Consumer Expenditures and Income

The **Consumer Expenditure Survey (CE)** is a nationwide household survey conducted by the U.S. Bureau of Labor Statistics (BLS) to find out how Americans spend their money. It is the only federal government survey that provides information on the complete range of consumers' expenditures as well as their incomes and demographic characteristics. BLS publishes 12-month estimates of consumer expenditures twice a year with the estimates summarized by various income levels and household characteristics. BLS also produces annual public-use microdata files to help researchers analyze the data in more detail.

The CE consists of estimates derived from two separate surveys, the Interview Survey and the Diary Survey. The Quarterly Interview Survey is designed to collect data on large and recurring expenditures that consumers can be expected to recall for a period of 3 months or longer, such as rent and utilities, and the Diary Survey is designed to collect data on small, frequently purchased items, including most food and clothing. Together, the data from the two surveys cover the complete range of consumers' expenditures. CE data are collected for BLS by the U.S. Census Bureau.

Table of Contents

CONCEPTS 3
DATA SOURCES 5
DESIGN 15
CALCULATION 18
CALCULATION METHODOLOGY .18
CALCULATION PRECISION 19
PRESENTATION 21
HISTORY 23
MORE INFORMATION 26

Quick Facts: Consumer Expenditures and Income	
Subject areas	Consumer spending , Income, Pay
Key measures	Income Spending
How the data are obtained	Survey of households
Classification system	Demographic
Periodicity of data availability	Annual, Semi-annual
Geographic detail	Census region, Metro area, National, Region
Scope	Civilian noninstitutional population

Quick Facts: Consumer Expenditures and Income	
Key products	• Consumer Expenditures Annual news release • Annual and Midyear Data Tables • Annual Public-Use Microdata • Consumer spending analytical articles and data comparisons
Program webpage	www.bls.gov/cex

Concepts

A consumer expenditure survey is a specialized study in which the emphasis is on data related to family expenditures for goods and services used in day-to-day living. In addition to data on family expenditures, the Consumer Expenditure Survey (CE) collects information on the amount and sources of family income, changes in assets and liabilities, and demographic and economic characteristics of family members.

The importance of the CE is its ability to allow data users to examine the association of expenditures and income of consumers to consumer characteristics. CE survey data are of value to government and private agencies interested in studying the welfare of particular segments of the population, such as the elderly, low-income families, urban families, and those receiving food stamps. Data also are used by economic policymakers interested in the effects of policy changes on levels of living among diverse socioeconomic groups, and econometricians find the data useful in constructing economic models. Market researchers find consumer expenditure data valuable in analyzing the demand for groups of goods and services. The Department of Commerce uses the survey data as a source of information for calculating the Supplemental Poverty Measure, and in revising its benchmark estimates of selected items in the expenditure and income components of the national accounts. The Department of Agriculture uses CE information to estimate the cost of raising a child from birth to age 18, and the Internal Revenue Service uses expenditures to calculate alternate sales tax standard deductions. The Department of Defense uses the data in determining cost-of-living allowances for military personnel living off military bases.

As in the past, the regular revision of the Consumer Price Index (CPI) remains a primary reason for undertaking the Bureau's extensive Consumer Expenditure Survey. Results of the CE are used to select new "market baskets" of goods and services for the index, to determine the relative importance of components, and to derive cost weights for the market baskets.

In August 2002, the Bureau of Labor Statistics began publishing another index, the Chained Consumer Price Index for All Consumers (C-CPI-U), which supplements the CPI for All Urban Consumers (CPI-U) index and the CPI for Urban Wage Earners and Clerical Workers (CPI-W) index. The use of expenditure data from different time periods distinguishes the C-CPI-U from the other two CPI measures, which use a single expenditure base period to compute price change over time. The chained index is designed to better measure the change in the cost of living, as compared with the CPI-U and CPI-W, which measure the change in a fixed market basket of goods and services in retail outlets. The C-CPI-U uses expenditure data from different time periods to reflect the effect of substitution that consumers make across item categories, in response to changes in the relative prices of goods and services. For more information, see: https://www.bls.gov/cpi/additional-resources/chained-cpi-introduction.pdf .

The measurement unit for which expenditure reports are collected is the set of eligible individuals constituting a consumer unit (CU). The CU is defined as 1) all members of a particular housing unit who are related by blood, marriage, adoption, or some other legal arrangement, such as foster children; 2) a person living alone or sharing a household with others, or living as a roomer in a private home, lodging house, or in permanent living quarters in a hotel or motel, but who is financially independent; or 3) two or more unrelated persons living together who pool their income to make joint expenditure decisions. Students living in university-sponsored housing are also included

in the sample as separate consumer units. Information on members living in the consumer unit is identified by their relationship to the reference person, who is defined as 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."

Survey participants report dollar amounts for goods and services purchased by any member of the consumer unit during the reporting period, regardless of whether payment was made at the time of purchase. Expenditure amounts for items purchased by the consumer unit include all applicable sales and excise taxes. Excluded from expenditure total amounts are any business-related expenditures and expenditures for which the family is reimbursed.

The Interview Survey is designed to collect data on the types of expenditures that respondents can be expected to recall for a period of 3 months or longer. In general, expenditures reported in the Interview Survey are either relatively large, such as those for property, automobiles, and major appliances, or that occur on a fairly regular basis, such as for rent, utilities, or insurance.

The primary objective of the Diary Survey is to obtain expenditure data on small, frequently purchased items, which can be difficult to recall even a few weeks later. These items include food and beverage expenditures at home and in eating places; housekeeping supplies and services; nonprescription drugs; most clothing items, and personal care products and services.

Last Modified Date: February 25, 2016

Data Sources

The concepts section gave some information on the two types of surveys the Consumer Expenditure (CE) program administers: the Interview Survey and the Diary Survey. This section will provide more detailed information on how these two surveys are administered.

Survey notification and collection method

CE data are collected for the Bureau of Labor Statistics (BLS) by the U.S. Census Bureau. A selected sample housing unit is notified in advance by a letter from the Census Bureau informing the occupants about the purpose of the survey and the upcoming visit by the interviewer. Both the Interview Survey and the Diary Survey are conducted primarily by personal visit with some telephone interviewing. The interviewer uses a structured questionnaire to collect both the demographic and expenditure data in the Interview Survey. The demographic data for the Diary Survey are collected by the interviewer, whereas the expenditure data are entered on the diary form by the respondent. Any eligible household member who is at least 16 years old can serve as the respondent in either survey.

