

BLS WORKING PAPERS



U.S. Department of Labor
U.S. Bureau of Labor Statistics
Office of Prices and Living Conditions

Distribution of U.S. Personal Consumption Expenditures for 2019: A Prototype Based on Consumer Expenditure Survey Data

Thesia Garner, Robert S. Martin, Brett Matsumoto, Scott Curtin
U.S. Bureau of Labor Statistics

Working Paper 557
August 8, 2022

Distribution of U.S. Personal Consumption Expenditures for 2019: A Prototype Based on Consumer Expenditure Survey Data¹

Thesia I. Garner, Robert S. Martin, Brett Matsumoto, Scott Curtin²

August 8, 2022

Abstract

We create a prototype procedure to allocate aggregate U.S. Personal Consumption Expenditures (PCE) across consumer units in the U.S. This allocation is based on data from the Consumer Expenditure Survey (CE). Since the CE collects data on out-of-pocket spending, to match PCE definitions, we assigned administrative and other household survey data for health expenditures to consumer units to reflect the full value of benefits received. Preliminary results for 2019 are that out of the total PCE (excluding expenditures by non-profits serving households), the bottom 10% accounted for 3.7%, while the top 10% accounted for 24.0%. The 90/10 ratio for equivalized PCE is 3.4, and the Gini coefficient is 0.28.

Key Words: Distribution, Personal Consumption Expenditures, Consumer Expenditure Survey

JEL Codes: D3 Distribution
E01 Measurement and Data on National Income and Product Accounts and Wealth & Environmental Accounts
E21 Consumption, Saving, Wealth

¹ We thank Dennis Fixler and Marina Gindelsky at the Bureau of Economic Analysis and Brian Adams at the Bureau of Labor Statistics for helpful comments and guidance. An earlier version of our analysis was presented at the June 15, 2022 meeting of the OECD Expert Group on Disparities within a National Accounts framework. This is a joint project between the Bureau of Labor Statistics and the Bureau of Economic Analysis and supports the work of the OECD Expert Group on Disparities within a National Accounts Framework.

² Division of Price and Index Number Research (Garner, Martin, Matsumoto), Division of Consumer Expenditure Surveys (Curtin), Bureau of Labor Statistics, 2 Massachusetts Ave., NE, Washington, DC 20212, USA. Emails: Garner.Thesia@bls.gov, Martin.Robert@bls.gov, Matsumoto.Brett@bls.gov, Curtin.Scott@bls.gov.

1. Introduction

The distribution of economic well-being has been the focus of researchers and policy makers for many years, with economic well-being most often defined in terms of income, expenditures or consumption, and wealth. Microeconomic analyses have focused on inequality in the distribution of these across the population, and poverty with respect to individuals and households. The primary data source for such studies are household surveys. In contrast, macroeconomic analyses traditionally have focused more on economic well-being defined in terms of national account aggregates or, for example, gross domestic product per capita. National account aggregates are based on statistical surveys (e.g., household surveys, wholesale and retail business surveys, economic censuses,) and government sources. With the production of the Stiglitz, et al. (2009), greater attention has focused the joint distribution of income, consumption, and wealth (ICW) to produce measures of the distribution of material living standards of households or individuals (see recommendation 3 therein). These distributions can be based on either household survey data alone or in combination with national account aggregates.³

This paper is an example of the latter in that we distribute U.S. Personal Consumption Expenditures (PCE) across households using micro data available from the U.S. Consumer Expenditure Survey (CE). In addition, this work supports the ongoing work of the Bureau of Labor Statistics to promote data quality through comparisons of CE data with data from other sources including the PCE.⁴ Our preliminary findings for 2019 are that out of the total PCE

³ For more information on the history of inquiry in this area, see Appendix B.

⁴ See <https://stats.bls.gov/cex/cecomparison.htm>

(excluding expenditures by non-profits serving households) the bottom 10% (as ranked by adult-equivalized PCE) accounted for 3.7%, while the top 10% accounted for 24.0%. The 90/10 ratio for equivalized PCE is 3.4, and the Gini coefficient is 0.28. Accounting for differences in scope between CE and PCE is important. For instance, if we only include out-of-pocket health care spending (as in the CE, and Gindelsky, 2020), then this component of PCE is much less equally distributed.

Most studies using household survey data examine income, expenditures/consumption, and wealth independently, in part because few household surveys collect data on all of these; the U.S. being an exception as the CE collects data on all three components.⁵ Two other exceptions are the U.S. Consumer Finances (SCF) and Panel Survey of Income Dynamics (PSID); however, neither collects the full range of expenditures available in the CE and neither for a continuous period of time.⁶ An example of the joint distribution of income, consumption, and wealth is a study by Fisher et al. (2022). However, rather than use only the CE for their study, these researchers used the SCF to study these components of economic well-being jointly; this is because the SCF, through its sample design, captures the top of the income and wealth distributions better than the CE.⁷ However, SCF expenditure data are very limited; as a solution,

⁵ The SCF can be used to measure income and wealth with limited information on expenditures for food, mortgage or rent, and the stock of vehicles (<https://www.federalreserve.gov/econres/scfindex.htm>). SCF data are collected periodically unlike the CE that has been collected on a continuous basis since 1980 and every 10 years for decades earlier. The PSID collects data on income and wealth, and expenditures but not expenditures or wealth for every wave (<https://psidonline.isr.umich.edu/Guide/default.aspx>). In the PSID expenditures are collected using global questions unlike the CE which provides much detail.

⁶ Historically another barrier, at least cross-nationally was the lack of an agreed-upon framework to examine these jointly even when the data were not from the same households. This changed with the development of the ICW Framework (OECD, 2013). See Appendix B

⁷ The Bureau of Labor Statistics is currently engaged in a project to impute wealth data when missing which is expected to improve data quality; however, concerns remain regarding missing wealth data from high wealth consumer units who do not respond to the CE at all.

Fisher et al. imputed a measure of consumption based on CE data to the SCF respondents. Most distributional studies using national accounts data have used per capita measures and sometimes decomposed aggregates by region or industry.⁸ However, a more recent focus of national accountants is the production of distributional accounts that integrates macro and micro data. Although household surveys can be used to show differences in economic well-being across households using various demographic variables, support for following a national accounts framework exists.

In contrast to household surveys, distributions of national accounts have a few advantages. To begin with, they are embedded in a larger framework that reflects a balance of accounts across sectors in the economy, and as such allows for what national accountants consider a comprehensive and consistent view of income, consumption, and wealth. Such a framework is useful, though from the household perspective, a comprehensive and consistent view of income, consumption, and wealth is reflected in the 2013 ICW Framework (OECD, 2013) and thus this reason is less salient than perhaps in the past. In addition, the distribution of national accounts approach allows for the alignment of household survey data with national accounts totals and thereby helps address concerns regarding reduced participation in household surveys and potential under-reporting. Nevertheless, just like household surveys, national accounts can also be subject to sampling and survey errors when based on survey data. And in addition, these can be subject to classification errors due to the estimation and

⁸ For U.S. NIPAs by region, see https://apps.bea.gov/iTable/index_regional.cfm and by industry, see https://apps.bea.gov/iTable/index_industry_gdplndy.cfm.

production or output to the personal sector and other sections of the economy.⁹ Finally, national accounts include items usually not covered in household surveys that are relevant to the economic well-being such as government provided health and education goods and services.¹⁰ This reason is a strong one as such benefits clearly support the economic well-being of households.

National accounts for the U.S. are produced by the Bureau of Economic Analysis (BEA) and have been used for distributional statistics fairly recently. Jorgenson and Slesnick (2014) developed a model of social welfare based on PCE. BEA researchers distributed Personal Income (Fixler et al., 2017 and 2020; Gindelsky, 2021) and PCE (Gindelsky, 2020) using household survey data. The BEA distributions are part of the larger OECD EG DNA to produce distributional national accounts. As yet, no study uses U.S. national accounts data reflecting the joint distribution of income, consumption, and wealth. However, research continues on the development of distributional statistics for each national accounts component that ultimately can be used together. This research is part of that endeavor.

This paper outlines a method developed by the U.S. Bureau of Labor Statistics (BLS) to distribute PCE as a counterpart to the BEA distribution of Personal Income (PI) product by Fixler et al. (2017, 2020) and Gindelsky (2021) who used Current Population Survey data, and as an update and improvement over the PCE distributional work of Gindelsky (2020) who used 2015 and 2016 CE data. The major contributions of the study, compared to the Gindelsky research

⁹ For example, because national accounts reflect a balancing across sectors, aggregates like those for household personal consumption expenditures are estimated as residuals from other accounts rather than expenditures specifically reflecting that for households.

¹⁰ In addition to the other references cited, see Fesseau and van de Ven (2014).

are two: (1) we augment CE Interview data with data from the Diary, and (2) we impute the value of health insurance from third party payers (e.g., employers, government) to CE respondents before scaling up to PCE totals. In contrast, Gindelsky only used Interview data, and allocated PCE health expenditures based on consumer unit (similar to a household) out-of-pocket spending. This latter approach resulted in an under-estimation of the health expenditures along the lower end of the distribution and an over-estimation at the upper end.

This study uses CE data to describe the distribution of PCE across people living in households, or in our case, consumer units. To do this, we need to not only match comparable categories across programs, but to also adjust less-comparable CE records so that they better match PCE definitions, delete records that are out-of-scope for PCE, and impute spending on items that are out-of-scope for CE. About 95% of CE spending can be sourced from the quarterly CE Interview survey, so we consider this sample to be the base of our analysis. For some categories, the CE Diary is either the sole (e.g., postage stamps, non-prescription drugs) or more reliable (e.g., clothing and footwear) source of information. We implement a statistical matching procedure to impute these expenditures for the CE Interview sample. Once we have our comparable and augmented CE data file, we produce PCE spending shares by deciles of equivalized total expenditures as well as distributional statistics (such as the ratio of the 90th to the 10th percentile and Gini coefficient) for total expenditures and total equivalized expenditures. In addition, we produce distributions by demographic subgroups. Our methodology prototype focuses on data for calendar year 2019.

An ultimate goal of this project, in coordination with BEA and the Federal Reserve Board, is to produce an integration of distributional statistics based on the U.S. National Income

and Product Accounts framework, in the style of other national statistical offices like the Australian Bureau of Statistics (2021). In addition, and as noted earlier, this work complements that of the BLS in assessing the quality of the CE Survey data in comparison to PCE.

2. Data and Methods

A first step in producing the distributional statistics was to adjust data from both the CE and PCE. This was necessary as there are differences in the CE and PCE that are based on the purpose of each and thereby the populations covered.¹¹ The CE is designed to collect out-of-pocket spending on goods and services by consumer units living in the U. S. in noninstitutional setting; the one exception is that students living in college or university housing (considered institutional settings) are included in the CE population. In contrast, the PCE reflects purchases of goods and services made by and on behalf of households and by non-profit institutions serving households (NPISHs). The PCE population includes all people who are “resident” in the U.S.¹² For this study we distribute PCE, as defined by Major Type of Product (NIPA Table 2.3.5) for 2019¹³ across consumer units using data from the CE. We restrict our analysis to components of the PCE that exclude expenditures by NPISHs to more closely match the types of expenditures and population covered by the CE; this follows the approach currently used by the

¹¹ See Passero et al. (2014) for a description of differences in concepts, measurement, and populations and a BLS approach to support comparisons of aggregate expenditures from the CE and PCE. For CE to PCE ratios based on aggregate expenditures for 2020, overall CE to PCE ratio is 0.62, while for comparable items it is 0.73 (see https://stats.bls.gov/cex/cecomparison/pce_profile.htm)

¹² PCE includes purchases for all people who are “resident” in the U.S. with “resident” defined as “those who are physically located in the U.S. and who have resided, or expect to reside, in this country for one year or more; U.S. government civilian and military personnel stationed abroad, regardless of the length of their assignments; and by U.S. residents who are traveling or working abroad for one year or less” (BEA 2019, p.2).

¹³ For the NIPA tables, see https://apps.bea.gov/iTable/index_nipa.cfm. In contrast, following the EG DNA methodology, Gindelsky (2020) allocated PCE aggregates using the Classification of Individual Consumer according to Purpose (COICOP) guidelines to produce a measure of final domestic consumption expenditures.

BLS to produce comparisons of CE and PCE aggregates.¹⁴ For this study, unlike for the BLS published CE-PCE comparisons, additional adjustments are also made; these are described in detail later in this section.

