Use with caution: interpreting Consumer Expenditure income group data

By Veri Crain and Taylor J. Wilson

Economic well-being has been an issue relevant to American public discourse for some time and has become a topic of particular interest among journalists, policymakers, and ordinary American families. The Consumer Expenditure Survey (CE) is often used as a tool to analyze the effects of economic well-being. Though the Census Bureau independently collects and processes income data—the CE is the only Federal survey that collects information on income, expenditures, and associated demographic characteristics from U.S. consumers. As a result, CE data are useful for answering many salient questions related to economic well-being. However, CE data are also complex and should be used with an understanding of the limitations of these data.

This **Beyond the Numbers** article examines assumptions users often make regarding how the CE measures household wealth, by providing examples of the nuances in the data and composition of five household groupings. The examples provide a clearer snapshot of economic well-being found in the income quintiles. The article uses tabulations of households by quintiles of income before taxes that many researchers use to identify the "poor" and the "rich" in the CE data. These terms are subjective and potentially pejorative. Another common set of terms for these types of analyses are "low-income" and "high-income." These are also subjective terms, and they are imprecise for communicating the concept of economic well-being, which involves other measures of wealth (e.g. stocks, bonds, cash assets, etc.). Misunderstandings arise potentially when researchers use terms like "rich" and "poor" to describe households rather than focus solely on income measures. So, what are the common assumptions made about the economic well-being of the "poor" and "rich" when using CE data to compare income groups?

Quintile construction

One must first consider how the quintiles of income are constructed in the CE. The quintiles are based on pre-tax income, which is composed of wages and salaries, self-employment income, retirement fund cash benefits including Social Security, investment income, public assistance cash payments including Supplemental Nutrition Assistance Program benefits, unemployment and workers' compensation, veteran's cash benefits, regular contributions for support, meals as pay, and rent as pay.[1] The sample households are ordered by income and an income rank is computed for each household (based on a cumulative distribution of each households' population weight). They are then assigned to 1 of 5 quintiles along the income distribution and are identified as lowest, second, third, fourth, and highest 20 percent.

When measuring the well-being of the households in each quintile, analysts should know what CE excludes from the definition of income as well as what it includes. The CE excludes noncash benefits such as housing subsidies, women, infants and children (WIC) benefits, school lunch subsidies, gifts received, waived tuition or fees, and professional services or medical benefits received from any source outside the household are not accounted for or identified anywhere in the survey and are not elements of pre-tax income. Owned assets and accumulated wealth are not considered pre-tax income, but these data are collected as part of the CE to provide a more complete picture of household wealth.

Discussion and examples

To understand the lowest 20 percent in relation to economic well-being, consider other socio-demographic factors such as the household composition of those who populate the quintile. The household characteristics data found on the first page of each published CE annual table provide information that can help the user understand the expenditure and income data.

Selected characteristics. Table 1 below uses selected items from the "consumer unit characteristics" section of table 1 from the 2015 Consumer Expenditure Survey report. The estimates cited are from the publically available CE microdata.[2] The microdata allow users to manipulate the data at the household level and to select or create additional variables for analysis.

Table 1. Quintiles of income before taxes: average annual expenditures and select characteristics, Consumer Expenditure Survey, 2015

Category	All Consumer Units	Income quintile				
		Lowest 20 percent	Second 20 percent	Third 20 percent	Fourth 20 percent	Highest 20 percent
Number of consumer units (in thousands)	128,427	25,627	25,562	25,700	25,730	25,773
Lower income limit	(1)	(1)	$19,572	$37,638	$62,587	$103,057
Income before taxes (mean)	$69,627	$10,916	$28,343	$49,606	$80,813	$177,851
Age of reference person	50.5	53.3	53.6	49.3	47.9	48.7
Average number in consumer unit						
People	2.5	1.7	2.2	2.5	2.9	3.1
Children under 18	0.6	0.3	0.5	0.6	0.7	0.8
Adults 65 and older	0.4	0.4	0.5	0.4	0.3	0.2
Earners	1.3	0.5	0.8	1.3	1.8	2.0
Housing tenure						
Homeowner	62	38	53	60	73	87
Homeowner with mortgage	35	11	18	32	52	65
Homeowner without mortgage	27	27	36	28	22	22
Fruits and vegetables	769	483	621	643	863	1,233
Fresh milk and cream	140	94	119	125	155	205

(1) not applicable
Source: U.S. Bureau of Labor Statistics.