Interview Survey details

The Interview Survey collects detailed data on an estimated 60 to 70 percent of total family expenditures. In addition, global estimates are obtained for food and other selected items such as alcoholic beverages and tobacco products. These global estimates account for an additional 20 to 25 percent of total expenditures. Global estimates are based on responses to global questions in the survey. A global question is designed to elicit comprehensive information about multiple subcategories using a single question. Global questions can be contrasted with detailed questions, in which a series of questions about specific types within a larger category are asked. For example, rather than asking individual detailed questions about purchases of pants, dresses, and footwear, and then aggregating to a total expenditure amount, a global question might ask for the total amount spent on clothing. On average, it takes 60 minutes to complete the interview.

In the Interview Survey, each family in the sample is interviewed every 3 months over four calendar quarters. The sample for each quarter is divided into three monthly panels, with consumer units (CUs) being interviewed every 3 months in the same panel of every quarter.

After the fourth interview, the sample unit is dropped from the survey and replaced by a new sample unit. For the survey as a whole, 25 percent of the sample in each quarter are new families introduced into the sample, to replace families that have completed their participation. Data collected in each quarter are treated independently, so that published 12-month estimates are not dependent upon a particular family participating in the survey for a full four quarters.

Exhibit 1 shows how Interview Survey consumer units rotate in and out of the sample. In this example, the first interviews start in April, May, and June 2015. Three months later, the second interviews begin. Consumer units first interviewed in April 2015 are re-interviewed in July, October, and January 2016. And while the second set of

interviews begins in July 2015 for the units first interviewed in April, a new set of addresses is starting their set of four interviews.

Exhibit 1. Quarterly Interview Survey rotation

Interview year and month		Interview set			
		1	2	3	4
See footnotes at end of table.					
2015	APR	a			
	MAY	b			
	JUN	c			
	JUL	d	a		
	AUG	e	b		
	SEPT	f	c		
	OCT		d	a	
	NOV		e	b	
	DEC		f	c	
2016	JAN			d	a
	FEB			e	b
	MAR			f	c
	APR				d
	MAY				e
	JUN				f
	JUL				
	AUG				
	SEPT				

Note: The column headings are the interview numbers. Each letter designates a panel or group of households. A household is part of a panel of four interviews conducted every 3 months.
Source: Bureau of Labor Statistics.

During the initial interview (interview set 1 on exhibit 1) information is collected on demographic and family characteristics and on the inventory of major durable goods of each consumer unit. Expenditure information that is based on 3 months of respondents' recall is also collected during this interview.

Data collected in these questionnaires are arranged by major expenditure component (for example, housing, transportation, medical, and education), and form the basis of the expenditure estimates derived from the Interview Survey. Wage, salary, and other information on the employment of each member of a consumer unit are also collected in the first interview and updated in the fourth interview. Expenditure data are collected via two major types of questions. The first set of questions asks the respondent for the month of purchase of each reported expenditure. The second asks for a quarterly amount of expenditures. The use of these two questions varies, depending on the types of expenditures collected. Most of the data are collected using the direct monthly method. A portion of the data are collected by asking for quarterly expenses, but this also includes asking for the amount that was spent in the current month so as to only have the expenses that occur in the 3-month reference period. In the final interview, an annual supplement is used to obtain a financial profile of the consumer unit. This profile consists of information on the income of the consumer unit as a whole, including unemployment compensation;

income from royalties, dividends, and estates; alimony and child support. A 12-month recall period is used in the collection of income- and asset-type data.

Diary Survey details

The separate Diary Survey is completed by the respondent family for two consecutive 1-week periods.

Two separate questionnaires are used by the Census Bureau to collect Diary Survey data: a Household Characteristics Questionnaire and a Record of Daily Expenses. In the Household Characteristics Questionnaire, the interviewer records information pertaining to age, sex, race, marital status, and family composition, as well as information on the work experience and earnings of each member of the consumer unit. This socioeconomic information is used by the Bureau to classify the consumer unit for publication of statistical tables, as well as for economic analyses. Data on household characteristics also provide the link in the integration of Diary Survey expenditure data with Interview Survey expenditure data that permit the publication of a full profile of consumer expenditures by demographic characteristics.

The daily expense record is designed as a self-reporting, product-oriented diary, in which respondents record a detailed description of all expenses for two consecutive 1-week periods. Diarykeeping can start on any day of the week. Data collected each week are treated as statistically independent – each week's diary is separately weighted to be representative of the sample. The diary is divided by day of purchase and by four classifications of goods and services—food away from home, food at home, clothing, and all other goods and services—a breakdown designed to aid the respondent in recording the entire consumer unit's daily purchases. The items reported are subsequently coded by the Census Bureau, so that BLS can aggregate individual purchases for representation in the CPI and for presentation in statistical tables.

The Diary Survey asks for almost all expenses that the consumer unit incurs during the survey week. Expenses incurred by family members while away from home overnight or on vacation, and for credit and installment plan payments are excluded.

It takes approximately 25 minutes over three visits for the interviewer to collect the demographic data and to instruct the respondent on how to keep the diary. It is estimated that it takes the respondent 15 minutes each day to complete the diary.

Integrated survey data

Integrated data from the BLS Diary and Interview Surveys provide a complete accounting of consumer expenditures and income, which neither survey component alone is designed to do. Most apparel and food expenditures in the integrated data tables come from the Diary, while rent, mortgage interest, utilities, and major appliances come from the Interview. Some expenditure items are collected only by the Diary or Interview Survey. For example, the Diary Survey collects data on detailed food expenditures and items, such as postage and nonprescription drugs, which are not collected in the Interview Survey. The Interview Survey collects data on expenditures for overnight travel and information on insurance reimbursements for medical-care costs and automobile repairs, which are not collected in the Diary Survey. Data on average annual expenditures that come

exclusively from the Interview Survey, including global estimates, such as those for food and alcoholic beverages, average about 95 percent of the total estimated spending, based on integrated Diary and Interview Survey data. For items unique to one or the other survey, the choice of which survey to use as the source of data is obvious. However, there is considerable overlap in coverage between the surveys. Because of the overlap, the integration of the data presents the problem of determining the appropriate survey component from which to select the expenditure items. When data are available from both survey sources, the more reliable of the two is selected, as determined by statistical methods. The selection of the survey source is evaluated every 2 years. https://www.bls.gov/cex/anthology11/csxanth3.pdf.

Quality control and confidentiality

Quality control and data integrity are provided by a re-interview program, which evaluates the performance of the individual interviewer, to determine how well the procedures are being carried out in the field. The re-interview is conducted by a Census Bureau supervisor or an interviewer at a National Processing Center Data Contact Center (a centralized telephone call center for conducting interviews). Subsamples of approximately 9 percent of households in both the Interview and Diary surveys are re-interviewed on an ongoing basis.