The PCE covers a broader set of goods and services than does the CE, with the focus on expenditures for what is produced within a year in line with the NIPA. The CE reflects consumer unit purchases from private businesses, and in addition, purchases from other consumer units through household-to-household transactions (e.g., for used cars). In contrast, the PCE reflects purchases from private businesses, costs incurred by NPISHs in providing goods and services on behalf of households, and purchases financed by third-party payers on behalf of households. Third-party payer expenditures include those for employer-paid health insurance, medical care financed through government programs, and financial services (such as banking services) that benefit households but for which they do not pay directly.¹⁵

The process of combining CE and PCE data to create distributions of PCE can be summarized as follows.

Process for Creating Distribution of PCE using CE Data

1. Map CE to PCE product categories.
2. Match CE Diary to Interview.

¹⁴ While the PCE includes expenditures made by resident households who normally live in the U.S. but who are temporarily abroad (see earlier footnote for the definition of “resident”), we did not adjust the PCE aggregates to deduct expenditures by these households for this analysis.

¹⁵ Neither the CE nor PCE includes in expenditures the value of in-kind transfers of goods and services such as government low-income food assistance (e.g., Supplemental Nutrition Program for Women, Infants, and Children), energy assistance (e.g., Low Income Home Energy Assistance Program, LIHEAP), and rental assistance (e.g., Housing Choice Voucher Program Section 8). However, these would be included as part of final national consumption expenditures when following the OECD EG DNA methodology.

3. Impute health expenditures.
4. Restrict CE sample to minimum of two interviews.
5. Annualize CE survey expenditures.
6. Scale CE estimates up to match PCE by major product.
7. Create adult-equivalized PCE and compute deciles and other statistics.

The following subsections describe many of the inherent issues and challenges in further detail.

2.A. CE Sample Selection

The CE comprises two surveys, an Interview and Diary, each with its own sample of consumer units. The Interview survey has a three-month recall period and is designed to cover major purchases (e.g., major appliances) and recurring items (e.g., rent, utilities). Consumer units are interviewed once every three months for up-to four consecutive quarters on a rolling basis. In contrast, the Diary is designed to cover more minor and frequently purchased items, such as detailed food, and consists of two consecutive one-week diaries. Data from both the Interview and Diary are integrated to produce published consumer unit means for total expenditures and subcomponents.¹⁶ Many studies on micro-level data use the Interview data alone since, with the use of global food questions, approximately 95% of CE defined total spending is represented.

The Interview and Diary overlap in coverage for some categories, though sometimes at differing levels of aggregation or frequency. As consumer unit-level expenditures are the

¹⁶ <https://stats.bls.gov/cex/>

building blocks of this analysis, we generally choose the Interview as the source when both surveys have coverage. We estimate total expenditures from the CE that are as close to PCE definitions as possible, and out of these we find about 95% can be represented by the Interview. This results in a remainder of about 5% for which the Diary is the only source or the most-reliable source by our judgment.¹⁷ Postage stamps and non-prescription drugs are examples of items that are only covered by the Diary. But expenditures for other items, like furnishings and durable household equipment, clothing and footwear, and motor vehicle insurance, are collected in both. For this latter set, we take data from the Diary because the CE aggregates for these are closer to the PCE aggregates. We start with the set of quarterly CE interviews that were collected from the first quarter of 2019 through the first quarter of 2021. The staggered timing of CE interviews, combined with sample attrition, means a relatively small sample give annual expenditure for calendar year 2019. Only about 300 consumer units participated in four consecutive interviews and had expenditure reference periods spanning January to December 2019. To form a larger sample, we include all consumer units whose expenditure reference periods started as early as November 2018 or ended as late as February 2020, provided they completed at least two quarterly interviews. This yields a sample of 7,171 consumer units. For consumer units who completed fewer than four quarters, we scale their expenditures up to represent one year of expenditure.

We also recalibrate the sampling weights to match 2019 counts and average demographic characteristics from the Current Population Survey (CPS). The BLS usually

¹⁷ These proportions are computed using all consumer units in 2019, before restricting to the set with two or more interviews.

calibrates CE data to 35 controls or known totals of people and households in the civilian non-institutional U.S. population, which it obtains from the CPS. These controls are counts based on certain geographic and demographic variables, with some being people counts and others being household counts. These include 14 age/race groups, 9 Census divisions, 9 urban Census divisions, the total number of Hispanics, the total number of owners, and the total number of households. The latter two controls are household counts, and CE converts these household counts to consumer unit counts by factoring in a relatively small adjustment.

This project required modifications to the normal CE calibration process. The CE normally treats each consumer unit independently in a quarterly weighting process, with a consumer unit defined as having participated in any one of its four quarterly interviews. However, we use interviews from the first quarter of 2019 through the first quarter of 2020, with a consumer unit defined as having participated during that period in either 2 of its 4 quarterly interviews, 3 of its 4 quarterly interviews, or all 4 of its quarterly interviews. This special way of defining a consumer unit means that we generate only one set of calibrated weights for each consumer unit, no matter if it completed two, three, or four quarterly interviews. Our calibrations averaged the 35 controls from the CPS data over the period that the data was collected, rather than over three months. Lastly, we adjusted the weights for the set of quarterly CE interviews used in this project since the data was a subsample from five quarters of CE's collected data. This adjustment ensures the sum of the weights match the number of consumer units in the population in 2019.

We start with the monthly expenditure (MTAB) files from the CE and map spending to PCE categories, as described in the following subsection. One limitation of using the CE to study

the distribution of expenditures is that expenditures at the top of the distribution are expected to be underreported in the CE (for example, see Fisher, et al 2022). After mapping CE expenditures to the PCE categories and imputing some missing categories, we scale expenditures so that population totals by major product type match PCE estimates from NIPA Table 2.3.5.

2.B. CE-PCE Concordance

As stated, we attempt to make the CE microdata match the definitions of the PCE categories. We start from a mapping of CE spending categories, called Universal Classification Codes (UCC) to PCE categories, called product types. BLS maintains the mapping for comparison purposes (Bureau of Labor Statistics, 2019). Some details on this mapping and imputations introduced in the subsections below.

2.B.1. Automobiles

The CE estimates for new and used automobiles subtract trade-in allowances from sales prices. To better reflect PCE aggregates, we add back in the trade-in values for new vehicle purchases. We then subtract these trade-in allowances from used vehicle spending under the assumption that vehicles traded in are eventually resold to the household sector, and as such, more closely align with the definition of new and used automobiles in the PCE. For this analysis, CE used vehicle expenditures do not include those that reflect household to household transactions for consistency with the PCE.

2.B.2. Health care

Table 1 (shown at the end of the report) summarizes the imputation procedures and data sources for health care. The comparison of CE and PCE health care expenditures is complicated by two issues. First, the scope of the CE and PCE differ. Before the CE can be mapped to the PCE, expenditures that are out of scope in the CE must be imputed. Second, the health care expenditure categories in the PCE are defined differently than in the CE, and adjustments must be made to the CE expenditures as some items that are classified as health expenditures in CE get mapped to non-health categories in the PCE.

The scope of the health care categories differs greatly between the CE and PCE. The CE only includes out-of-pocket spending for health insurance and health care goods and services, while the PCE in addition includes expenditures by employers and the government on behalf of consumers. For employer contributions to health insurance premiums, we use Medical Expenditure Panel Survey – Insurance Component (MEPS-IC) data to compute average employer contributions by plan type (self, plus one, and family). For plans purchased in the individual market, we add the average tax credit to the out-of-pocket premiums for individuals who report receiving a subsidy. No adjustments are made for those who purchase unsubsidized plans in the individual market. For Medicaid and Child Health Insurance Program (CHIP), we use the average expenditures per enrollee from the National Health Expenditure (NHE) Table 21. For Medicare, we compute the average expenditure less premiums per enrollee for traditional Medicare (parts A and/or B), Medicare Advantage (part C), and prescription drug coverage (part D) using Boards of Trustees (2020) and add these amounts to the out-of-pocket premiums included in the CE data. Other government programs (Indian Health Service, CHAMP-VA, and

Tricare) provide much of the care at government owned and operated facilities, which are out of scope for the PCE. We use the average expenditure for private providers per enrollee for these programs.¹⁸ The per enrollee amounts are multiplied by the number of covered members in the consumer unit. These initial results use a single national average amount for each of the programs. Future versions will impute at a more granular level based on household characteristics.¹⁹

Once the total premium expenditures are imputed for the CE, some adjustment must be made to CE health expenditures to match the category definitions in the PCE. With the imputations for employer and government contributions to health insurance, CE health insurance expenditures include the total premiums and are categorized as health expenditures along with out-of-pocket spending on medical goods and services. Health insurance in the PCE measures net premiums (premiums minus benefits) and is categorized as a financial service. As defined for the PCE, spending on medical goods and services includes the total amount paid to the providers (out-of-pocket payments plus any insurance reimbursement), while CE only includes out-of-pocket spending. Also, medical goods are categorized as non-durable goods in the PCE. The health care product category in the PCE includes non-insurance medical services only.

¹⁸ Care purchased from private providers is identified as a separate line item in the budgets for these programs. We ignore any out-of-pocket premiums reported in the CE for these programs to avoid double counting, which also allows us to avoid the issue of determining how much of the premiums go towards out-of-scope care.

¹⁹ In the EG DNA methodology, this type of adjustment is referred to as Method C where missing components are imputed according to exogenous data, e.g., sociodemographic data or in our case, reports of different types of health insurance coverage (see de Queljoe et al. forthcoming 2022). However, in our case, Method A, proportional scaling, is applied after the health expenditure imputations.

Before the CE can be mapped to PCE, health insurance premiums that go towards benefits need to be reassigned to medical goods and non-insurance medical services. The remaining premium amount reflects the net premiums and can be mapped to the health insurance category in the PCE. Thus, total health insurance premiums in the CE are allocated across three different PCE categories: pharmaceutical and other medical products (non-durables), health care services (health care), and net health insurance (financial services). Data from the NHE are used to reassign total premiums into these three categories based on the type of insurance.

2.C. Remaining CE-PCE discrepancies

For this study, we were able to adjust CE to account for some differences, but not all. We redefined CE expenditure categories to follow those of the PCE as much as possible. For example, for comparability to the PCE, rental equivalence as collected in the CE replaces out-of-pocket shelter expenditures as defined for CE publications,²⁰ and household-to-household transactions (e.g., for vehicles) are dropped from the CE expenditures. Not included in CE, but included in the PCE, are health and medical expenditures made on behalf of households – those made through employer provided health insurance and by the federal government. Thus, for our analysis, we supplement the CE with health care spending data from the Center for Medicare and Medicaid Services (CMS) NHE data, as well as the Medical Expenditure Panel Survey (MEPS) from the Agency for Health care Research and Quality. Gindelsky (2020) did not adjust CE data before scaling up CE reported health expenditures to PCE aggregates. We nor

²⁰ See the following for the definition of shelter for owner occupied housing:
<https://stats.bls.gov/cex/csxgloss.htm#housing>

Gindelsky adjusted CE data for education spending from NPISHs or the government; instead, we both allocated PCE education aggregates based on the distribution of education spending in the CE. While the value of home-produced food on farms is included in PCE food expenditures, the value of this is not included in the CE. The adjustment procedure we used to scale up CE data to match PCE aggregates accounts for this under-representation.

We were only able to partially adjust for the differences due to CE and PCE population coverage. Since the BEA produces aggregates for households and NPISHs, we were able to focus on household expenditures only. In other words, we excluded final consumption expenditures identified as the net expenses of NPISHs in Table 2.3.4; these are defined as gross output of NPISHs minus the receipt of sales of goods and services by NPISHs. This approach contrasts with that used by Gindelsky (2020) in her production of 2015 and 2016 distributional PCE. Gindelsky allocated net NPISH expenses to household expenditures by commodity type, for example, health.

An additional adjustment made by Gindelsky that is not made in this study is with reference to expenditures by resident and non-resident households. The PCE population includes among resident households, government employees and private employees living abroad; these households are not included in the CE population. However, included in the CE population but not the PCE are foreign students if they have a U.S. address or live in U.S. college or university housing. For the EG DNA work, Gindelsky adjusted PCE (defined as expenditures made by households and on behalf of households by NPISHs) by subtracting from this the difference in final consumption expenditures of resident households traveling abroad and those

living abroad from expenditures of non-resident households in the U.S.²¹ For our analysis, the PCE product aggregates that we allocate are not adjusted for resident households.

