Chapter 16. Consumer Expenditures and Income

Consumer expenditure surveys are specialized studies in which the primary emphasis is on collecting data related to family expenditures for goods and services used in day-to-day living. Consumer expenditure surveys of the Bureau of Labor Statistics (BLS, the Bureau) also collect information on the amount and sources of family income, changes in savings and debts, and demographic and economic characteristics of family members.

IN THIS CHAPTER

Background

The Bureau's studies of family living conditions rank among its oldest data-collecting functions. The first nationwide expenditure survey was conducted in 1888–91 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 century, a second survey was administered in 1901. The resulting data provided the weights for an index of prices of food purchased by workers, which was used as a deflator for workers' incomes and for expenditures for all kinds of goods until World War I. A third survey, conducted in 1917–19, provided weights for computing a cost-of-living index, now known as the Consumer Price Index (CPI). The next major survey, covering only urban wage earners and clerical workers, was carried out in 1934–36, primarily to revise the CPI weights.

During the economic 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 estimates of consumption for both urban and rural segments of the population. The sixth survey, in 1950, was an abbreviated version of the 1935–36 study, covering only urban consumers. The seventh survey, the 1960–61 Survey of Consumer Expenditures, which once again included both urban and rural families, provided the basis for revising the CPI weights and also supplied material for broader economic, social, and market analyses.

The next major survey to collect information on expenditures of 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 way in which earlier surveys were carried out, the Census Bureau, under contract with the Bureau of Labor Statistics, conducted all sample selection and fieldwork. Another significant change was the use of two independent surveys—a Diary survey and an Interview panel survey—to collect the information. Yet 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 the CPI weights.

The Current Survey

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. Since then, data have been available annually. The objectives of the survey remain the same: To provide the basis for revising the 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. In the Interview survey, each family

in the sample is interviewed every 3 months over five calendar quarters. The sample for each quarter is divided into three panels, with consumer units being interviewed every 3 months in the same panel of every quarter. The Diary (or recordkeeping) survey is completed at home by the respondent family for two consecutive 1-week periods.

The sample housing unit is notified in advance by a letter informing the occupants about the purpose of the survey and the upcoming visit by the interviewer. Both the Interview and the Diary survey are conducted by personal visits, with telephone usage limited to scheduling appointments and placing followup calls to gather information missed at the time of the proposed interview. The interviewer uses a structured questionnaire to collect both the demographic and expenditure data in the Interview survey. The demographic data in the Diary survey are collected by the interviewer, whereas the expenditure data are entered on the diary form by the respondent. If, after a few attempts to contact the respondent, no adult is available, both surveys accept proxy responses from any eligible household member who is at least 16 years old.

The unit for which expenditure reports are collected is the set of eligible individuals constituting a consumer unit, which 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, in a 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 also are included in the sample as separate consumer units.

Survey participants record dollar amounts for goods and services purchased during the reporting period, whether payment was or was not made at the time of purchase. The expenditure amounts include all sales and excise taxes for all items purchased by the consumer unit for itself or for others. Excluded from both surveys are all business-related expenditures and expenditures for which the family is reimbursed.

The Interview survey collects detailed data on an estimated 60 to 70 percent of total family expenditures. In addition, global estimates—that is, estimated average expenditures for a 3-month period—are obtained for food and other selected items. These global estimates account for an additional 20 to 25 percent of total expenditures. On average, it takes approximately 90 minutes to complete the interview.

In the Diary survey, detailed data are collected on all expenditures made by consumer units during their participation in the survey. It is estimated that it takes approximately 15 minutes per visit during each of the three visits for the interviewer to collect the demographic data and to instruct the respondent on how to keep the diary. It also is estimated that it takes the respondent about 15 minutes each day to complete the diary.

Quality control is provided by a reinterview program, which constitutes a means of evaluating the performance of the individual interviewer to determine how well the procedures are being carried out in the field. The reinterview is conducted by a member of the supervisory staff. A subsample of approximately 6 percent of households in the Interview survey and 17 percent in the Diary survey are reinterviewed 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. All employees have taken an oath to that effect.

Interview survey

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 either are relatively large, such as expenditures for property, automobiles, or major appliances, or occur on a fairly regular basis, such as expenditures for rent, utility bills, or insurance premiums.

Each occupied sample unit is interviewed once per quarter for five consecutive quarters. After the fifth interview, the sample unit is dropped from the survey and replaced by a new consumer unit. For the survey as a whole, 20 percent of the sample is replaced each quarter. Data collected in each quarter are considered independently, so that estimates are not dependent upon a family participating in the survey for a full five quarters.

For the initial interview, information is collected on demographic and family characteristics and on the inventory of major durable goods of each consumer unit. Expenditure information also is collected in this interview, based on a 1-month recall period, but is used, along with the inventory information, solely for bounding purposes—that is, to classify the unit for analysis and to prevent duplicate reporting of expenditures in subsequent interviews.