All data collected in both surveys are subject to Census Bureau and BLS confidentiality requirements that prevent the disclosure of the respondents' identities. The information that respondents provide is used solely for statistical purposes. All Census Bureau and BLS employees who work with the CE data take an oath of confidentiality and are subject to fines and imprisonment for improperly disclosing information provided by respondents. Confidentiality certification training is required annually.

Names and addresses are removed from all forms and datasets prior to transmission from the Census Bureau to BLS and are not included in any statistical releases. At BLS, the data are processed and stored on secure servers, with access limited to employees having security clearances. As a further precaution, BLS applies certain restrictions to the microdata available on the public-use files. These include geographical and value restrictions that prevent identification of respondents.

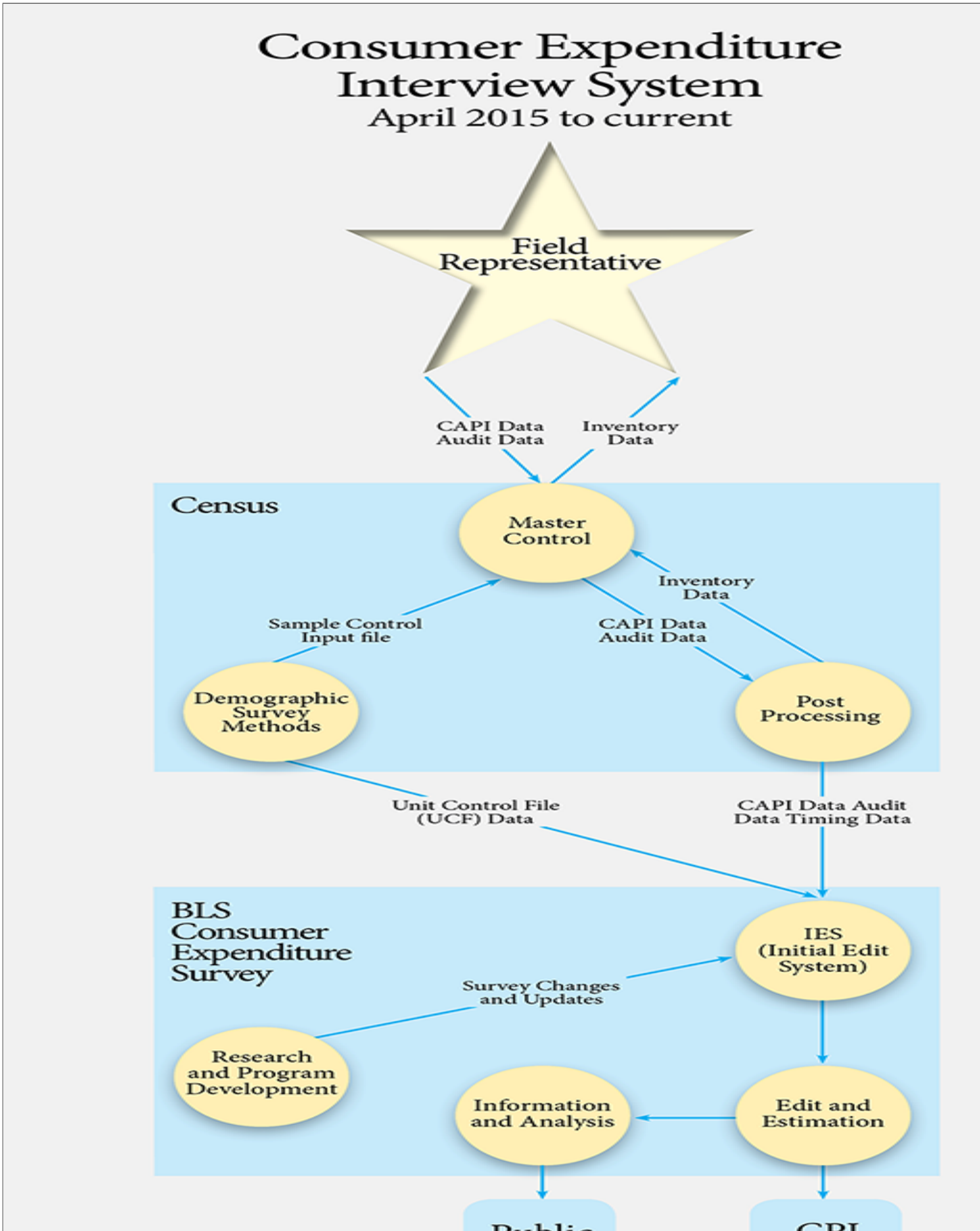
Data collection and processing

Due to differences in format and design, the Interview Survey and the Diary Survey are collected and processed separately. The Census Bureau, under contract with BLS, carries out data collection for both surveys. In addition to its collection duties, the Census Bureau does field editing and coding, checks consistency, ensures quality control, and transmits the data to BLS. In preparing the data for analysis and publication, BLS performs additional review and editing procedures.

Interview Survey. Beginning April 2003, census field representatives started collecting the Interview Survey data using a Computer Assisted Personal Interview (CAPI) instrument. This was a major improvement from the paper-and-pencil data collection that had been in place since 1980. The CAPI instrument enforces question skip patterns, allows for data confirmation of high expenditure values, and reduces processing time. The CAPI instrument allows the field representative to see and select from a list the category that an expense would fit into as well as any classifying information about the expense. Categories of expenses would be pants, oil changes, or trips. Examples

of classifying information consist of age and sex of clothing purchase, make and model of vehicles owned, and trip destination. See Exhibit 2 for the diagram of the process flow from collection by the field representative (FR) to delivering the data to the CPI program, and then to publishing the information for the public.

Exhibit 2.


Consumer Expenditure Interview System
April 2015 to current
Field Representative
CAPI Data Audit Data
Inventory Data
Census
Master Control
Inventory Data
Sample Control Input file
CAPI Data Audit Data
Demographic Survey Methods
Post Processing
Unit Control File (UCF) Data
CAPI Data Audit Data Timing Data
BLS Consumer Expenditure Survey
IES (Initial Edit System)
Survey Changes and Updates
Research and Program Development
Information and Analysis
Edit and Estimation

Data are electronically transferred from a field representative's laptop at completion of the interview to the Census Master Control System. The Census Bureau does post processing by reformatting the data into datasets and performing special processing, including converting missing values to special characters and merging data records into the required BLS output structure. Some inventory data, such as those for vehicle and mortgage records, are copied into an input file that is loaded onto the laptops for subsequent interviews during the next quarter. This way, a few fields are updated each quarter, rather than an entire data record. As mentioned earlier, names and addresses of respondents are not transmitted to BLS.