2.D. Statistical Matching to Impute Diary Expenditures

We implement a statistical matching procedure based on Hobijn et al. (2009) to impute the remaining 5% of PCE not sourced from the Interview Survey. Similar donors from the Diary sample provide the missing data for each Interview consumer unit, where similarity is determined by a model of monthly expenditure as a function of demographic characteristics.²² The model provides a convenient method of weighting a relatively large number of characteristics by doing so according to which linear combination most strongly predicts expenditures. We then use the predicted values to form measures of distance between Interview consumer units and prospective Diary donors. The only characteristic guaranteed to match between donor and recipient is quintile of the annual before-tax income distribution.²³ The matching procedure is many-to-one, as we draw four donor Diaries for each Interview in each month with replacement.²⁴

²¹ In other words, expenditures by those living or traveling abroad are not counted in Gindelsky's measure of final domestic consumption expenditure. The adjustment is the difference in the sum of expenditures abroad by U.S. residents (line 144 in NIPA Table 2.4.5U) and foreign travel by U.S. residents (line 330) and the sum of personal remittances in-kind to nonresidents (line 147) and expenditures in the U.S. by nonresidents (line 334).

²² A similar procedure is being used in ongoing research to create household-weighted Consumer Price Indexes (Martin, 2022).

²³ The Diary samples are small enough, particularly on a monthly basis, that conditioning on multiple characteristics quickly leads to empty cells. See Hobijn, et al. (2009) for more discussion.

²⁴ This is in contrast with the one-to-one "optimal transport" method that Blanchet, et al. (2022) uses to match CPS observations with public-use tax data. A many-to-one match is more appropriate for our purposes because there are many more Interview observations than Diary observations in a given month.

First, we stratify both Interview and Diary consumer unit samples by quintiles of annual before-tax income. The rankings of income are done for each survey. For each expenditure reference month t and quintile q , we use the Interview sample to estimate the regression

$$y_{ht} = \mathbf{x}_{ht}\boldsymbol{\beta}_{qt} + u_{ht},$$

where y_{ht} is logged expenditure of consumer unit h , u_{ht} is an error term, and \mathbf{x}_{ht} include Census region, urban/rural, age, race, sex, and education of the reference person, consumer unit size, and the prior year's annual before-tax income.²⁵ For the model, the measure of expenditure we use is the total monthly expenditure of the Interview household after mapping to PCE product definitions, but before imputations for employer or publicly provided health care.²⁶ We use the least squares estimator weighted by the CE sampling weight, `finlwt21`. A total of 80 models are estimated—one for each of five quintiles over 16 months (Nov. 2018 to Feb. 2020). On average, variation in characteristics explains about 30% of variation in logged expenditures across observations within the same month and income quintile.²⁷

Let $\hat{\boldsymbol{\beta}}_{qt}$ be the slope estimate for quintile q in month t . As household characteristics are available and comparably defined in both surveys, we calculate predicted values $\hat{y}_{ht} = \mathbf{x}_{ht}\hat{\boldsymbol{\beta}}_{qt}$

²⁵ These demographic variables technically pertain to the collection quarter or some other reference period. For instance, in the first interview, income represents the 12 months prior to the interview date, and this value is repeated for the second and third interview records. In the fourth interview, income is collected again for the 12 months prior to the interview date. We implicitly assume the demographic variables are representative of characteristics from the expenditure reference months.

²⁶ Alternatively, it might be attractive to use the Diary sample to estimate Diary expenditures as a function of demographic characteristics, as we intend to impute these expenditures for the Interview sample. However, we find that characteristics explain relatively little variation in Diary expenditures, perhaps due to the short (week-long) recall period.

²⁷ Income quintile itself has explanatory power when looking over the entire Interview samples. For reference, regressing total expenditures on indicators for income quintiles yields R-squared values that average about 0.4.

for each Diary and Interview observation. For a given Interview observation h and Diary observation k , the distance metric is defined as

$$\delta_t(h, k) = |\hat{y}_{ht} - \hat{y}_{kt}|.$$

Within each month and income quintile, we calculate $\delta_t(h, k)$ for all $\{h, k\}$ pairs. Then for each Interview observation h , we randomly select (with replacement) four k from the twenty smallest $\delta_t(h, k)$ out of all the Diary observations from the same month and income quintile.

The random component is intended to ensure a more even distribution of matches across Diary observations. The detailed set of expenditures (after adjustments to match PCE definitions) of the donor Diary is then assigned to the recipient Interview. As one donor Diary is intended to represent one quarter of one month of expenditure, but Diaries correspond to a one-week reference period, the donor Diary expenditures are scaled by 13/12. This process is repeated for each Interview observation, for each month it is in the sample. Since the Interview sample is much larger than the Diary on a per-month basis, draws are performed with replacement. As a result, each Diary is matched with several Interviews.

2.E. Scaling Estimates to Match PCE

After allocating CE to PCE categories and imputing missing items, CE aggregates differ from PCE aggregates for most categories. Table 2 describes the extent to which CE and PCE differ after adjustments to health-related allocations and imputations. For example, for all categories, the CE/PCE ratio is 0.71 after imputing health expenditures like Medicare, Medicaid, and employer contributions. For Durable Goods, it is 0.56, for Non-Durable Goods, it is 0.62, and for Services it is 0.76. Prior to these imputations, the ratio is only 0.58.

To allocate PCE expenditures across the distribution of equivalized total expenditures, as defined above, a next step is needed. CE expenditures are scaled up to consumer unit-level PCE estimates so that major product group totals match BEA's published estimates. This scaling is referred to as a proportional adjustment and reflects the allocation of the gap between CE and PCE totals in proportion to the underlying micro data.²⁸ For example, if aggregate PCE health care expenditures equal \$2,458 billion and consumer unit X had 0.000005% of total CE health care expenditures (after imputations for health insurance and allocations to PCE categories), consumer unit X is assigned $0.000005\% * \$2,458$ billion as its total health care expenditure.

2.F. Equivalence Scales

To create the deciles across which to distribute the PCE expenditures, we rank people based on their consumer unit's total expenditures equivalized by the square root of family size as the equivalence scale. In creating the ranking (as a precursor to creating quantile groupings), consumer units are weighted by the product of consumer unit size and their sampling weight (finlwt21). Total expenditures are defined as the sum of aggregate expenditures based on PCE major product type after all the processing described in Section 2.B to match PCE definitions and impute spending not covered by CE (e.g., health care), and after scaling up expenditures to PCE totals as described in Section 2.E. For consumer units whose reported consumer unit size varied across collection quarters, we use the mean of consumer unit size to produce the number of adult equivalents in each consumer unit.

²⁸ This is referred to as the Method A adjustment of micro data to macro totals. The EG DNA recommends this method be used when macro and micro data only show relatively small gaps (see de Queljoe et al. forthcoming 2022).

3. Results

Table 3 and Table 4, as well as Figure 1 through Figure 4 (given at the end of the document), describe our main results concerning the distribution of PCE. Table 3 is analogous to Table 1 of BEA's Distribution of Personal Income report for 2019.²⁹ By major product type, this table gives total PCE, and the share accounted for by each decile of the distribution of equivalized expenditure. Out of total PCE (excluding expenditures by non-profits serving households) in 2019, we estimate only 3.7% was accounted for by the bottom 10%, while the top 10% (the 90th percentile and above) accounted for 24.0%. The top 20% accounted for 38.9% of PCE, roughly equivalent to the bottom 60% (38.7%). The distribution of PCE is more equal than the distribution of PI, where the top 10% account for 36.1% and the bottom 10% account for 1.8% (Bureau of Economic Analysis, 2021b). Spending shares for services follow a similar pattern as overall PCE, with the bottom 10% accounting for 3.5% and the top 10% accounting for 24.2%. Spending on Durable goods is weighted more towards the higher deciles, with the top 10% making 37.7% of PCE versus only 1.5% for the bottom 10%. Spending on Non-Durables is slightly more even, with the top 10% accounting for 16.2%, versus 5.4% for the bottom 10%.

Table 4 presents some distributional statistics for PCE and equivalized PCE, similar to those presented with Table 3 of BEA's Distribution of Personal Income report (Bureau of Economic Analysis, 2021b). Statistics for equivalized PCE are weighted by a population weight equaling the product of the sampling weight and the consumer unit size.³⁰ The mean and

²⁹ See <https://www.bea.gov/data/special-topics/distribution-of-personal-income>

³⁰ In the Appendix, we re-compute deciles of equivalized PCE using other weights to rank consumer units, as well as using other weights to compute statistics like the 90/10 ratio. The alternative weights are weights `finlwt21` alone and `finlwt21` times the square root of consumer unit size. We find only a minor quantitative impact.

median equivalized PCE were \$69,064 and \$60,246, respectively, while the mean and median (non-equivalized) PCE were \$105,947 and \$91,505, respectively. We also compute five inequality indicators—the ratio of the 90th to the 10th percentiles (3.4), the Gini coefficient (0.28), the Theil index (0.14), the Log-Deviance (0.13), and the coefficient of variation (CV, 0.17). These again reflect a more equal distribution of PCE than for PI, where the Gini for 2019 was 0.44 (Bureau of Economic Analysis, 2021b). Figure 1 presents the shares of equivalized PCE accounted for by quintiles and other percentiles of equivalized expenditure. For example, the top 1% accounts for 4.2% of equivalized PCE, and the 80th to the 99th percentiles account for 33.1%. On the other hand, the bottom 20% accounts for 9.2%, while the 20th to the 40th percentiles accounts for 13.6%.

Figure 2 shows the Lorenz curve for equivalized PCE, corresponding to the Gini index of 0.28 as shown in Table 4. Figure 3 shows different concentration curves for PCE product categories. The Lorenz and concentration curves are based on population (i.e., person) weighting. The convexities of the curves indicate the degrees of inequality reflected in the distributions, with the 45-degree line corresponding to perfect equality. Health Care Services is the most equally distributed, followed closely by Nondurables, while Durable Goods is the least equally distributed.

Figure 4 describes the composition of PCE for the five quintiles of the equivalized expenditure distribution. For instance, out of the bottom quintile's total PCE, 29.0% was spent on Nondurable Goods, whereas this figure for the top quintile is only 16.3%. Lower quintiles also spend a higher share on health care services (23.2% for the bottom quintile and 12.6% for the top quintile) and housing services (19.8% for the bottom quintile versus 18.4% for the top

quintile). However, because higher quintile households account for a greater share of total PCE, higher quintiles still account for a disproportionate amount of total Health Care and Nondurable Goods PCE. On the other hand, the lower quintiles spend a lower share on Durable Goods (4.3% for the bottom quintile versus 15.5% for the top quintile) and Other Services (23.7% for the bottom quintile versus 37.2% for the top quintile).

The results concerning health care warrant further discussion given the additional steps needed to estimate the distribution of PCE. Prior to allocation of health insurance premiums and imputations for employer-provided and government-provided coverage, the CE only accounted for about 5% of PCE for health care services. The missing health care expenditures are not likely to follow the same distribution as the out-of-pocket expenditures captured by the CE. For instance, we see in Figure 5, Panel a that the share of households with members who are Medicare recipients or have employer-sponsored health insurance generally increases as we move up the distribution of equivalized PCE. Only 9.8% and 25.8% of the bottom decile have Medicare and Employer-provided, respectively, versus 41.9% and 66.1% of the top decile. In contrast, the share with members who are Medicaid recipients decreases—31.8% for the bottom decile and 2.0% of the top decile. It is important to note that the imputed values of these programs are large enough to move many participants out of the lowest decile. Panel b of Figure 5 shows smaller, but nonnegligible proportions of consumer units which included members of other government health insurance programs, across particularly the lower expenditure deciles. This motivates our imputation of government and employer-provided health care benefits, without which CE's coverage of PCE health care spending is much lower. Without the imputation, our adjustment process (to match PCE totals) would implicitly assume

that the expenditures missing from CE are distributed identically to the non-missing health care expenditures; this is the approach used by Gindelsky (2020) whose results showed that health care expenditures are more concentrated among consumer units in the upper end of the expenditure distribution. However, Figure 6 implies this is not the case, as the proportion of imputed health care expenditures is significantly higher for the lowest decile (90.8%) than it is for the highest decile (60.6%). Figure 7 shows how distribution of Health Care Services is impacted significantly by the imputations, without which the distribution is much more unequal. The results would be even more extreme if we had not allocated CE out-of-pocket health insurance premiums between the PCE categories for medical goods and services, health care services, and net health insurance.

Finally, we discuss the joint distribution of PCE and several consumer unit characteristics. Figure 8 shows that the share of renters (defined as those paying rent, living rent-free, or living in student housing) and owner-occupiers by decile of equivalized PCE. Figure 9 shows the distribution of equivalized PCE across different attributes, including family composition, housing tenure, race, and education. Figure 10 compares the breakdown of membership in the quintiles of equivalized PCE for the same characteristics. The share of renters decreases with equivalized PCE, while the share of homeowners increases.³¹ Of the bottom decile, 76.8% are renters, while 25.3% are owner-occupiers. For the top decile, only 11.2% are renters, while 89.0% are owner-occupiers. This is important for considering how, for example, the CE/PCE ratio for owner equivalent rent affects the distributional estimates. From

³¹ Figure 5 and Figure 8 track shares of consumer units who ever met the criteria at some point during our 2019 sample period, so totals may exceed 100%.

