher Carroll, Thomas F. Crossley, and John Sabelhaus. Chicago: University of Chicago Press, for the National Bureau of Economic Research.

Stiglitz, Joseph E., Amartya Sen, and Jean-Paul Fitoussi. 2009. *Report by the Commission on the Measurement of Economic Performance and Social Progress*. New York: United Nations Press.

Theil, Henri. 1967. *Economics and Information Theory*. Amsterdam: North-Holland Publishing Company.

Zwijnenburg, Jorrit, Sophie Bournot, and Federico Giovannelli. 2016. "Expert Group on Disparities in a National Accounts Framework: Results from the 2015 Exercise." OECD Statistics Working Papers, No. 2016/10. Paris: Organisation for Economic Co-operation and Development Publishing.