The 2015 CE public-use microdata and CE quintiles of income before taxes tables were used to develop the examples to examine the demographic composition of the lowest and highest quintiles. Thirteen percent of the reference persons in the lowest 20 percent quintile identified themselves as college students.[3] An additional 33 percent of the reference persons reported that they were retirees. Accordingly, 46 percent of the households ranked in the lowest quintile are at stages in the life-cycle where one would assume lower incomes might prevail. Homeownership data found on table 1 provides more clues to economic well-being. Table 1 shows that in the highest 20 percent quintile, 87 percent of households report as homeowners, with 22 percent of the total reporting as homeowners without a mortgage. The lowest quintile reports 38 percent as homeowners, with 27 percent of the total reporting as homeowners without a mortgage. Given the proportion of college students and retirees, who are assigned to the lowest 20 percent, one can assume that not all households within the lowest quintile represent the usual image of a "poor" household.

Homeownership is an example for which the use of "poor" versus "rich" labels can be misleading. In this case, the "poor" top the "rich" by 5 percentage points when it comes to homeownership without a mortgage. Homeownership, especially without a mortgage, is an indication of wealth that would not be reflected in the income category that establishes the basis for the quintiles.

Cash contributions. Investigation of the "cash contributions" expenditure category found in the published data table (table 1101) provides further evidence that one must use caution when using CE-defined income as a measure of well-being. For example, the 2015 detailed unpublished version of the quintiles of income before taxes

table (available by request) reports that the lowest 20 percent had an average expenditure for gifts of stocks, bonds, and mutual funds to persons outside of the household of $140.51. Less than 0.7 of 1 percent of these households reported such gifts, indicating that a small number of households categorized in this quintile collectively gave $3.61 billion in stocks, bonds, and mutual funds to persons outside of the household. An example from the CE microdata reveals that one household reported just over $10,000 in pre-tax income and a contribution of stocks, bonds, and mutual funds worth $150,000. (See table 2.) This occurred only once in 2015, and not in any other year. One should examine patterns over time before drawing conclusions from just 1 year of the CE quintile table.