At BLS, a series of automated edits are applied to monthly data. These edits check for inconsistencies; identify missing expenditure amounts for later imputation; impute values for missing demographic variables; calculate weights; and adjust data to include sales tax, exclude business expenses, and net out reimbursed expenditures.

Monthly data files then are combined into quarterly databases, and a more extensive data review is carried out. During this data review, BLS conducts the following steps: verifies counts and means by region, checks family relationship coding inconsistencies, and inspects selected extreme values for expenditure and income categories. Outlier values are confirmed or corrected. Other adjustments convert mortgage and vehicle loan payments into principal and interest (using associated data on the interest rate and term of the loan). In addition, BLS verifies the various data transformations it performs. As with outlier values, other cases of questionable data values or relationships are investigated, and errors are corrected before release of the data for public use.

Three major types of data adjustment routines—imputation, allocation, and time adjustment—improve estimates derived from the Interview Survey. Data imputation routines account for missing or invalid entries and address all fields in the database, except assets. Allocation routines are applied, when respondents provide insufficient detail to meet tabulation requirements. For example, combined expenditures for the fuels and utilities group are allocated among the components of that group, such as natural gas and electricity. Time adjustment routines are used to classify expenditures by month, prior to aggregation of the data to calendar-year expenditures. Tabulations are made before and after data adjustment routines, to analyze the results.

The CE implemented multiple imputations of income data starting with the publication of 2004 data. Before that, only income data collected from complete income reporters (those CUs whose reference person provided answers to most major sources of income) were published. However, even complete income reporters may not have provided information on all sources of income for which they reported receipt. With the collection of bracketed income data starting in 2001, this problem was reduced but not eliminated. One limitation was that bracketed data only provided a range in which income falls, rather than a precise value for that income. In contrast, imputation allows income values to be estimated when they are not reported. In multiple imputations, several estimates are made for the same consumer unit, and the average of these estimates is published. See https://www.bls.gov/cex/csximpute.htm and https://www.bls.gov/opub/mlr/2009/08/art3full.pdf.

Beginning with the 2014 published data, the CE imputes all state and federal income taxes for all consumer units. The CE uses an internal BLS version of the National Bureau of Economic Research's TAXSIM software in estimating tax liabilities. Starting in 2015, respondents are no longer asked questions about their income taxes, which reduced respondent cognitive burden.

The CE Interview questionnaire is revised every 2 years to incorporate new products and services, to clarify instructions, to improve navigation through the instrument, to incorporate changes required for the CPI, and to streamline the interview by deleting outdated items. Whereas changes to the questionnaire are made biennially, CE staff continuously monitors the emergence of new goods and services available in the marketplace, as well as changes in the relative importance of existing items in consumers' budgets.

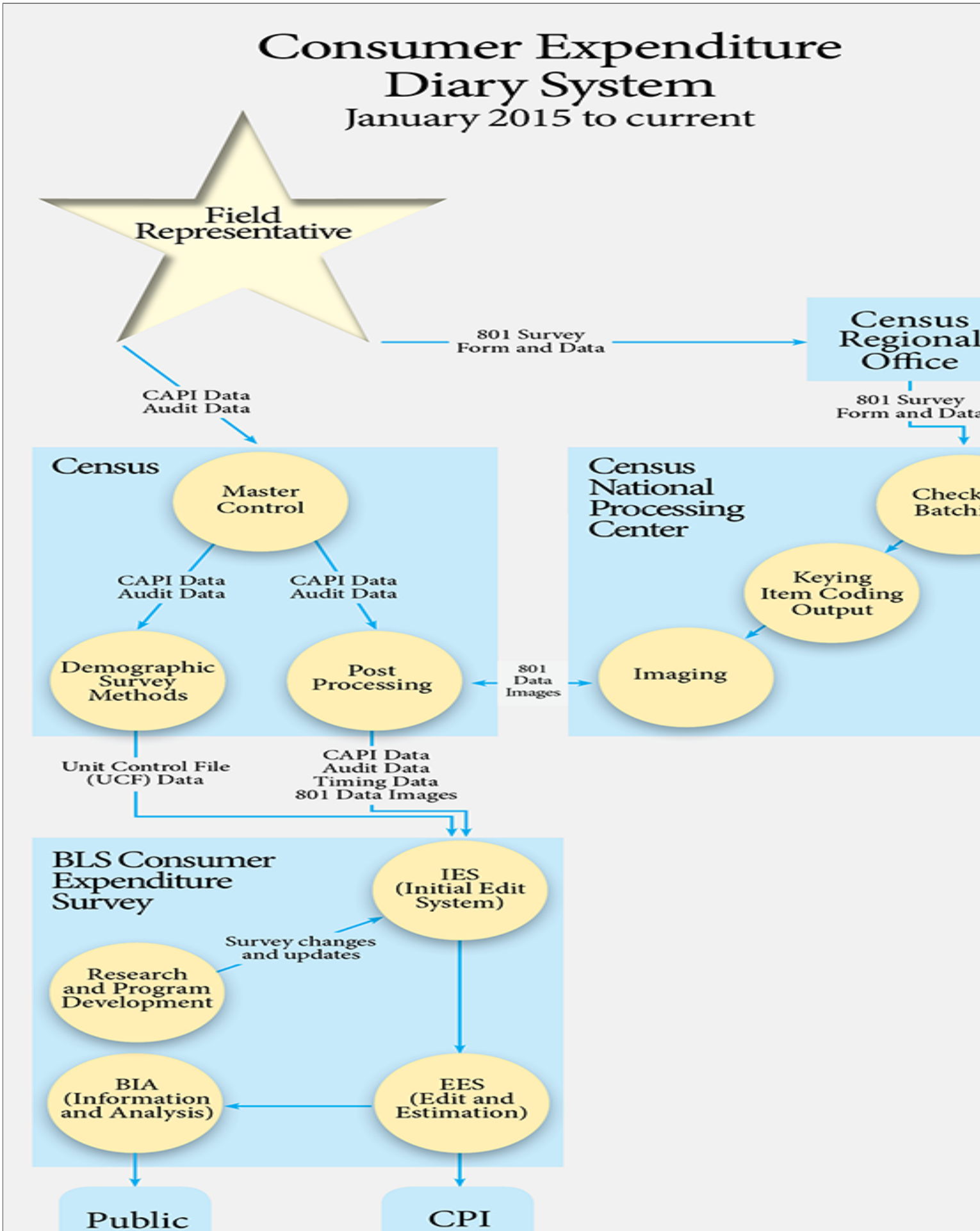
Updated information on how to report new goods and services is provided to the field representatives on a regular basis. Also, new items are incorporated in a product index that enables Census field representatives to classify these new items by the appropriate item codes. The product index is a file that contains a list of items that could be found under a particular item code. Given the space limitation in the information book shown to respondents, the Census Bureau is not able to include a lot of items under each category. The product index contains more descriptions or examples of items. For example, camping equipment has "tent" as an example in the information booklet, but not "tent poles." So the field representative could look up tent poles in the product index and see which item code it belongs under. This is a file that is accessible through the CE Quarterly Interview Survey API instrument. It is maintained by the BLS CE Production and Control staff as part of normal revisions to the survey.