Figure 9 and Figure 10, owners with a mortgage accounted for the largest share of equivalized PCE (49.4%) and were disproportionately likely to be in the top quintile (28.1%). Renters paying rent, in contrast, accounted for 23.8% of equivalized PCE and disproportionately fell in the bottom quintile (36.8%). The Other category in Figure 9 and Figure 10 includes students and those living rent-free, which make up a small proportion of equivalized PCE (1.1%), and are more likely to be in the bottom quintile (57.3%).

Figure 9 and Figure 10 explore further the relationship between PCE and other consumer unit characteristics as well. For example, consumer units comprising a married couple and additional members made up the largest share of equivalized PCE (45.8%) and were approximately evenly distributed across quintiles. On the other hand, consumer units with single parents accounted for 4.5% of equivalized PCE and such units were most likely (44.5%) to be in the bottom quintile. Consumer units with a reference person that identified as White, non-Hispanic account for 68.8% of equivalized PCE, versus 10.4% for Black, non-Hispanic, 13.6% for Hispanic of any race, and 7.2% for Other. Consumer units with reference persons identifying as White or Other were more likely to be in the top quintile (25.6% and 21.6%, respectively) than the bottom quintile (13.4% and 10.9%), respectively. In contrast, consumer units with reference persons identifying as Black or Hispanic were less likely to be in the top quintile (8.8% and 8.0%, respectively), and more likely to be in the bottom quintile (35.2% and 35.6%, respectively). As for education, consumer units whose reference person completed a bachelor's degree accounted for 26.9% of equivalized PCE, and those with a master's degree or higher accounted for 19.1%. These two groups were more likely to be in the top quintiles (31.2% and 47.6%, respectively), and unlikely to be in the bottom quintiles (7.6% and 3.6%). On the other

hand, consumer units whose reference person's highest degree was high school accounted for 18.6% of equivalized PCE. Only 8.4% of these consumer units were in the top quintile, while 27.6% were in the bottom quintile. Of consumer units whose reference person did not complete high school, 45.7% were in the bottom quintile of equivalized PCE.

4. Conclusions

We develop a prototype methodology to estimate the distribution of PCE across consumer units or individuals using CE data. This paper describes several challenges in reconciling different definitions across data programs, combining multiple surveys, and imputing missing expenditures. Ideally, we want to minimize these differences before scaling CE expenditures to match PCE product category totals. In terms of population, we can easily focus on the non-NPISH portion of PCE, but we would need to also understand better the BEA data used to produce the PCE in order to adjust PCE aggregates to omit purchases by residents living outside the U.S. As for concepts, we aim to match PCE classification as much as possible and include substantial imputation to account for spending on health care goods and services on behalf of consumers. However, we have not yet attempted to account for other missing expenditures, such as financial services which are free to households and education services. Specific data challenges for which we propose solutions include the sampling of consumer units in the CE survey not aligning with the calendar year, the CE Interview reference periods being quarters rather than years, and the out-of-pocket definition of CE. In addition, we create custom sampling weights to account for the fact we use a specific subsample of CE Interview consumer units.

The challenges noted above suggest many possible improvements and avenues for future research, some of which is ongoing. To begin with, we plan to incorporate more sophisticated imputations for Medicare, Medicaid, and private insurance (employer provided or subsidized individual coverage) which incorporates information on geography and corrects for underreporting of Medicaid. Furthermore, we have yet to include educational expenditures beyond what the consumer unit spends out-of-pocket, nor have we attempted to impute whether such expenditures are made for a consumer unit member's benefit or for a student outside the consumer unit.

Another unresolved issue recommended by the EG DNA is to adjust the CE data for possible under coverage of consumer units in the upper end of the expenditure or income distribution. While the CE sample is designed to include consumer units all along the expenditure and income distribution, it is expected that the expenditures of those at the top of these distributions could be under-represented due to survey sample non-response, item non-response, and/or item under-reporting. We could correct for underreporting of expenditures at the top of the distribution using the method of Fisher, et al (2022); as noted earlier, they use the Survey of Consumer Finances (SCF), which does a better job capturing the top of the distribution, to correct for the under-estimation of expenditures. Another option might be to use generalized Pareto curves in the style of Blanchet, et al. (2022). In a recent OECD study an adjustment was made to consumption expenditures that was based on the ratio of Pareto-adjusted income to non-Pareto-adjusted income (de Queljoe forthcoming 2022). Finally, we plan to explore alternative methods to impute missing quarters of expenditure for consumer units which completed only two or three interviews.

References

- Agency for Healthcare Research and Quality. 2021. *Medical Expenditure Panel Survey*. Accessed April 13, 2022. <https://www.meps.ahrq.gov/mepsweb/>.
- Australian Bureau of Statistics. 2021. *Australian National Accounts: Distribution of Household Income, Consumption and Wealth methodology*. June 18. Accessed July 12, 2022. <https://www.abs.gov.au/methodologies/australian-national-accounts-distribution-household-income-consumption-and-wealth-methodology/2019-20>.
- Blanchet, Thomas, Emmanuel Saez, and Gabriel Zucman. 2022. "Real-Time Inequality." Unpublished manuscript. <https://eml.berkeley.edu/~saez/BSZ2022.pdf>.
- Blanchet, Thomas, Thomas Piketty, and Juliette Fournier. 2022. "Generalized Pareto Curves: Theory and Applications." *Review of Income and Wealth* 68 (1): 263-288. doi:10.1111/roiw.12510.
- Boards Of Trustees, Federal Hospital Insurance and Federal Supplementary Medical Insurance Trust Funds. 2020. *2020 Annual Report of the Boards of Trustees of the Federal Hospital Insurance and Federal Supplementary Medical Insurance Trust Funds*. Washington, DC: Center for Medicare and Medicaid Services. <https://www.cms.gov/files/document/2020-medicare-trustees-report.pdf>.
- Bureau of Economic Analysis. 2021a. "Chapter 5: Personal Consumption Expenditures." In *Handbook of Methods*, by Bureau of Economic Analysis, 1-69. Washington, DC: Bureau of Economic Analysis. <https://www.bea.gov/resources/methodologies/nipa-handbook/pdf/chapter-05.pdf>.
- . 2021b. *Distribution of Personal Income*. December 15. Accessed May 14, 2021. <https://www.bea.gov/data/special-topics/distribution-of-personal-income>.
- . 2021c. *Table 2.3.5. Personal Consumption Expenditures by Major Type of Product*. April 30. Accessed May 13, 2021. https://apps.bea.gov/iTable/index_nipa.cfm.
- Bureau of Labor Statistics. 2019. *Comparing expenditures from the Consumer Expenditure Survey with the Personal Consumer Expenditures: Results of the CE/PCE Concordance*. Accessed May 9, 2022. <https://www.bls.gov/cex/cepceconcordance.htm>.
- Canberra Group. 2011. *Canberra Group Handbook on Household Income Statistics, 2nd Edition*. Geneva: UNECE. <https://unece.org/statistics/publications/canberra-group-handbook-household-income-statistics-2nd-edition>.
- Centers for Medicare and Medicaid Services. 2021. *National Health Expenditure Data*. Accessed July 29, 2021. <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData>.

- Coli, Alessandra, Radoslav Istatkov, Hakam Jayyousi, and Friderike, Tsigkas, Orestis Oehler. 2022. *Distributional national account estimates for household income and consumption: methodological issues and experimental results*. Luxembourg: Publications Office of the European Union. <https://ec.europa.eu/eurostat/web/products-statistical-working-papers/-/ks-tc-21-010>.
- de Queljoe, Matthew, Joseph Grilli, and Jorrit Zwignenburg. 2022 (forthcoming). *OECD Centralised Approach to Calculate Distributional Results in line with National Accounts Totals*. Paris: OECD.
- Fesseau, Maryse, and Maria L. Mattonetti. 2013. *Distributional Measures Across Household Groups in a National Accounts Framework*. Working Paper 53, Paris: OECD Publications. doi:<https://doi.org/10.1787/5k3wdjqr775f-en>.
- Fesseau, Maryse, and Peter van de Ven. 2014. "Measuring inequality in income and consumption in a national accounts framework." *OECD Statistics Brief* 19. <https://www.oecd.org/sdd/na/Measuring-inequality-in-income-and-consumption-in-a-national-accounts-framework.pdf>.
- Fesseau, Maryse, Florence Wolff, and Maria L. Mattonetti. 2013. *A Cross-country Comparison of Household Income, Consumption and Wealth between Micro Sources and National Accounts Aggregates*. Working Paper 52, Paris: OECD Publications. doi:<https://doi.org/10.1787/5k3wdjrn7mv-en>.
- Fisher, Jonathan D., David S. Johnson, Timothy M. Smeeding, and Jeffrey P. Thompson. 2022. "Inequality in 3-D: Income, Consumption, and Wealth." *Review of Income and Wealth* 68 (1): 16-42. doi:10.1111/roiw.12509.
- Fixler, Dennis, David Johnson, Andrew Craig, and Kevin Furlong. 2017. "A Consistent Data Series to Evaluate Growth and Inequality in the National Accounts." *Review of Income and Wealth* S437-S459. doi:<https://doi.org/10.1111/roiw.12324>.
- Fixler, Dennis, Marina Gindelsky, and David Johnson. 2020. *Measuring Inequality in the National Accounts*. BEA Working Paper 2020-3, Washington, DC: Bureau of Economic Analysis. https://www.bea.gov/system/files/papers/measuring-inequality-in-the-national-accounts_0.pdf.
- Gindelski, Marina. 2020. *United States: Household distributional results in line with national accounts, experimental statistics*. Accessed July 28, 2022. <https://www.oecd.org/sdd/na/household-distributional-results-in-line-with-national-accounts-experimental-statistics.htm>.
- Gindelsky, Marina. 2021. *Technical Document: An Updated Methodology for*. Washington, DC: Bureau of Economic Analysis. https://apps.bea.gov/data/special-topics/distribution-of-personal-income/technical_document.pdf.

- Hobijn, Bart, Kristin Mayer, Carter Stennis, and Giorgio Topa. 2009. "Household Inflation Experiences in the U.S.: A Comprehensive Approach." *Working Paper 2009-19*. Federal Reserve Bank of San Francisco.
<http://www.frbsf.org/publications/economics/papers/2009/wp09-19bk.pdf>.
- International Labour Organization. 2003. *Report II: Household Income and Expenditure Statistics*. Prepared for the 17th International Conference of Labour Statisticians, Geneva, 24 November-3 December 2003, Geneva: International Labour Office.
<https://www.ilo.org/public/english/bureau/stat/download/17thicls/r2hies.pdf>.
- Jorgenson, Dale W., and Daniel T. Slesnick. n.d. "Measuring Social Welfare in the U.S. National." In *Measuring Economic Sustainability and Progress*, by J. Steven Landefeld, Paul Schreyer Eds. Dale W. Jorgenson, 43-88. University of Chicago Press.
<http://www.nber.org/chapters/c12825>.
- Martin, Robert. 2022. "Democratic Aggregation: Issues and Implications for Consumer Price Indexes." Unpublished manuscript.
- OECD. 2021. *Household distributional results in line with national accounts, experimental statistics*. Accessed May 9, 2022. <https://www.oecd.org/sdd/na/household-distributional-results-in-line-with-national-accounts-experimental-statistics.htm>.
- . 2013. *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth*. Paris: OECD Publishing. doi:<https://doi.org/10.1787/9789264194830-en>.
- Passero, William, Thesia I. Garner, and Clinton McCully. 2014. "Understanding the Relationship: CE Survey and PCE." In *Improving the Measurement of Consumer Expenditures*, 181-203. University of Chicago Press. <http://www.nber.org/chapters/c12659>.
- Stiglitz, Joseph E., Amartya Sen, and Jean-Paul Fitoussi. 2009. "Report by the Commission on the Measurement of Economic Performance and Social Progress." www.stiglitz-sen-fitoussi.fr.
- Törmälehto, Veli-Matti. 2019. *Reconciliation of EU statistics on income and living conditions (EU-SILC) data and national accounts*. Luxembourg: Publications Office of the European Union. <https://ec.europa.eu/eurostat/en/web/products-statistical-working-papers/-/ks-tc-19-004>.
- United Nations. 1977. *Provisional Guidelines on Statistics of the Distribution of Income, Consumption and Accumulation of Households*. Studies in Methods Series M No. 61, New York: United Nations. <https://digitallibrary.un.org/record/479675?ln=en>.
- Zwijnenburg, Jorrit, Sophie Bournot, David Grahn, and Emmanuelle Guidetti. 2021. *Distribution of household income, consumption and saving in line with national accounts: Methodology and results from the 2020 collection round*. Paris: OECD.
<https://doi.org/10.1787/615c9eec-en>.