Number of earners. Table 1 lists the average number of earners per household in each quintile. These data show a higher number of earners corresponds to higher average household income for each successive quintile. The lowest 20 percent has an average of 0.5 earners per household, while the highest 20 percent has an average of 2.0 earners per household. Because income quintiles are based on household income, one might assume that households with more earners would report more income than households that have fewer earners. We can use the CE microdata to capture a clearer picture of earners. Chart 1 illustrates the 2015 frequency count of the number of earners in the lowest 20 percent and the highest 20 percent.
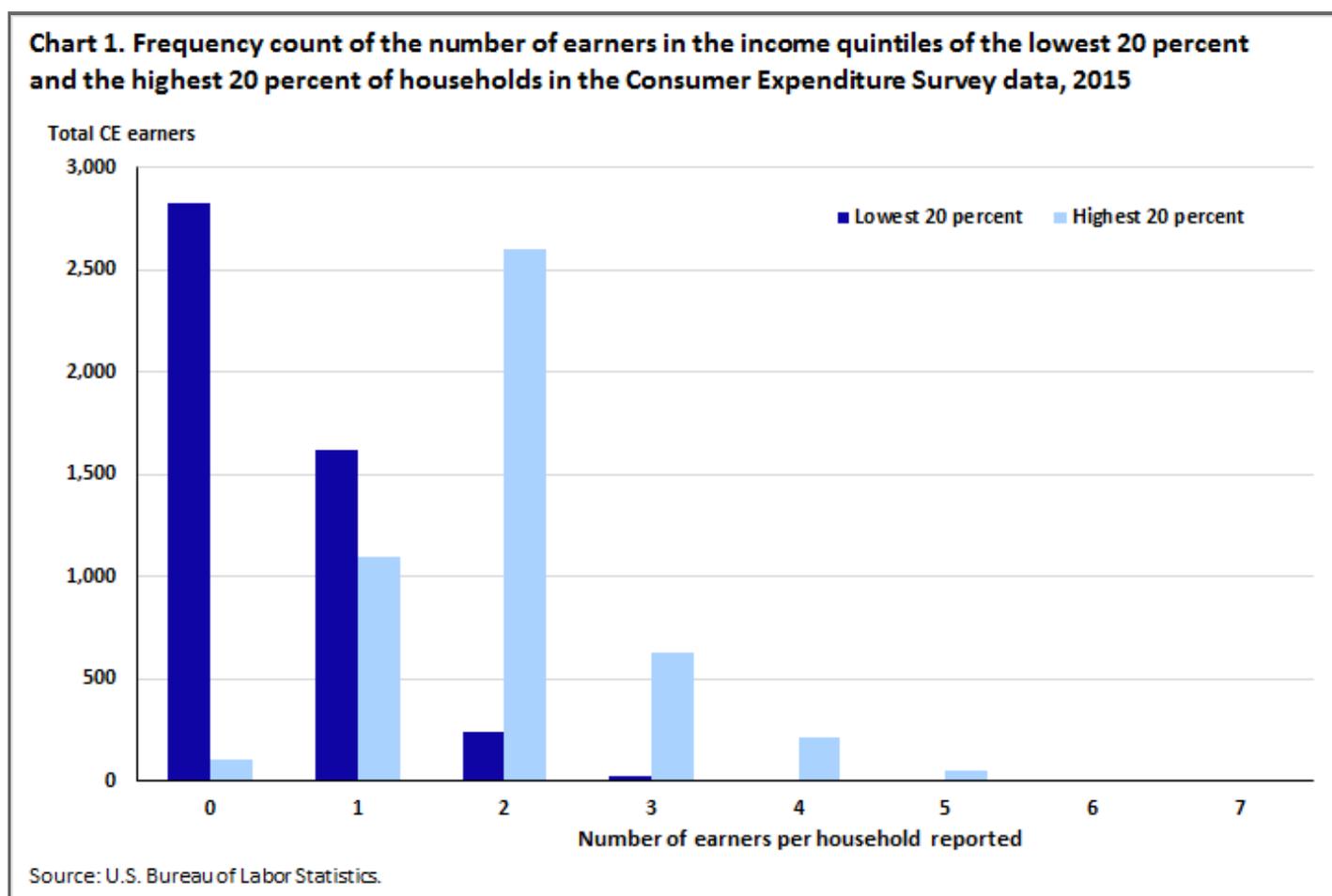

Chart 1. Frequency count of the number of earners in the income quintiles of the lowest 20 percent and the highest 20 percent of households in the Consumer Expenditure Survey data, 2015

Source: U.S. Bureau of Labor Statistics.

Table 1, it shows that the household average income before taxes in the highest 20 percent is $177,851. This figure might lead researchers to classify this group as the "rich" in their analysis.

Table 2. Select examples of consumer units from the Consumer Expenditure Survey microdata, 2015

Household identification number	Number of earners	Income before taxes	Income per earner	Gift to non-consumer unit of stocks, bonds , and mutual funds
2995452	1	103,000	103,000	0
3091693	4	103,000	25,750	0
3105982	0	10,536	0	150,000
Source: U.S. Bureau of Labor Statistics.				