Diary Survey. At the beginning of the 2-week collection period, the Census Bureau interviewer, using the Household Characteristics Questionnaire (a CAPI instrument), records demographic information on members of each sampled consumer unit.

At the end of the first week, the interviewer collects the diary, reviews the entries, answers any questions that the respondent may have, and leaves a second diary. At the end of the second week, the interviewer picks up the second diary and reviews the entries. During this time, the interviewer again uses the Household Characteristics Questionnaire to collect previous-year information on work experience and income.

Diaries for both week 1 and week 2 may be placed with the consumer unit at the first week interview under certain circumstances depending on distance and special circumstances for the respondent. When the interviewer places both diaries at the first week dropoff, the interviewer conducts all first week dropoff procedures and informs the respondent of the need to keep 2-weeks of diaries. At the end of the 2-week collection period, the interviewer conducts all other necessary interview activities that occur at the end of the 2-week period. Each week of a consumer unit's participation in the survey is treated as a separate entity. See Exhibit 3 for a description of the process starting from data collection by the field representative (FR) to the regional office (RO) to the National Processing Center (NPC) or directly to the Census headquarters, all the way to delivering the data to BLS for the CPI, and making the information available to the public.

Exhibit 3.


Consumer Expenditure Diary System
January 2015 to current
Field Representative
801 Survey Form and Data
Census Regional Office
CAPI Data Audit Data
801 Survey Form and Data
Census
Master Control
CAPI Data Audit Data
CAPI Data Audit Data
Demographic Survey Methods
Post Processing
801 Data Images
Census National Processing Center
Keying Item Coding Output
Imaging
Unit Control File (UCF) Data
CAPI Data Audit Data Timing Data 801 Data Images
BLS Consumer Expenditure Survey
IES (Initial Edit System)
Survey changes and updates
Research and Program Development
BIA (Information and Analysis)
EES (Edit and Estimation)
Public
CPI

The Census Bureau performs preliminary processing activities, including a number of data edits and adjustments. Data in the diaries are reviewed during a field edit for completeness and consistency. After the diaries are sent to the Census Bureau's National Processing Center, expenditure data captured in the diaries are key-entered into electronic formats assisted by auto-coding to fill in typical descriptions; and a computer file of the database containing these data is produced and transmitted monthly to Census headquarters, along with image files of the diaries. Census headquarters merge the expenditure data with the data collected in the Household Characteristics Questionnaire, removes personal identifying information, and transmits the merged file monthly to BLS. At BLS, data are processed by computer to calculate population weights, to impute demographic characteristics for missing or inconsistent demographic data, to impute values for weeks worked when nonresponse is encountered, and to apply appropriate sales taxes to the expenditure items.

The Diary Survey combines three monthly diary data files into quarterly databases; and BLS screens the data for invalid coding and inconsistent relationships, as well as for extreme values recorded or keyed erroneously. BLS then corrects any coding and extreme-value errors found.

Two types of data adjustment routines—allocation and imputation—improve the Diary Survey estimates. Allocation routines transform reports of nonspecific items into specific ones. For example, when respondents report expenditures for meat rather than beef or pork, allocations are made, using proportions derived from item-specific reports in other completed diaries. BLS imputes missing attributes, such as age or sex, or a product's type of packaging needed for mapping Diary Survey expenditures. Income data from the Diary Survey are processed in the same way as in the Interview Survey. Income tax data are not estimated for the Diary Survey because the published integrated information uses Interview Survey after-tax income information, and because of the limited amount of background detail collected in the Diary Survey.

Last Modified Date: February 25, 2016

Design

Design selection of households

The Consumer Expenditure Survey (CE) is a nationwide household survey representing the entire U.S. civilian noninstitutional population. It includes people living in houses, condominiums, apartments, and group quarters such as college dormitories. It excludes military personnel living overseas or on base, nursing home residents, and people in prisons. The civilian noninstitutional population represents more than 98 percent of the total U.S. population.

The selection of households for the survey begins with the definition and selection of primary sampling units (PSUs). PSUs are small clusters of counties grouped together into geographic entities called “core-based statistical areas” (CBSAs), which are defined by the Office of Management and Budget (OMB) for use by federal statistical agencies in collecting, tabulating, and publishing federal statistics. The CE currently uses OMB definitions from 2012. There are two types of CBSAs: metropolitan and micropolitan. Metropolitan CBSAs are areas that have an urban “core” of 50,000 or more people, plus the adjacent counties that have a high degree of social and economic integration with the core as measured by commuting ties. Micropolitan CBSAs are similar to metropolitan CBSAs but they have an urban core of 10,000 to 50,000 people. Areas without an urban core or whose urban core is under 10,000 people are called non-CBSA areas. See https://www.census.gov/programs-surveys/metro-micro/about.html.

Starting in 2015, the geographic sample used in the survey consists of 91 PSUs based on population numbers from the 2010 Decennial Census that are classified into three categories:

- 23 “S” PSUs, which are metropolitan CBSAs with a population over 2.5 million people (self-representing PSUs)
- 52 “N” PSUs, which are metropolitan and micropolitan CBSAs with a population under 2.5 million people (nonself-representing PSUs)
- 16 “R” PSUs, which are non-CBSA areas (“rural” PSUs)

The 23 “S” PSUs are the largest CBSAs in the country, and they were selected with certainty for the CE sample. The 52 “N” and 16 “R” PSUs are smaller CBSAs that were randomly selected from the rest of the country, with their probabilities being proportional to their populations. The 23 “S” and 52 “N” PSUs are also used by the Consumer Price Index program, but not the 16 “R” PSUs because the CPI measures inflation only in urban areas of the country.