Tables

Table 1: Imputations of Health Care Expenditures

Type	For Whom	Value of Imputation	Source of data
<i>Public</i>			
Medicare	Assigned to consumer unit (CU) members enrolled in Medicare	National average benefits net of premiums	CMS 2020 Medicare Trustees Report (https://www.cms.gov/files/document/2020-medicare-trustees-report.pdf)
Medicaid	Assigned based on the number of CU members identified as participating	National average expenditures per enrollee	CMS National Health Expenditures (NHE)
CHIP	Assigned based on the number of CU members identified as participating	National average expenditures per enrollee	CMS National Health Expenditures (NHE)
Other public (VA, Tricare, and other military, and IHS)	Assigned based on the number of CU members identified as participating	National average expenditures on private care	Agency Budgets
<i>Private</i>			
Employer provided	Assigned to each CU reporting employer coverage based on the number of individuals covered: self only (number covered = 1), self plus one (number covered = 2), family (number covered >2)	National average premium paid by employer for health insurance by policy type (self only, self plus one, family)	MEPS-IC (https://datatools.ahrq.gov/meps-ic)
Individual	Assigned to each CU reporting individual coverage and receiving a subsidy	National average premium tax credit among those who receive a credit	CMS (https://www.cms.gov/files/document/2016-2021-1h-effectuated-enrollment-tables.xlsx)

Notes: For private and Medicare, the imputed amounts are added to out-of-pocket premiums reported in the CE.

For other public plans, only care provided in non-government facilities is in scope for the PCE. The VA and IHS budgets include these amounts as a separate line item. The Department of Defense submits a separate budget for care purchased from private providers for Tricare.

Table 2: CE Coverage of PCE Total After Adjustments and Imputations, 2019

Category	Total (billions)	Unadj.	CE/PCE	
			After allocation	After allocation and imputation
PCE less final CE of nonprofit institutions serving households	\$13,989	0.58	0.58	0.71
Durable goods	\$1,513	0.56	0.56	0.56
Motor vehicles and parts	\$515	0.95	0.95	0.95
Furnishings and durable household equipment	\$360	0.56	0.56	0.56
Recreational goods and vehicles	\$421	0.24	0.24	0.24
Other durable goods	\$218	0.27	0.27	0.27
Nondurable goods	\$2,966	0.54	0.55	0.62
Food and beverages purchased for off-premises consumption	\$1,031	0.77	0.77	0.77
Clothing and footwear	\$398	0.47	0.47	0.47
Gasoline and other energy goods	\$338	0.87	0.87	0.87
Other nondurable goods	\$1,199	0.26	0.30	0.47
Household consumption expenditures (for services)	\$9,510	0.59	0.58	0.76
Housing and utilities	\$2,572	1.12	1.12	1.12
Health care	\$2,458	0.05	0.20	0.83
Transportation services	\$495	0.47	0.47	0.47
Recreation services	\$583	0.40	0.40	0.40
Food services and accommodations	\$1,008	0.52	0.52	0.52
Financial services and insurance	\$1,172	0.63	0.28	0.41
Other services	\$1,222	0.69	0.69	0.69

Note: Category totals from the BEA NIPA Table 2.3.5 (June 29, 2022 release).

Table 3: PCE Total and Share by Decile of Equivalized Expenditure, 2019

Category	Total (bil.)	Spending share by decile of equivalized expenditure									
		0-10%	10-20%	20-30%	30-40%	40-50%	50-60%	60-70%	70-80%	80-90%	90-100%
PCE less final CE of nonprofit institutions service households	\$13,989	3.7%	4.9%	6.0%	6.9%	8.2%	9.1%	10.3%	12.1%	14.9%	24.0%
Durable goods	\$1,513	1.5%	1.9%	2.6%	4.2%	5.5%	6.6%	9.6%	12.4%	17.9%	37.7%
Motor vehicles and parts	\$515	0.4%	0.8%	1.4%	3.1%	4.8%	4.4%	10.1%	13.9%	20.4%	40.8%
Furnishings and durable household equipment	\$360	2.3%	3.2%	3.8%	5.1%	7.2%	8.5%	10.7%	12.9%	17.2%	29.1%
Recreational goods and vehicles	\$421	2.0%	2.2%	3.1%	4.8%	5.0%	6.8%	8.7%	10.3%	16.4%	40.8%
Other durable goods	\$218	1.8%	2.0%	2.6%	4.0%	5.6%	8.4%	8.4%	12.5%	16.3%	38.3%
Nondurable goods	\$2,966	5.4%	6.3%	7.5%	8.4%	9.4%	10.2%	10.9%	12.1%	13.6%	16.2%
Food and beverages purch. for off-premises consumption	\$1,031	6.6%	7.3%	8.1%	8.5%	9.4%	9.9%	10.3%	11.7%	12.8%	15.4%
Clothing and footwear	\$398	5.0%	5.6%	6.6%	7.5%	9.2%	9.8%	11.0%	12.2%	14.6%	18.6%
Gasoline and other energy goods	\$338	5.5%	6.2%	7.9%	9.0%	9.6%	10.6%	10.7%	12.0%	13.4%	15.2%
Other nondurable goods	\$1,199	4.4%	5.8%	7.2%	8.4%	9.4%	10.5%	11.5%	12.3%	14.1%	16.5%
Household consumption expenditures (for services)	\$9,510	3.5%	4.9%	6.1%	6.9%	8.2%	9.1%	10.2%	12.0%	14.8%	24.2%
Housing and utilities	\$2,572	4.2%	5.0%	6.1%	6.7%	8.2%	8.6%	10.2%	12.0%	14.4%	24.5%
Health care	\$2,458	4.2%	7.0%	8.3%	9.2%	9.9%	10.4%	11.1%	11.8%	13.0%	15.0%
Transportation services	\$495	1.9%	3.0%	4.0%	5.0%	7.3%	8.0%	10.0%	13.1%	17.5%	30.3%
Recreation services	\$583	2.0%	2.9%	4.3%	5.2%	6.6%	8.4%	9.7%	12.6%	17.2%	31.0%
Food services and accommodations	\$1,008	3.0%	3.4%	4.6%	5.7%	6.9%	8.2%	9.2%	12.6%	16.4%	29.9%
Financial services and insurance	\$1,172	3.1%	4.9%	6.3%	7.3%	8.8%	9.7%	10.9%	12.5%	15.4%	21.0%
Other services	\$1,222	2.6%	3.3%	4.5%	5.0%	6.4%	8.2%	9.0%	10.9%	15.2%	34.9%

Note: Category totals from the BEA NIPA Table 2.3.5 (June 29, 2022 release).

Table 4: PCE Statistics, 2019

Statistic	Estimate
Equivalized total mean	\$69,064
Equivalized total median	\$60,246
Equivalized total: 90/10	3.37
Equivalized total: Gini index	0.28
Equivalized total: Thiel index	0.14
Equivalized total: Log-Deviance	0.13
Equivalized total: CV	0.17
Total mean	\$105,947
Total median	\$91,505