Table 2 shows two observations from the 2015 public use microdata that fall into the highest 20 percent quintile. Both households report $103,000 pre-tax income, but one household has four earners while the other has one earner. The average income per earner of $25,750 in the former example may paint a far different picture of their lifestyle, compared to the latter. The microdata provide many variables one might choose for analysis; however, for this study, we hypothesize that the four-earner household consists of four sisters who pool their resources. They share with their four children a rented house located in a city with a high cost of living where the demand for domestic workers is high. The sisters are each self-employed, earning income by cleaning houses. They each work 50 hours a week and travel by public transportation to 3 different houses each day. The sisters have no assets other than their income. They receive no benefits from their employment and do not qualify for any public assistance. This scenario differs from what might be called the typical scenario of a "rich" family.

Family size. Another important characteristic to consider when using expenditures for food as a measure of well-being may be family size. Table 1 lists the average number of people per household in each quintile. The data show that in 2015, the average household size in the lowest 20 percent was 1.7 people, and the average household size of the highest 20 percent was 3.1 people. As with earners, household size grows with income across the quintile groups. Table 1 reports that households in the lowest 20 percent spent $483 for fruits and vegetables and households in the highest 20 percent spent $1,233. When one uses the average number of persons in each quintile to calculate a per capita average, the values change to $284 and $398 respectively. If one were to choose fresh milk and cream and apply the same calculation, the results change from $94 and $205 to $55 per capita in the lowest 20 percent and $66 per capita for the highest 20 percent. If one were conducting research about nutritional needs of children, the results could be further adjusted to take the average number of children under 18 into account. Using the average of 0.3 children under 18 from table 1 for the lowest 20 percent and 0.8 children for the highest 20 percent, one finds a per child average of $314 for fresh milk and cream in the lowest 20 percent and a per child average of $256 for the highest 20 percent. Of course, those researching nutritional needs might also want to consult other sources to assess the impact of food banks and the WIC program on meeting nutritional goals, which is not covered by the CE.

Words of caution

CE data are complex and detailed. The data can be used for various types of analyses across disciplines. However, it is important for users to understand the caveats and nuances associated with the dataset, because these nuances often go undetected and lead to drawing inaccurate conclusions. The purpose of this article is to delve into the assumptions users often make about how the survey measures and presents income. The data are compiled through an objective collection process, whereas most analyses contain subjective inferences based on the data.

Should researchers use terms like "rich" and "poor" to accurately describe income groups? The CE does not assign subjective titles to the data like "rich" and "poor." These kinds of subjective titles can only be determined by the data user. Having clear definitions is a key part of the research process. As such, it is important that all inferences that use CE data are formed with a complete understanding of what the data contain and how they are presented. In the case of income, the presentation in the tables is a strictly mathematical way of breaking down a single variable. If researchers' definitions are more complex than that, it may be in their best interest to look more deeply into the microdata in order to fully explore their topic rather than adopt the objective assumptions (e.g., high asset and low income individuals not existing in the data sample) to make the table values fit their subjective definition.

This **Beyond the Numbers** article was prepared by Vera Crain and Taylor J. Wilson, economists in Office of Prices and Living Conditions, U.S. Bureau of Labor Statistics. Email: crain.vera@bls.gov; telephone: (202) 691-5640.

Information in this article will be made available upon request to individuals with sensory impairments. Voice phone: (202) 691-5200. Federal Relay Service: 1-800-877-8339. This article is in the public domain and may be reproduced without permission.

NOTES

[1] For more information about the detailed unpublished prepublication table data that is available, see https://www.bls.gov/cex/csxresearchtables.htm. These data are available by email request at CEXINFO@bls.gov.

[2] For more information about microdata and how to obtain these data that is available, see https://www.bls.gov/cex/pumd.htm.

[3] The first member mentioned by the respondent when asked to "Start with the name of the person or one of the persons who owns or rents the home." It is with respect to this person that the relationship of the other consumer unit members is determined. For more information, see https://www.bls.gov/cex/csxgloss.htm.

SUGGESTED CITATION