Standard Errors in the 2015 Consumer Expenditure Survey

by David Swanson

"Standard errors" are a measure of the uncertainty in a survey's estimates caused by the use of data from a representative sample of households instead of the complete universe of households when making the survey's estimates. The United States has approximately 128 million households, and the Consumer Expenditure (CE) Survey interviews approximately 21,000 households per year.[1] The fact that the CE does not interview every household means its estimates have a standard error or margin-of-error associated with them. This article presents the standard errors of the CE's expenditure and income estimates for 2015.

Standard errors are the most common measure of sampling variability of a survey's estimates. Defined as the square root of the survey estimates' variances, they measure how much the CE's estimates would vary if the survey could be repeated over-and-over using a different sample of households every time. Of course it is not feasible to repeat the survey over-and-over, but statistical theory allows standard errors to be estimated anyway.

Standard errors serve two important purposes. First, they provide a general measure of the accuracy of the CE's estimates. And second, they are used to determine whether differences between various expenditure estimates are statistically significant. The statistical significance of the estimates is usually described with a 95% confidence interval, which is informally called the estimate's margin-of-error.

Presentation of the estimates

The CE's expenditure tables for 2015 show four numbers for every item category: the mean, the share, the standard error, and the coefficient of variation. The tables look like this:

Item	All consumer units
Average annual expenditures	
Mean	$55,978.46
Share	100.0
SE	594.00
CV (%)	1.06
Food	
Mean	$7,022.59
Share	12.5
SE	77.17
CV (%)	1.10

The "mean" is an estimate of the average annual expenditure per household on a particular item category. As mentioned above, it comes from interviewing 21,000 households per year. In 2015 the average annual

[1] The CE collects data from approximately 7,000 households per quarter in the Interview survey and 7,000 households per year in the Diary survey. Each household in the Interview survey is interviewed four times, and each household in the Diary survey fills out two weekly diaries, making the total amount of data collected 28,000 quarterly interviews and 14,000 weekly diaries per year. Because of the rotating nature of the Interview survey's sample design, with old addresses constantly dropping out of the sample and new addresses replacing them, there are about 14,000 unique households in the Interview survey each year. Adding this to the 7,000 households in the Diary survey gives 21,000 unique households per year in both surveys.

expenditure per household on all items ("average annual expenditures") was estimated to be $55,978.46, and on food it was estimated to be $7,022.59. These are sometimes called the survey's "point estimates."

The "share" is an item category's expenditure amount expressed as a percent of total expenditures. For example, in 2015 the mean expenditure per household for "food" was $7,022.59, and for all items it was $55,978.46, so the share of total expenditures for "food" was 12.5 percent ($7,022.59 divided by $55,978.46 equals 0.125).

The "standard error" (SE) is a measure of how much the mean expenditure would vary if the survey could be repeated over-and-over using a different sample of households every time. In 2015 the standard error for all items was $594.00, and for food it was $77.17. That means if the CE survey could have been repeated over-and-over with different samples of households, the mean expenditure for all items would have varied by approximately plus-or-minus $594.00, and the mean expenditure for food would have varied by approximately plus-or-minus $77.17.

In mathematical terms, the variability of a survey's estimates is usually described by a 95 percent confidence interval, which is equal to the survey's point estimate plus-or-minus 1.96 times its standard error. Thus the usual way of describing the variability of these point estimates is $55,978.46 plus-or-minus $1,164.24 (= $55,978.46 ± 1.96×594.00) and $7,022.59 plus-or-minus $151.25 (= $7,022.59 ± 1.96×77.17).

The last number is the "coefficient of variation" (CV). It is the standard error divided by the mean expenditure, and it is expressed as a percent. It gives the relative amount of variability instead of the absolute amount of variability in the expenditure estimates. It is useful when comparing expenditure categories whose mean and standard error differ in magnitude. For example, in 2015 the average annual expenditure per household on all items was $55,978.46 and its standard error was $594.00, so its coefficient of variation was 1.06 percent ($594.00 divided by $55,978.46 equals 0.0106). Likewise, the average annual expenditure per household on food was $7,022.59 and its standard error was $77.17, so its coefficient of variation was 1.10 percent ($77.17 divided by $7,022.59 equals 0.0110). These two item categories have approximately the same coefficient of variation, which means they have approximately the same amount of variability relative to the size of their point estimates.

Analysis of the estimates

An examination of the data reveals three significant observations. First, standard errors tend to be smaller for the nationwide estimates than for the individual regions of the country. This is primarily due to their different sample sizes. In general standard errors decrease as the sample size "*n*" increases, and the CE interviews more households in the whole country than in each individual region of the country.

Second, standard errors increase as the number of individual item categories in the broader item category increases. For example, the standard error of the "all items" category is larger than the standard error for "food," which in turn is larger than the standard error for "fruits and vegetables." This is primarily due to the heterogeneity of items in broader item categories.

And third, standard errors decrease as the frequency of purchases increases. For example, the standard error for "telephone services" is smaller than the standard error for the purchase of new cars and trucks even though both item categories have approximately the same mean expenditure. This is due to the tendency of households to spend roughly the same amount of money every month on frequently purchased items like telephone services, while for infrequently purchased items like new cars and trucks there are a lot of $0 expenditures in the CE database interspersed with a small number of large expenditures, causing the variance of infrequently purchased items to be large.