Figures

Figure 1: Share of Total Equivalized Personal Consumption Expenditure, 2019

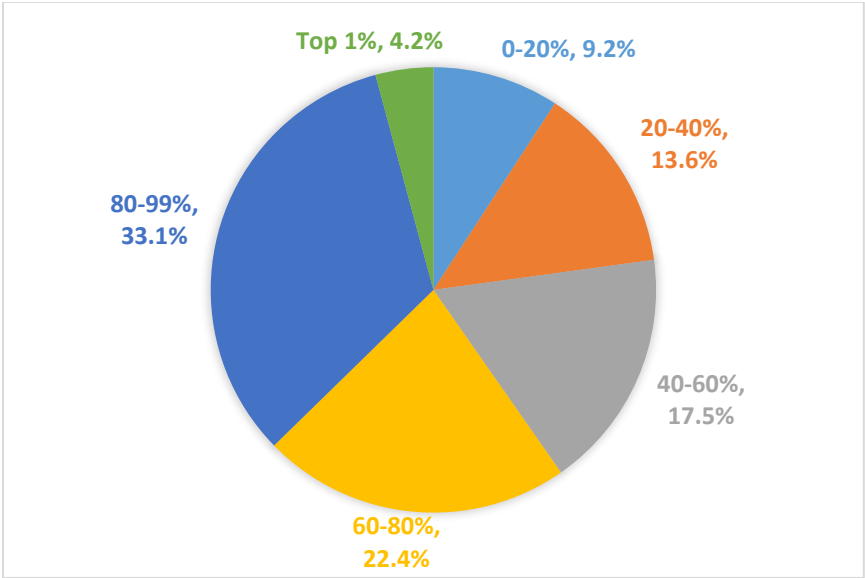


Figure 2: Lorenz Curve of PCE, 2019

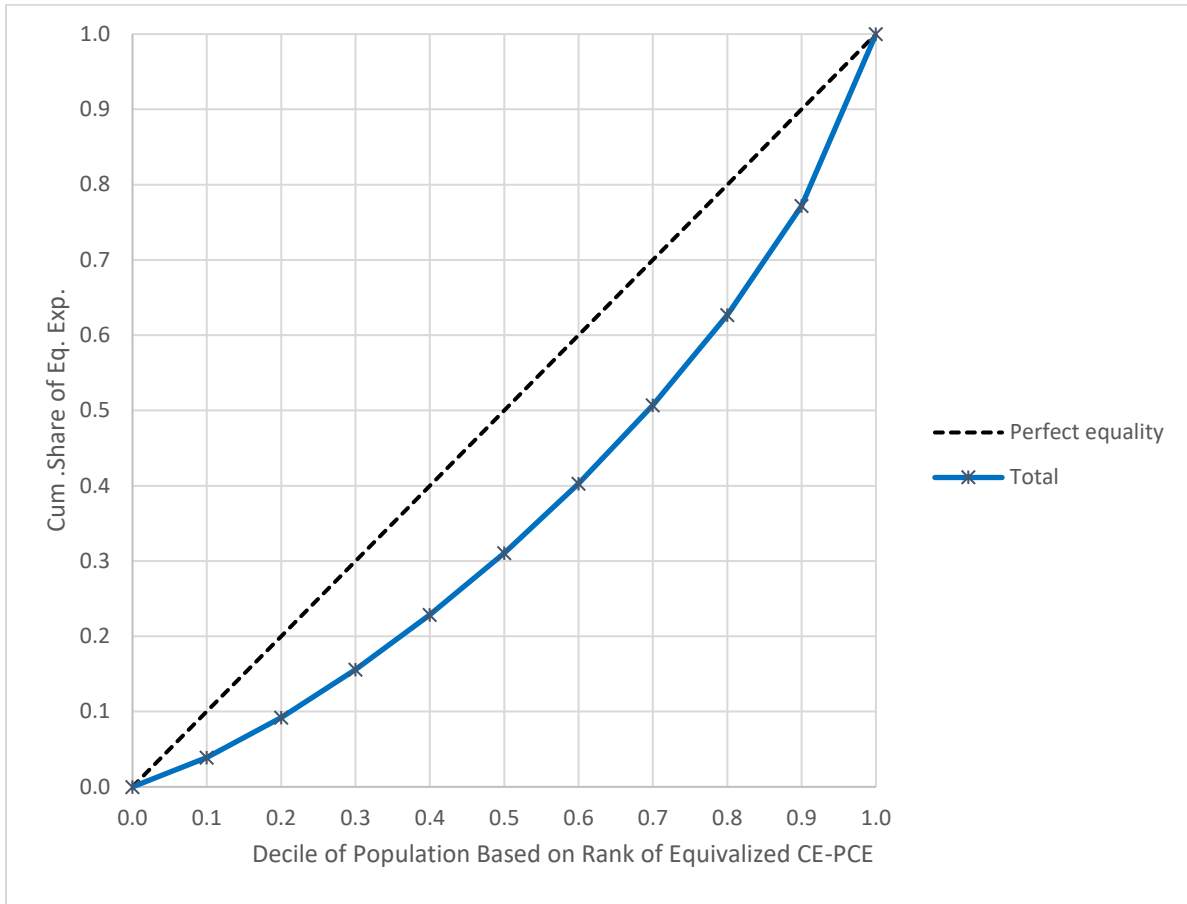


Figure 3: Concentration Curves of PCE Products, 2019

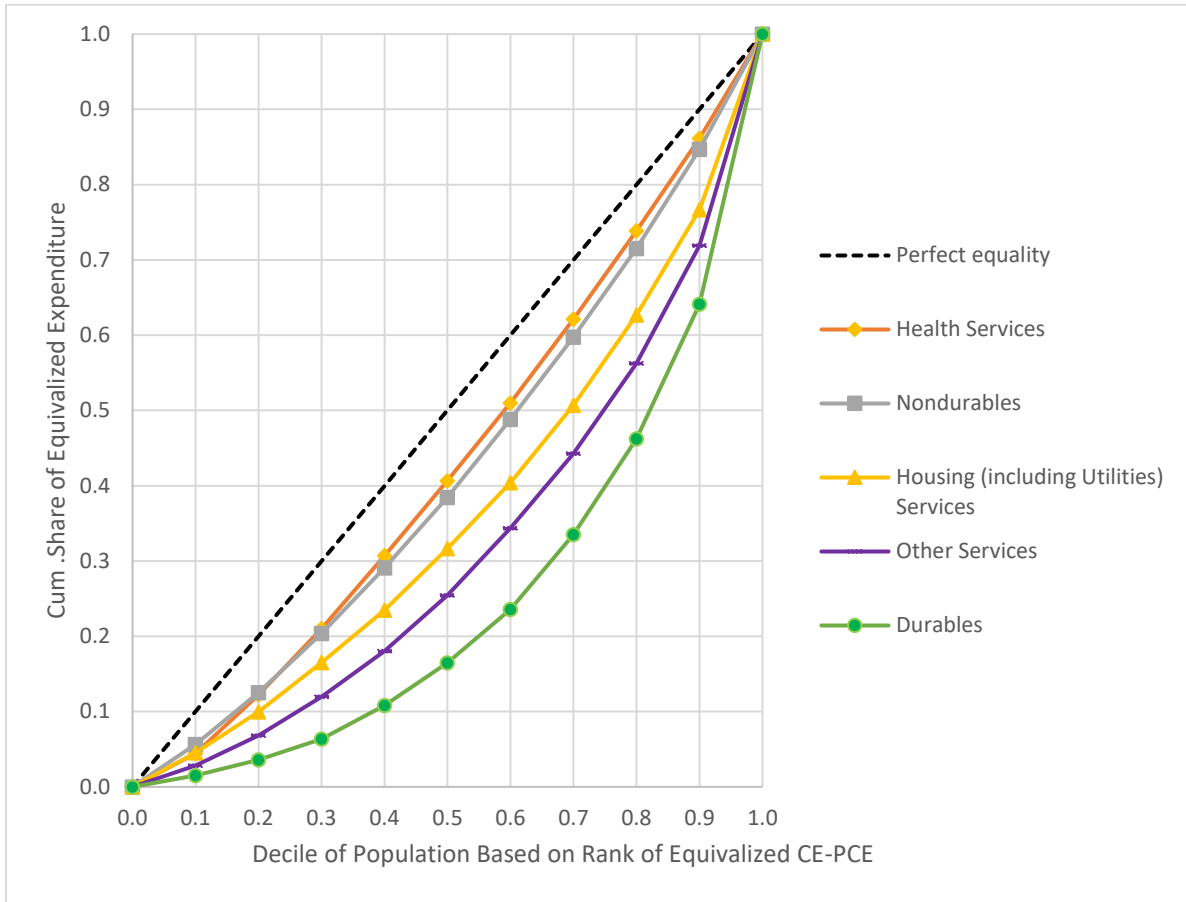


Figure 4: CE-PCE: Composition of Personal Consumption Expenditure, 2019

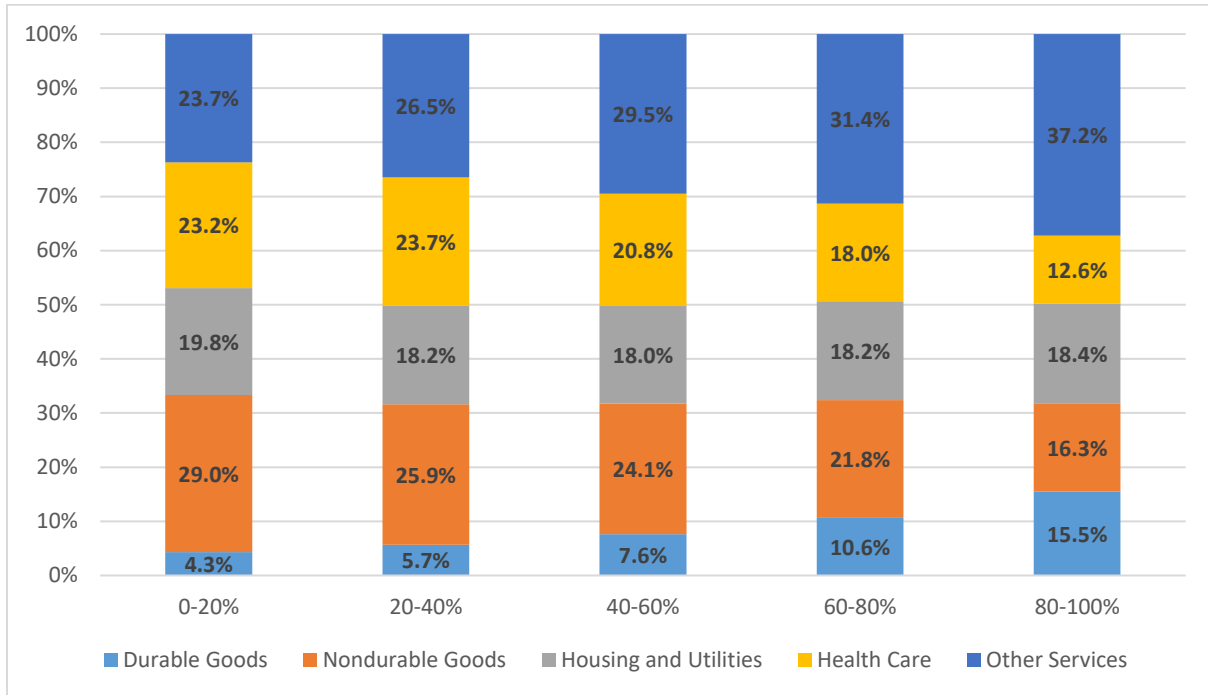
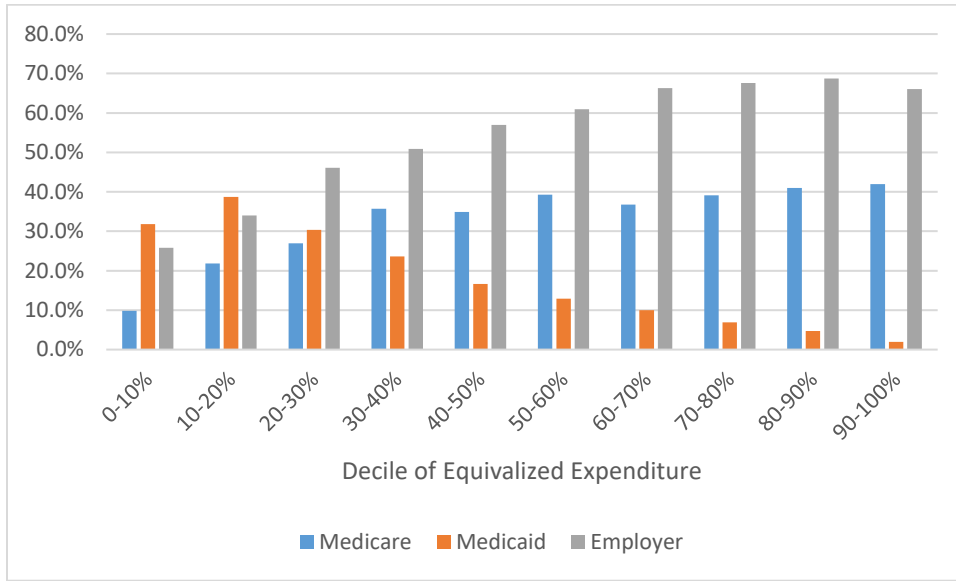
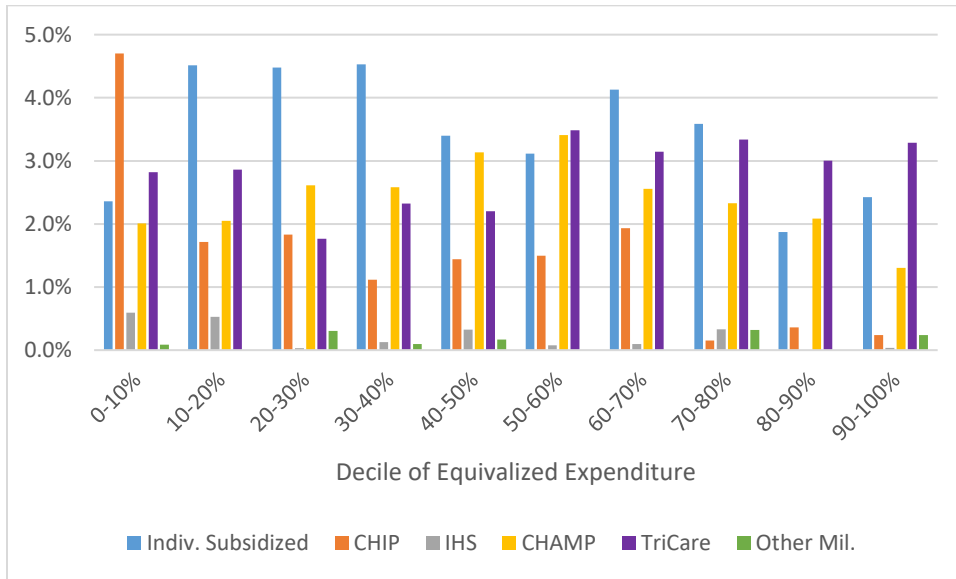


Figure 5: Share of Consumer Units with Different Health Insurance Types, 2019



Panel a. Medicare, Medicaid, and Employer-sponsored



Panel b. Individual subsidized and other government insurance

Note: Share is of consumer units with a member who met criteria at some point in 2019, so sums may exceed 100%.

Figure 6: Imputed Proportion of CE Spending on PCE-defined Health Care Services (before Scaling Up), 2019