Within these 91 PSUs, the list of addresses from which the sample is drawn comes from two sources called “sampling frames.” The primary sampling frame for both the Diary Survey and the Interview Survey is the Census Bureau’s Master Address File (MAF). That file has all residential addresses identified in the 2010 census and is updated twice per year with the U.S. Postal Service’s Delivery Sequence File. Over 99 percent of the addresses used in the survey come from the MAF. It is supplemented by a small Group Quarters frame, which is a list of housing units that are owned or managed by organizations for residents who live in group arrangements such as

college dormitories and retirement communities. Less than 1 percent of the addresses used in the CE come from the Group Quarters frame.

The Census Bureau selects a sample of approximately 12,000 addresses per year from these two frames to participate in the Diary Survey. Usable diaries (two 1-week diaries per household) are obtained from approximately 6,900 households at those addresses. Diaries are not obtained from the other addresses due to refusals, vacancies, ineligibility, or the nonexistence of a housing unit at the selected address. The placement of diaries is spread equally over all 52 weeks of the year.

The Interview Survey is a rotating panel survey in which approximately 12,000 addresses are contacted each calendar quarter of the year for the survey. One-fourth of the addresses that are contacted each quarter are new to the survey. Usable interviews are obtained from approximately 6,900 households at those addresses each quarter of the year. After a housing unit has been in the sample for four consecutive quarters, it is dropped from the survey, and a new address is selected to replace it. Before 2015, the Interview Survey included a preliminary bounding interview, and each CU could be contacted up to five times over five quarters. The bounding interview, which recorded recent major expenditures for comparison with subsequent purchases, was determined to be unnecessary, and was dropped at the beginning of 2015 to save money and reduce respondent burden and collection costs. See https://www.bls.gov/cex/research_papers/pdf/Recommendation-Regarding-the-Use-of-a-CE-Bounding-Interview.pdf.

Cooperation levels

Response data for the 2013 CE are shown in table 1. For the Interview Survey, each unique housing unit provides up to four usable interviews per year. For the Diary Survey, each unique housing unit provides up to two usable diaries (weeks 1 and 2). Most Diary Survey respondents participate in both weeks.

There are three general categories of nonresponse:

- *Type A nonresponses* are refusals, temporary absences, and noncontacts.
- *Type B nonresponses* are vacant housing units, housing units with temporary residents, and housing units under construction.
- *Type C nonresponses* are nonresidential addresses, such as destroyed or abandoned housing units, and housing units converted to nonresidential use.

Response rates are defined as the percentage of eligible housing units (that is, the designated sample less Type B and Type C nonresponses) from which usable interviews are collected by the Census Bureau. In the 2013 Interview Survey, there were 48,920 eligible housing units from which 32,822 usable interviews were collected, resulting in a response rate of 67.1 percent. In the 2013 Diary Survey, there were 19,252 eligible housing units from which 12,511 usable interviews were collected, resulting in a response rate of 65.0 percent. (See table 1.)

Table 1. Analysis of response in the Consumer Expenditure Survey, 2013

Sample unit	Interview Survey	Diary Survey

See footnotes at end of table.

Table 1. Analysis of response in the Consumer Expenditure Survey, 2013

Sample unit	Interview Survey	Diary Survey
Housing units designated for survey	60,166[1]	25,480[2]
Less: Type B and type C nonresponses	11,246	6,228
Equals: Eligible units	48,920	19,252
Less: Type A nonresponses	16,098	6,741
Equals: Interviewed units	32,822	12,511
Percentage of eligible units interviewed	67.1	65.0

[1]In 2015, the number of interviews per household decreased from five to four. The data in this table are from 2013, so the numbers in the Interview Survey column are approximately 25 percent larger than they will be starting in 2015.

[2]The number of Diary Survey addresses (12,740) multiplied by two weekly diaries.

Note: Type A nonresponses are refusals, temporary absences, and noncontacts. Type B nonresponses are vacant housing units, housing units with temporary residents, and housing units under construction. Type C nonresponses are nonresidential addresses, such as destroyed or abandoned housing units, and housing units converted to nonresidential use.

Source: U.S. Bureau of Labor Statistics.

Last Modified Date: February 25, 2016

Calculation

Calculation methodology

The use of Consumer Expenditure Survey's (CE) sample data to estimate population quantities of interest, such as the average expenditure per consumer unit on a particular item category, is achieved through the use of weights. Each consumer unit in the survey is assigned a weight that is the number of similar consumer units in the U.S. civilian noninstitutional population the sampled consumer unit represents. Using these weights, the average expenditure per consumer unit on a particular item category is estimated with the standard weighted average formula:

where

$\bar{y}$ = the average expenditure per consumer unit on the item category,

y_i= the expenditure made by the ith consumer unit on the item category,

w_i= the weight of the ith consumer unit in the sample, and

S = the sample of consumer units that participated in the survey.

For example, if y_i is the expenditure on eggs made by the ith consumer unit in the sample during a given time period, then $\bar{y}$ is an estimate of the average expenditure on eggs made by all consumer units in the U.S. civilian noninstitutional population during that period.

If one wants to estimate the proportion of consumer units that purchased eggs during a given period, then the same formula is applied, where y_i is set equal to 1 if the ith consumer unit purchased eggs during the time period, and 0 if it did not. When this binary definition of y_i is used, $\bar{y}$ is an estimate of the proportion of all consumer units in the U.S. civilian noninstitutional population that purchased eggs during the given period.

Several factors are involved in computing the weight of each consumer unit from which a usable interview is received. Each consumer unit is initially assigned a *base weight* equal to the inverse of the consumer unit's probability of being selected for the sample. The total U.S. target population counts for these base weights come from the Census Current Population Survey. Base weights in the CE are typically around 10,000, which means a consumer unit in the sample represents 10,000 consumer units in the U.S. civilian noninstitutional population?itself

plus 9,999 other consumer units that were not selected for the sample. The base weight is then adjusted by the following factors to correct for certain nonsampling errors:

Weighting control factor. This adjusts for subsampling in the field. Subsampling occurs when a data collector visits a particular address and discovers multiple housing units where only one housing unit was expected.

Noninterview adjustment factor. This adjusts for interviews that cannot be conducted in occupied housing units due to a consumer unit's refusal to participate in the survey or the inability to contact anyone at the housing unit in spite of repeated attempts. This adjustment is based on region of the country, consumer unit size, number of contact attempts, and the average adjusted gross income in the consumer unit's zip code according to a publicly available database from the Internal Revenue Service.

Calibration factor. This adjusts the weights to 24 "known" population counts to account for frame undercoverage. These known population counts are for age, race, household tenure (owner or renter), region of the country, and urban or rural. The population counts are updated quarterly using the Current Population Survey estimates. Each consumer unit is given its own unique calibration factor. There are infinitely many sets of calibration factors that can make the weights add up to the 24 known population counts, and the CE uses nonlinear programming to select the set that minimizes the amount of change made to the "initial weights" (initial weight = base weight x weighting control factor x noninterview adjustment factor).