All three of these observations can be explained by a single statistical formula. Let x_{hi} be the expenditure of an individual household on an individual item category (e.g., "beef" or "bakery products"), let "n" be the number of households in the sample, and let $\bar{x}_I$ be the average expenditure per household on an aggregate item category (e.g., "food"). Then a simple calculation from elementary statistics shows

$$SE(\bar{x}_I) = \sqrt{V(\bar{x}_I)} = \sqrt{\frac{1}{n}\sum_{i\in I} V(x_{hi})} .$$

The term on the right-hand side has three components corresponding to the three observations made here. The first observation noted that the variance decreases as the sample size "n" increases. This is clear from the formula by observing that "n" is in the denominator. The second observation noted that the variance increases as the number of individual item categories in the broader item category increases. This is clear from the formula by the number of terms in the sum ($i\in I$) increasing as the number of individual item categories increases. The third observation noted that the variance decreases as the frequency of purchases increases. This can be seen in the formula by the variance of individual item categories ($V(x_{hi})$) being smaller for frequently purchased items.

Standard errors change from year to year as the range of expenditures reported by the households in the survey changes. Although it is not often discussed, every statistic has a standard error, and since the standard error is a statistic it has a standard error as well. In recent years (2011-2015) the standard error for the nationwide all items category ranged from \$519.52 to \$603.93, which gives a rough idea of the true standard error – \$560 plus-or-minus \$50. That means the standard error of the standard error is approximately \$50. From this it can be seen that the estimated mean is generally more stable than the estimated standard error in relative terms, which is a well-known fact from statistical theory.

Nonsampling error

Sampling error is not the only source of uncertainty in a survey's estimates. There is also "nonsampling error," which is everything else. It includes incorrect information provided by the survey's respondents, households missing from the survey's master address file, data processing errors, and so forth. Nonsampling errors occur regardless of whether data are collected from a sample of households or the complete universe of households. Standard errors do not measure these other sources of error.

Variance estimation methodology

Standard textbook formulas for computing variances usually assume the survey's data come from a "simple random sample" of households. Those formulas do not apply to the CE because, like most real-world surveys, the CE does not collect data from a simple random sample of households. Instead the CE draws a stratified random sample of geographic areas around the United States, and then draws a systematic sample of households within the selected areas. As a result, a different method of variance estimation is needed.

There is a class of variance estimation techniques that use "replicates" to produce unbiased variance estimates for surveys with any sample design, simple or complex. In those techniques a number of subsamples are drawn from the survey's full sample of data, and the mean expenditure is computed for both the full sample and each subsample. Then the variance of the full sample expenditure estimate is computed by looking at the variability of all the different subsample expenditure estimates.

The specific technique used by the CE is called "balanced repeated replication" (BRR). In this technique the CE divides the geographic areas in which it collects data into 43 subgroups called "strata." Then the households within each stratum are randomly divided into two "half samples," with half of the households assigned to one half sample, and the other half assigned to the other half sample. There are 43 strata with 2 half samples per stratum, making 2^{43} (approximately 9 trillion) different estimates of the mean expenditure that can be computed from exactly half of the data.

Fortunately it is not necessary to compute all 9 trillion estimates of the mean expenditure to get a reasonable variance estimate. By carefully picking 44 estimates in a balanced way (hence the name of the technique), the same variance estimate can be obtained with considerably less work. Then after generating 44 balanced estimates, the variance of the full sample estimate is computed as the average squared difference between the half sample and full sample estimates:

$$SE(\bar{x}) = \sqrt{V(\bar{x})} \approx \sqrt{\frac{1}{44}\sum_{HS=1}^{44}(\bar{x}_{HS} - \bar{x}_{FullSample})^2} .$$

This gives an unbiased estimate of the CE's variance.

To be more precise, the CE uses a 44×44 Hadamard matrix to pick its 44 balanced estimates. A Hadamard matrix is an $n \times n$ matrix whose entries are ± 1 and where the product of the matrix and its transpose is equal to n times the identity matrix. If $H = [h_{ij}]$ is such a matrix, $\bar{x}_{i1}$, $\bar{x}_{i2}$ are the mean expenditures from half samples 1 and 2 in the i-th stratum, and w_i is the weight of the i-th stratum ($w_i \geq 0$ and $w_1 + \ldots + w_{43} = 1$), then

$$\bar{x}_j = \sum_{i=1}^{43} w_i \left(\frac{1+h_{ij}}{2} \bar{x}_{i1} + \frac{1-h_{ij}}{2} \bar{x}_{i2} \right)$$

is one of the 44 balanced estimates of the mean expenditure, and

$$\hat{V}(\bar{x}) = \frac{1}{44}\sum_{j=1}^{44}(\bar{x}_j - \bar{x})^2 .$$

is the unbiased estimate of the variance used by the CE.[2]

More Information

For more information, contact the Consumer Expenditure Survey Division by telephone at (202) 691-6000, by email at cexinfo@bls.gov, or online at http://www.bls.gov/cex/.

[2] The CE uses a 44×44 Hadamard matrix, but only 43 of its rows. One of its rows cannot be used because it consists of all +1's, which means it does not satisfy the full orthogonal balancing property needed for the technique.