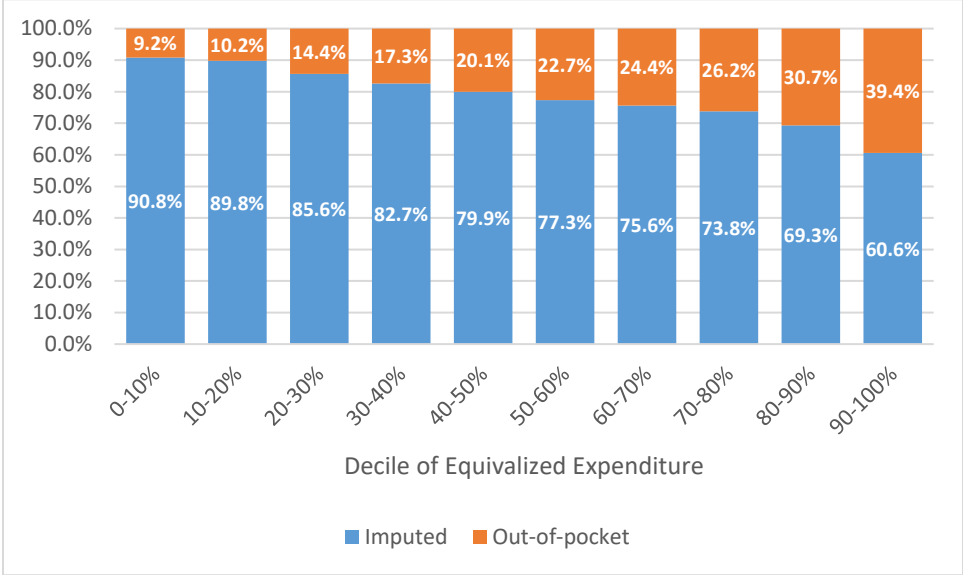


Figure 7: Concentration Curves of PCE Health Care Services, 2019

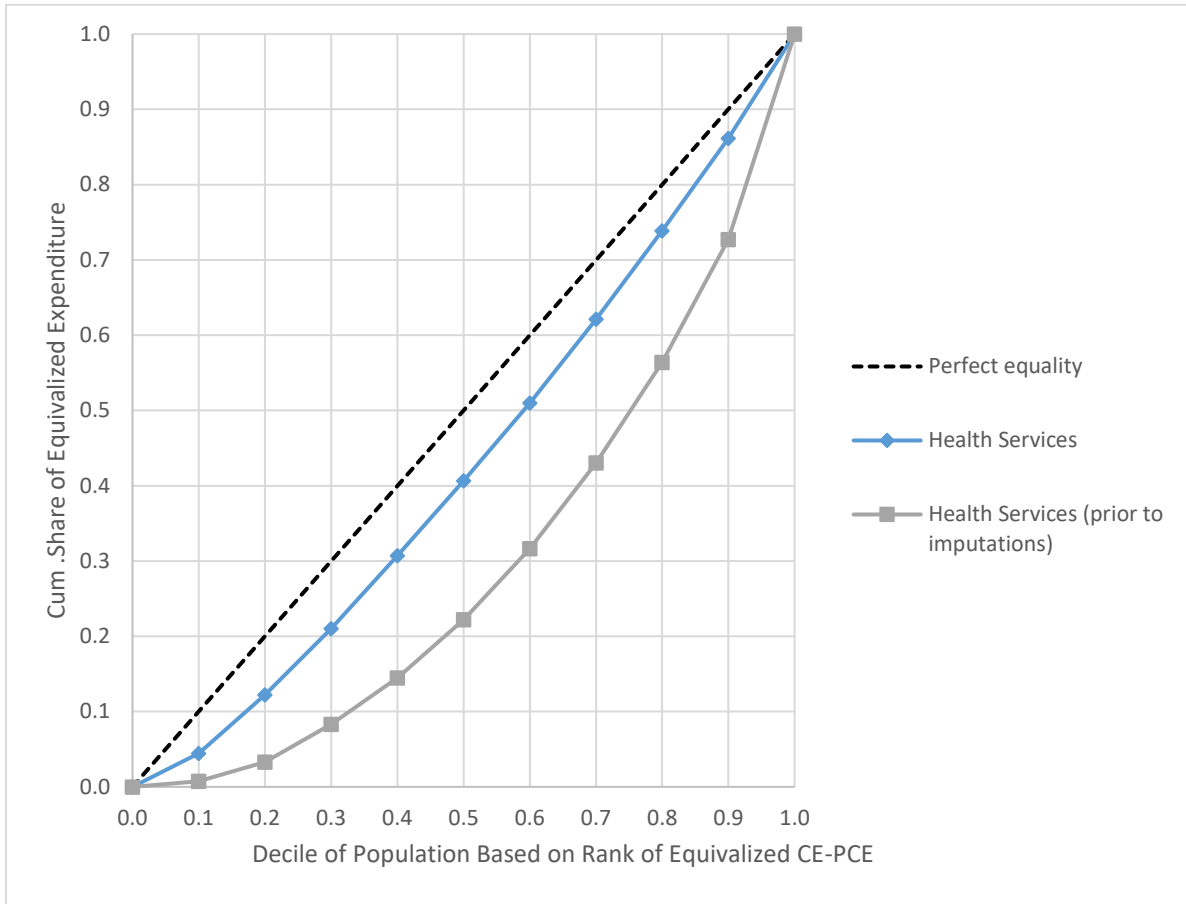
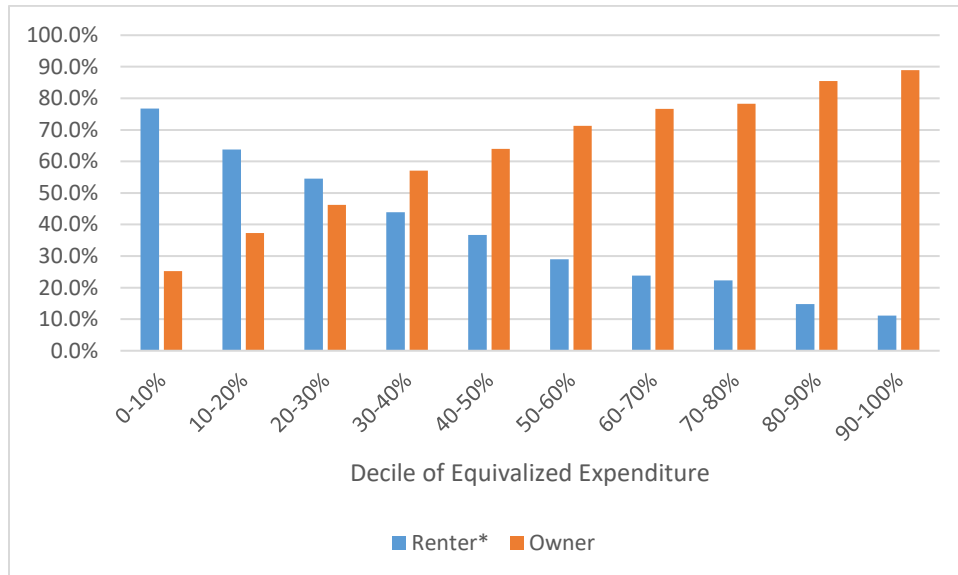
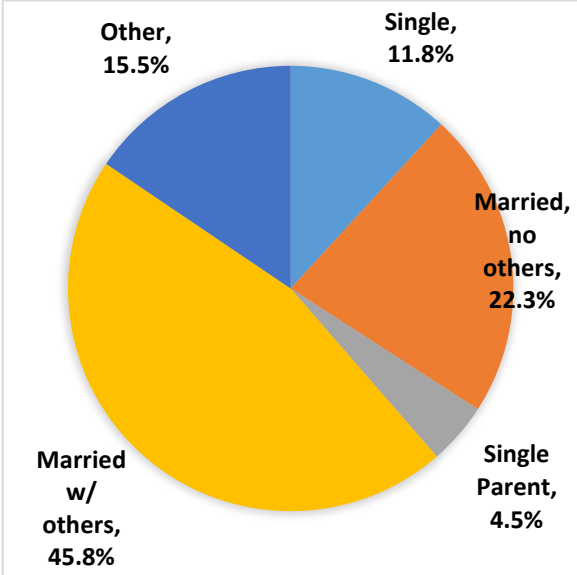


Figure 8: Share of Households that are Renters or Homeowners, 2019

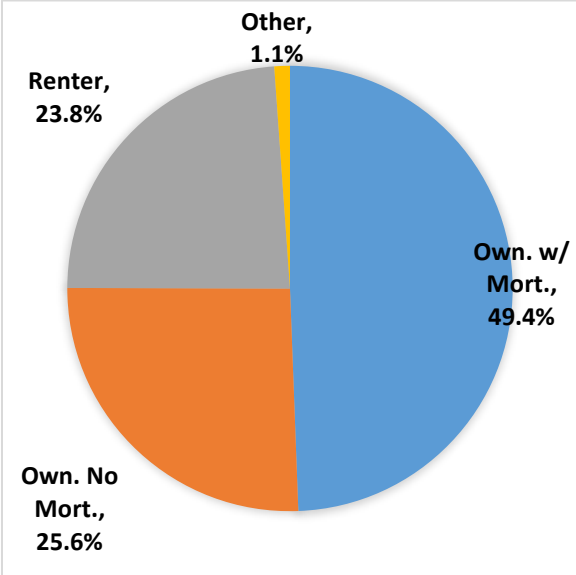


Note: Share is of consumer units who met criteria at some point in 2019, so sums may exceed 100%. *Renter group includes student housing and those living without payment of rent.

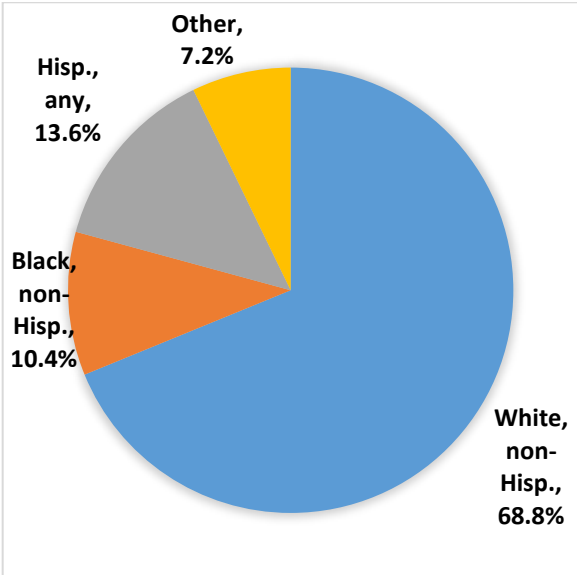
Figure 9: Distribution of Equivalized PCE by Demographic Characteristics, 2019



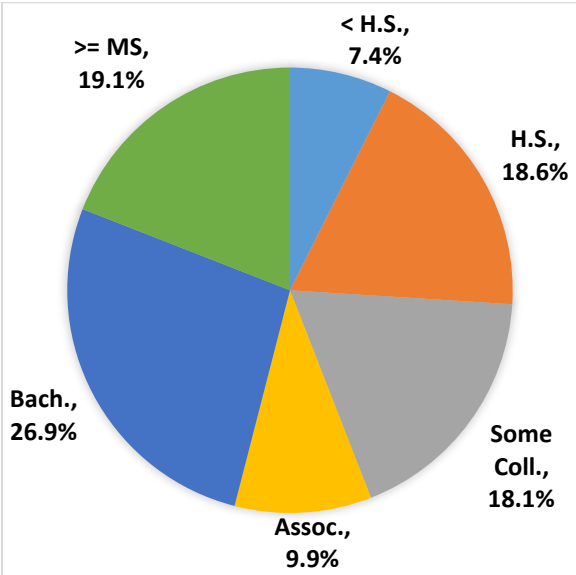
Panel a: Consumer Unit Composition



Panel b: Housing Tenure

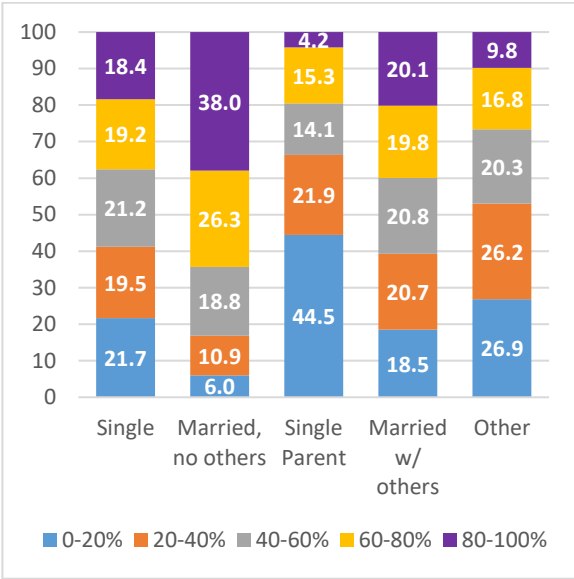


Panel c: Race/Ethnicity of CU Ref. Person

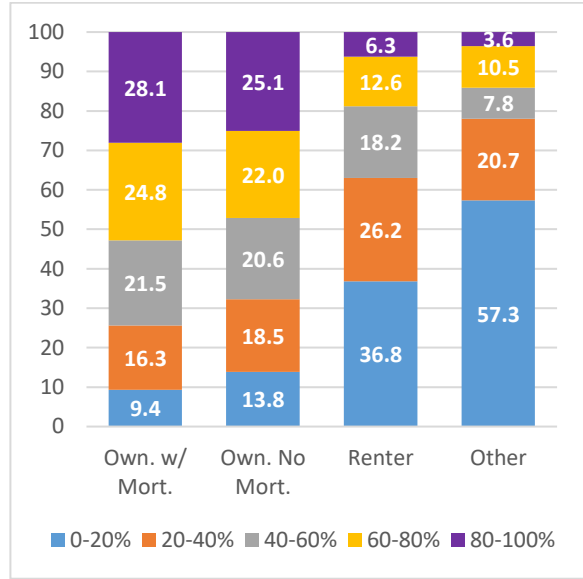


Panel d: Education of CU Ref. Person

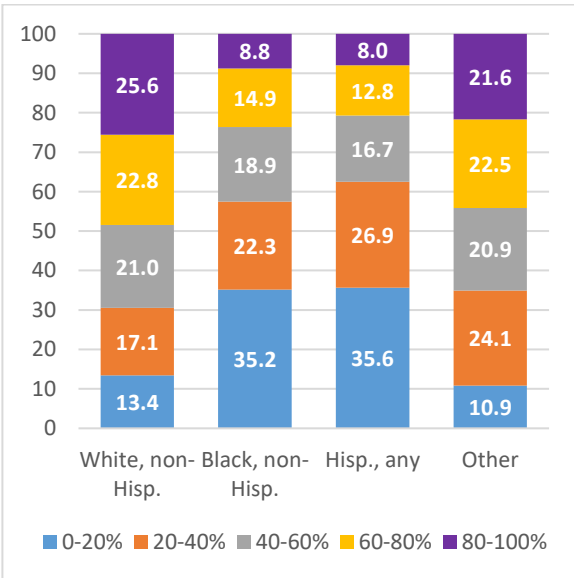
Figure 10: Quintile of Equalized PCE Membership by Demographic Characteristics, 2019