After adjusting the base weights by these factors, the *final weights* are typically around 17,000, which means an interviewed consumer unit represents 17,000 consumer units in the U.S. civilian noninstitutional population?itself plus 16,999 other consumer units that did not participate in the survey.

Calculation precision

The precision of the estimator $\bar{y}$ is measured by its standard error. Standard errors measure the sampling variability of the CE estimates. That is, standard errors measure the uncertainty in the survey estimates caused by the fact that a random sample of consumer units from across the United States is used instead of every consumer unit in the nation. See table 2.

The CE's standard errors are estimated by using the method of "balanced repeated replication." In this method, the sampled PSUs are divided into 43 groups (called *strata*), and the consumer units within each stratum are randomly divided into two *half samples*. Half of the consumer units are assigned to one half sample, and the other half are assigned to the other half sample. Then 44 different estimates of $\bar{y}$ are created using data from only one half sample per stratum. There are many combinations of half samples that can be used to create these replicate estimates, and the CE uses 44 of them that are created in a "balanced" way with a 44x44 Hadamard matrix. The standard error of $\bar{y}$ is then estimated by:

where

$\bar{y}_r$ is the rth replicate estimate of $\bar{y}$.

The coefficient of variation is a related measure of sampling variability that measures the variability of the survey estimate relative to the mean. It is defined by the equation:

.

Table 2. Precision of the Consumer Expenditure Survey expenditure estimates, integrated Diary and Interview Survey data, 2013

Item category	Average annual expenditure per consumer	Standard error, SE($\bar{y}$)	Coefficient of variation, CV ($\bar{y}$) (in percent)
See footnotes at end of table.			
Total expenditures	$51,100	$520	1.02
Food	6,602	100	1.51
Housing	17,148	171	1
Apparel	1,604	36	2.26
Transportation	9,004	148	1.64
Healthcare	3,631	53	1.47
Entertainment	2,482	49	1.98
Personal care	608	15	2.46
Cash contributions	1,834	97	5.29
Personal insurance and pensions	5,528	116	2.09
Other	2,659	59	2.21

Source: U.S. Bureau of Labor Statistics.

Last Modified Date: February 25, 2016

Presentation

Information from the Consumer Expenditure Survey (CE) is available in press releases, reports, and analytical papers. Public use microdata from the survey from 1996 to the present are available for free download on the CE section of the BLS website (https://www.bls.gov/cex), and older microdata are available for purchase on USB flash drives. For microdata information and to purchase USBs, see: https://www.bls.gov/cex/pumd.htm. Tabular data also are available at the same location on the BLS website and by contacting the BLS Consumer Expenditure Survey Division directly at CEXInfo@BLS.gov or 202-691-6900.

Publications from the CE generally include tabulations of average expenditures and income arrayed by consumer unit characteristics, such as consumer unit size, age of reference person, or income. Tabulations by two variables (cross-tabulations) are available for selected characteristics, such as age by income and consumer unit size by income.

Integrated Diary and Interview Survey data tables covering 12 months of data are published on a twice a year basis, and tabulations for 1960–61, 1972–73, and 1984 onward are available on the BLS website.

The Diary and Interview Survey public use microdata contain files of expenditure and income reports of each consumer unit. To protect the identities of respondents, selected geographic detail is eliminated, and selected income and expenditure variables may be topcoded. Topcoding refers to a confidentiality protection method were a subset of extremely high or low values are averaged together and the original values are replaced with the average amount. The Interview Survey files contain expenditure data in two formats: MTAB files that present monthly values in an item-coding framework, based on the CPI pricing scheme; and EXPN files that organize expenditures by the section of the Interview Survey instrument in which they are collected. Expenditure values on the EXPN files cover different time periods, depending on specific questions asked; these files also contain relevant non-expenditure information not found on the MTAB files. The public use microdata files include quarterly expenditure summary variables at the consumer unit level. The annual Interview and Diary Survey microdata files are available beginning with 1990, as well as for selected earlier years.

Each year, the CE conducts a free summer public use microdata users' workshop to provide training on how to work with the microdata files. See: https://www.bls.gov/cex/csxannualworkshop.htm.

Articles that include analyses of CE data are published online in the *Monthly Labor Review* (*MLR*), in the quarterly publication *Beyond the Numbers*, and in CE data comparisons and research reports. Other survey information is available on the Internet, including answers to frequently asked questions, copies of the Interview and Diary Survey instruments, a glossary of terms, and order forms for survey products. Starting with the 2000 data, estimates of standard errors for integrated Diary and Interview Survey data are available on the BLS website.

More detailed expenditure tables are available upon request. These tables are sorted by the same demographic variables as the standard tables on the CE website, but have more expenditure subcategories. Estimates for these

subcategories have higher variance than the standard published tables. These detailed tables are not online, but will be emailed upon request.

Last Modified Date: February 25, 2016

History

Timeline Events:

1888–1891: The first nationwide expenditure survey conducted to study workers' spending patterns as elements of production costs

1901: The second nationwide expenditure survey conducted in response to rapid price changes prior to the turn of the 20th century. It provided the weights for an index of food prices purchased by workers

1917–1919: The third expenditure survey conducted. It provided weights for computing a cost-of-living index, now known as the Consumer Price Index (CPI)

1934–1936: Expenditure data collected from only urban wage and clerical workers used in revising CPI weights

1935–1936: The first-ever nationwide rural and urban expenditure survey was collected

1941–1942: Urban and rural expenditure survey conducted during World War II to measure domestic household expenditures during wartime

1944: The wartime expenditure survey repeated for just urban households

1950: Expenditure survey conducted for urban households

1960–1961: Expenditure survey for both urban and rural households conducted

1972–1973: First survey collected by the Census Bureau for BLS. Began the first use of two collection instruments: a weekly Diary Survey and the 3-month recall Interview Survey

End of 1979: The CE begins annual data collection in urban and rural areas. Annual collection continues to the present

1984: Beginning of annual published data tables

2002: CPI begins using biennial Consumer Expenditure (CE) weights to update CPI cost weights every 2 years instead of updating the CPI every 10 years using 3 years of CE expenditures

2003: Computer Assisted Personal Interview (CAPI) starts

2004: Introduction of Imputed Income to fill in all missing income values

2005: Introduction of a user-friendly Diary Survey

2009: Gemini CE Redesign long-term research begins. The primary mission of the Gemini Project is to improve data quality through a verifiable reduction in measurement error, with a particular focus on underreporting

March 2013: First publication of midyear data every 12 months

2013: Introduction of estimated federal and state income taxes with the published 2013 data tables. Replaced all collected and missing amounts with estimated amounts

2015: Noninterview adjustment calculations include income as a weighting variable

2015: Initial 'Bounding' Interview Survey dropped. Number of interviews per household drops from five to four

The Bureau's studies of family living conditions rank among its oldest data-collecting functions. The first nationwide expenditure survey was conducted during 1888–1891 to study workers' spending patterns as elements of production costs. With special reference to competition in foreign trade, the survey emphasized the worker's role as a producer, rather than as a consumer. In response to rapid price changes prior to the turn of the 20th century, a second survey was administered in 1901. The resulting data provided the weights for an index of prices of food purchased by workers that was used until World War I as a deflator for workers' incomes and expenditures. A third survey, conducted during 1917–19, provided weights for computing a cost-of-living index, now known as the Consumer Price Index. The Bureau conducted its next major survey, covering only urban wage earners and clerical workers, during 1934–36, primarily to revise CPI weights.

The Great Depression

During the Great Depression of the 1930s, the use of consumer expenditure surveys extended from the study of the welfare of selected groups to more general economic analysis. Concurrent with its 1934–36 investigation, the Bureau cooperated with four other Federal agencies in a fifth survey, the 1935–36 study of consumer purchases, which presented consumption estimates for both urban and rural segments of the population.

World War II

During World War II, a 1941–42 survey for urban, nonfarm, and farm households was conducted, with a followup survey of only urban households collected in 1944. See: https://www.bls.gov/opub/mlr/2015/article/consumer-spending-in-world-war-ii-the-forgotten-consumer-expenditure-surveys.htm.

The next survey in 1950, which covered only urban consumers, was an abbreviated version of the 1935–36 study. The 1950 CE data were used in the 1953 CPI revision.

Other significant improvements

The 1960–61 Survey of Consumer Expenditures once again included both urban and rural families and provided the basis for revising the CPI weights, while supplying material for broader economic, social, and market analyses.

Eleven years later, the next survey collecting information on expenditures of urban and rural householders in the United States was conducted in 1972–73. That survey, while providing continuity with the content of the Bureau's previous surveys, departed from the past in its collection techniques.

Unlike the previous surveys, the U.S. Census Bureau, under contract to BLS, conducted all sample selection and field work. Another significant change was the use of two independent surveys to collect the information—a diary survey and an interview panel survey. A third major change was the switch from an annual recall to a quarterly recall in the Interview Survey, and daily recordkeeping of expenditures in the Diary Survey. Again, the resulting data were used to revise CPI weights.

The need for more timely data than could be supplied by surveys conducted every 10 to 12 years—intensified by the rapidly changing economic conditions of the 1970s—led to the initiation of the current continuing survey in late 1979. From 1984 to 2011, annual calendar year data tables have been available. Beginning with July 2012 data, both annual data plus a second set of 12-month tables covering the second half of one published year with the first half of the subsequent year have been published. The objectives of the CE remain the same: to provide the basis for revising weights and associated pricing samples for the CPI and to meet the need for timely and detailed information on the spending patterns of different types of families. Like the 1972–73 survey, the current survey consists of two separate surveys, each with a different data collection technique and sample. However, as described in the more information section, BLS began a project in 2009 to look at the need for a complete redesign of the survey.

Last Modified Date: February 25, 2016

More Information

Survey methods research

Consumer expenditure surveys undergo continuous evaluation, by comparing results with other sources and by performing internal statistical, qualitative, and cognitive analyses to address current methodological concerns. To improve expenditure estimates, the Consumer Expenditure Survey (CE) program began research that was related to the data collection instruments, field procedures, and sources of potential survey error in the mid-1980s, and has since become standard practice. In 1999, BLS established a separate Branch of Research and Program Development (BRPD) within the Division of Consumer Expenditure Survey, with the mission of developing and conducting methodological studies to improve survey instruments, field procedures, and overall survey data quality. In recent years, BRPD has focused on three core areas: the Gemini Project to redesign the survey, analyzing historical data in support of ongoing methodological improvements, and field testing alternative data collection methods.

The Gemini Project. The BLS began the Gemini Project in 2009 with a goal of redesigning the CE. The Gemini Project was created in response to increasing evidence of measurement error, declining response rates, the emergence of new data collection technologies, and the need for more flexibility in addressing changes in the interviewing environment. The primary mission of the Gemini Project is to improve data quality through a verifiable reduction in measurement error, with a particular focus on underreporting. Early stages of the project focused on gathering facts to inform redesign decisions. This included conducting and reviewing research on survey methodologies and prioritizing user needs.

In addition, in 2010 the BLS contracted with the Committee on National Statistics (CNSTAT) to convene an expert panel charged with recommending different CE design options that would meet the project goals. The CNSTAT panel presented three alternate designs in September 2012. In 2013, the CE program approved a comprehensive redesign proposal based on 3 years of information gathering, inquiry, and synthesis, including a review of the CNSTAT recommendations. The redesign proposal meets key stakeholder requirements and addresses three factors believed to affect the survey's ability to collect high quality data; specifically, measurement error, environmental changes, and flexibility. For further information on the Gemini Project, including information about current research studies and the project's timeline, see https://www.bls.gov/cex/geminiproject.htm.

Research overview. BRPD conducts ongoing research, both in support of the redesign effort, and as an effort to improve data quality while balancing survey costs. Current research has focused on analyzing historical data in support of methodological improvements, and field testing alternative data collection methods. The first area is useful for reviewing the existing survey protocol and considering the potential impact of design changes. The second area provides empirical insight for decisions on implementing future protocol improvements. Details about ongoing and recently completed research projects are provided in CE's annually released Methods Research Program, accessible from https://www.bls.gov/cex/research_papers/research-paper-catalog.htm. In addition, the September 2013 *Monthly Labor Review* article "Research Highlights of the Consumer Expenditure Survey Redesign," https://www.bls.gov/opub/mlr/2013/article/ce-survey-redesign.htm, provides highlights of recent

research projects, along with a summary of major research project findings as they relate to the objectives of the survey redesign process.

Other ongoing survey improvements

In a collaborative effort headed by the CE branch of Production and Control involving the different CE branches and divisions, there are regular biennial Interview questionnaire revisions and other improvements. These improvements include adding new products into the survey, deleting outdated wording or categories, improving non-interview adjustment through the inclusion of income data at the zip code level, using TAXSIM in order to provide estimated income taxes, dropping the initial bounding interview, and publishing new tables.

Last Modified Date: February 25, 2016