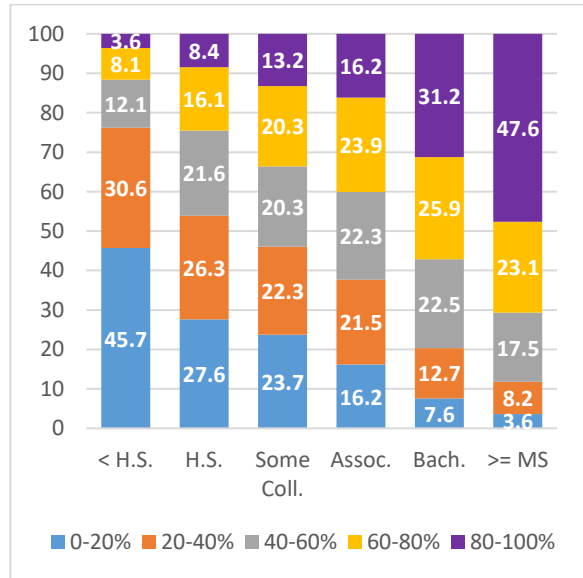
Panel a: Consumer Unit Composition



Panel b: Housing Tenure



Panel c: Race/Ethnicity of CU Ref. Person



Panel d: Education of CU Ref. Person

Appendix A. Impact of Weighting

Researchers have used various weighting options for the production of distributional statistics (e.g., percentile ratios like the 90/10 and aggregate indexes like the Gini). For example, de Queljoe et al. (forthcoming 2022) and Gindelsky (2020) weighted equivalized expenditures using household/consumer unit weights (`finlwt21` for the CE) following the EG DNA guidelines. The reasoning for selecting this option is that households, not people, are the focus of income and consumption expenditure national accounts. In contrast, the OCED ICW expert group (OECD 2013) recommended that the use of person-weighting when producing distributional statistics based on an economic unit (e.g., household, consumer unit, family) or any other statistical unit that combines individuals. The assumption with this weighting option is that equivalized household income or expenditures are available to each person in the household. In the case of the CE, the person-weight would be defined as `finlwt21` times consumer unit size. Another option is to weight by the number of “consumption units”; this is the number of equivalized adults times the household or consumer unit weight. The number of equivalent units is a function of the equalize scale chosen. For example, with the equivalence scale being the square root of consumer unit, a 2-adult household is represented by 1.41 equivalent units (times the consumer unit population weight). When conducting distributional statistics, this latter weighting option ensures that the sum of equivalized expenditures within a decile, for example, equals the total expenditure for the decile.

In our main results, we create a new weight variable equal to `finlwt21` (the CE sampling weight) times the family size. We use this new weight for computing statistics concerning

equivalized PCE (e.g., from Table 4), and for the creation of decile and other quantile groups (e.g., the 10%-20% group).³² The alternative weights are weights `finlwt21` alone and `finlwt21` times the square root of family size. We find only a minor quantitative impact, which is shown in Table 5, Figure 11, and Figure 12 below. From Figure 11, we see that including family size in the weight lowers the higher deciles slightly more than the lower deciles. From Figure 12, we see that weighting by family size lowers the number of consumer units in the bottom deciles and increases the number of consumer units in the top deciles. As for the statistics compared in Table 5, the effects of including family size in the weight are a slight decrease in the mean and median and a slight decrease in the share accounted for by the top decile and percentiles. Summary measures like the Gini index are constant (to two decimal places), while the 90/10 ratio is slightly lower (3.42 vs. 3.46).

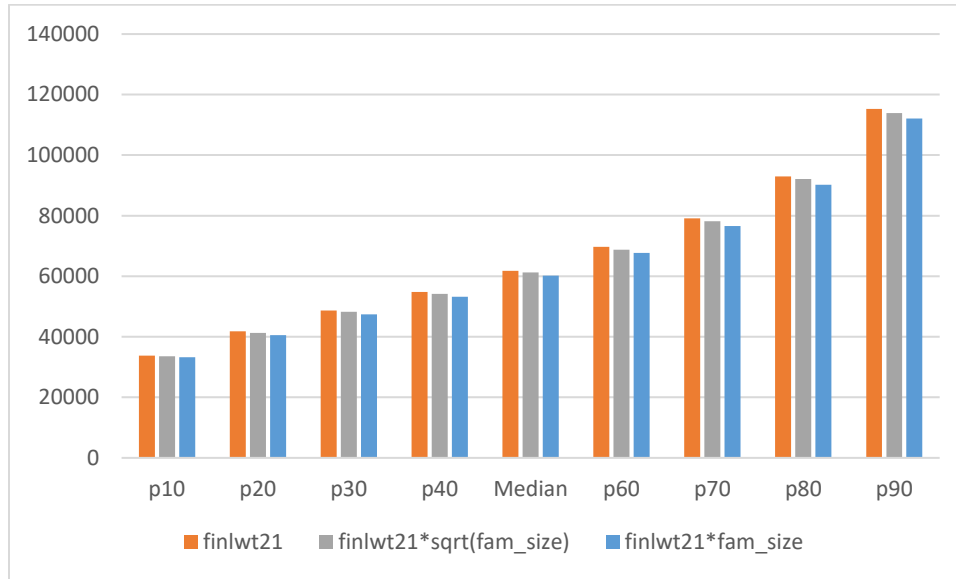
³² Quantile groups are created using PROC UNIVARIATE in SAS to compute the quantiles themselves, and then assigning group membership based on consumer unit-level equivalized PCE relative to these quantiles.

Table 5: Impact of Weighting on Quantiles and Distributional Statistics

Statistic / Weight	finlwt21	finlwt21*sqrt(fam_size)	finlwt21*fam_size
Equivalentized total mean	\$71,135	\$70,359	\$69,064
Equivalentized total median	\$61,818	\$61,275	\$60,246
Equivalentized total: 0-20% share	9.0%	9.1%	9.2%
Equivalentized total: 20-40% share	13.6%	13.6%	13.6%
Equivalentized total: 40-60% share	17.4%	17.4%	17.5%
Equivalentized total: 60-80% share	22.4%	22.4%	22.4%
Equivalentized total: 80-100% share	37.5%	37.4%	37.3%
Equivalentized total: 80-99% share	33.2%	33.2%	33.1%
Equivalentized total: Top 1% share	4.3%	4.2%	4.2%
Equivalentized total: Top 5% share	14.0%	13.9%	13.9%
Equivalentized total: 90/10	3.42	3.40	3.37
Equivalentized total: Gini index	0.28	0.28	0.28
Equivalentized total: Thiel index	0.14	0.14	0.14
Equivalentized total: Log-Deviance	0.13	0.13	0.13
Equivalentized total: CV	0.18	0.17	0.17

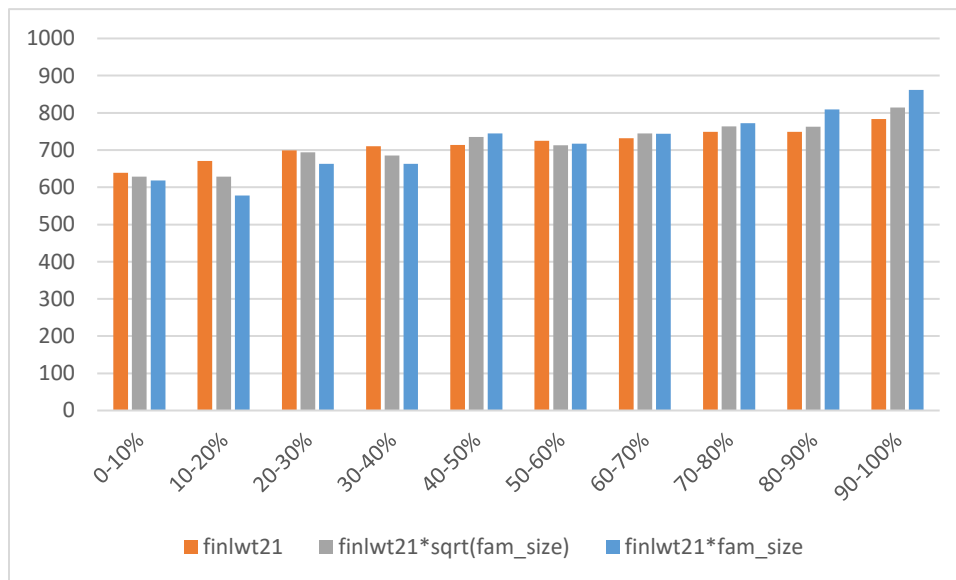
Notes: Columns track changes in distributional statistics when different weights are used to both to create quantile groups and weight the statistics. Finlwt21 is the sampling weight used by CE, adjusted for our particular selection of interviews. “fam_size” refers to the number of members of the consumer unit. “finlwt21*fam_size” corresponds to the weight used in the main results.

Figure 11: Deciles of Equivalized PCE using Different Weights



Notes: Finlwt21 is the sampling weight used by CE, adjusted for our particular selection of interviews. “fam_size” refers to the number of members of the consumer unit. “finlwt21*fam_size” corresponds to the weight used in the main results.

Figure 12: Number of consumer units in Deciles of Equivalized PCE using Different Weights



Notes: Finlwt21 is the sampling weight used by CE, adjusted for our particular selection of interviews. “fam_size” refers to the number of members of the consumer unit. “finlwt21*fam_size” corresponds to the weight used in the main results.

Appendix B. History of Research on Joint Distributions of Income, Consumption, and Wealth

The need for distributional statistics that reflect the well-being of households in terms of the income, consumption, and wealth (ICW) jointly was recognized as early as 1960. For example, following the 14th session of the United Nations Statistical Commission (UN) a system of distributional statistics that covered all three components of economic well-being focused on households was to be developed by the UN. The result was a set of guidelines published in 1977 (United Nations, 1977). This was followed by additional meetings and guidelines, for example, those published by the ILO (2003) and the Canberra Group (2011). While the UN primarily was tied to the System of National Accounts (SNA), the more recent guidelines were more focused on distributions based on household level data with most detail on income. The need to consider the distribution of income, consumption, and wealth jointly – with an emphasis on households -- was also highlighted by the 2009 Commission on the Measurement of Economic Performance and Social Progress in their report (Stiglitz et al., 2009). An aim of the Commission “to identify the limits of GDP as an indicator of economic performance and social progress.”³³ Motivated by the Stiglitz et al. report, two OECD expert groups were launched, one focused on micro-based measures of distribution and another on macro-based measures.³⁴

³³ Stiglitz et al., 2009, p. 7.

³⁴ The national accounts methodology to produce distributional accounts differs from that based solely on household surveys as outlined in the 2013 ICW Framework. For example, the ICW Framework focuses on the value of consumption goods and services used or paid for by the household to directly meet its needs; expenditures for goods and services given to other households are considered transfers and not consumption expenditures. Such transfers are ignored in the SNA as the focus is on total spending for all households, not just the household to meet its own needs. The SNA (and NIPA PCE) includes the purchase price of durables including vehicles while the ICW Framework includes the flow of services from these in consumption expenditures. Also included in ICW Framework consumption expenditures are those for goods and services received through barter, and the value of home-produced domestic services. To derive the value of final consumption in the ICW Framework, social transfer

For micro-based measures, an OECD expert group was convened to develop a harmonized ICW statistical framework focused on micro statistics of households and to conduct work that focuses on cross-national comparisons. This group is referred to as the EG ICW. An early product of the EG ICW was the *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth* (2013). Cross-national work based on this framework continues as part of joint work with Eurostat.³⁵ For macro-based measures, another OECD task force, starting in 2011, has been developing a methodology for the production of distributional national accounts for income, consumption, and wealth. This group is referred to as the expert group on disparities in national accounts framework (EG DNA) (see Zwijnenburg et al., 2021). Cross-national work based on the methodology being developed by this group also continues.³⁶

in-kind (STiK) provided by the government and NPISHs serving households are added to consumption expenditures. In contrast, the provision of goods and services by NPISHs is included in consumption expenditures in the national accounts; see distributional accounts produced by Australia as an example (ABS 2021).

³⁵ For updates on the Eurostat work on the joint ICW distribution, see

<https://ec.europa.eu/eurostat/web/experimental-statistics/income-consumption-and-wealth>

³⁶ In 2011, the OECD and Eurostat started a joint Expert Group on Disparities in a National Accounts framework (EG DNA) with the aim to develop such a methodology and guide countries in their production of distribution accounts for income, consumption, and wealth; this work continues today. For example, see the OECD compilation of experimental statistics of the distribution of income and consumption or consumption expenditures following a national accounts framework: https://stats.oecd.org/Index.aspx?DataSetCode=EGDNA_PUBLIC