The second through fifth interviews use uniform questionnaires to collect expenditure information in each quarter. Data collected in these questionnaires, which are arranged by major expenditure component (for example, housing, transportation, medical, and education), 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 also is collected or updated during each of these interviews. The expenditure data are collected via two major types of questions. The first asks directly for the month of purchase for each reported expenditure. The second type of question asks for a quarterly amount of expenditures. The use of these two types of questions varies, depending on the types of expenditures collected. Approximately 65 percent of the data are collected with the use of the direct monthly method, whereas about 35 percent are collected with the quarterly recall approach.

In the fifth and final interview, an annual supplement is used to obtain a financial profile of the consumer unit. The profile consists of information on the income of the con-

sumer unit as a whole, including unemployment compensation; income from royalties, dividends, and estates; alimony and child support; and so forth. A 12-month recall period is used to collect income- and asset-type data.

Diary survey

The primary objective of the Diary survey is to obtain detailed expenditure data on small, frequently purchased items, because such data normally are difficult to recall. These items include food and beverages, both at home and in eating places; housekeeping supplies and services; nonprescription drugs; and personal care products and services. The Diary survey is not limited to these types of expenditures, but, rather, includes all expenses that the consumer unit incurs during the survey week. Expenses incurred by family members while away from home overnight and for credit and installment plan payments are excluded.

Two separate questionnaires are used to collect Diary 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 the twin purposes of publication of statistical tables and of economic analysis. Data on household characteristics also provide the link in the integration of Diary expenditure data with Interview expenditure data that permits 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 on which respondents record a detailed description of all expenses for two consecutive 1-week periods. Data collected each week are considered independently. The diary is divided by day of purchase and by broad classifications of goods and services—a breakdown designed to aid the respondent in recording the entire consumer unit's daily purchases, as well as household utilities and other expenses. 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.

Integrated survey data

The 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. Some expenditure items are collected only by either the Diary or the Interview survey. For example, the Diary collects data on detailed food expenditures and items such as postage and nonprescription drugs, which are not collected in the Interview. The Interview collects data on expenditures for overnight travel and information on reimbursements, such as for medical-care costs or automobile repairs, which are not collected in the Diary. 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 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 this 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 surveys, the more reliable of the two, as determined by statistical methods, is selected. The selection of the source survey is evaluated periodically.

Processing

Due to differences in format and design, Diary and Interview survey data are processed separately. Diary questionnaires are coded and keyed at the Census Bureau Processing Center in Jeffersonville, IN, and are then transmitted monthly to BLS for further editing and processing. Missing or invalid data on demographic or work experience questions are imputed. No imputation is done for missing data on expenditures or income. The households that are surveyed are assigned weights, so that estimates can be derived that represent the total civilian noninstitutional population.

BLS combines the monthly data into separate databases that relate to calendar quarters. The data on these quarterly tapes are screened selectively for invalid coding and inconsistent relationships, as well as for extreme values that might affect the reasonableness of estimates after the data are aggregated. All errors of coding or extreme value are corrected before further processing.

Selected portions of the Diary data also are adjusted by automated imputation and allocation routines when respondents report insufficient detail to meet publication requirements. These routines (imputation and allocation) are performed annually on the data. The imputation routines assign qualifying information to data items when there is clear evidence of invalid nonresponses. For example, the qualifiers classify food expenditures by type of processing (that is, fresh or frozen) and apparel expenditures by age and sex groupings of the members in the consumer units. Allocation routines are a means of transforming nonspecific reports of items into specific ones. For example, when respondents report expenditures for meat rather than beef or pork, proportions derived from specific reports in other completed diaries for the same year are used to distribute the expenditures reported for meat between beef and pork.

Census Bureau processing of the Interview Survey data is minimal, once Computer-Assisted Personal Interview (CAPI) data collection is complete. There is no separate process required for keying and coding the data, because this is done as the data are collected. The data are transmitted directly to the Census Bureau headquarters, where the information needed for the next quarter's interview is captured

and stored and where output files are created for BLS.

The monthly data files are then transmitted to BLS, where they go through a series of complex computer edits and adjustments that include the identification and correction of irregularities and inconsistencies throughout the questionnaire. Other adjustments convert mortgage and vehicle payments into principal and interest (given associated data on the interest rate and term of the loan), eliminate business and other reimbursed expenses, apply appropriate sales taxes, and derive weights for individual questionnaires. In addition, demographic and work-experience items (except income) are imputed when missing or invalid.

The Bureau then constructs quarterly databases and conducts an extensive review to ensure that severe data aberrations are corrected. The review takes place in several stages: A review of counts and means by region; a review of coding of family relationships for inconsistencies; a review of selected extreme values for expenditure and income categories; and a verification of the various data transformations performed by the Bureau. Errors are corrected prior to release of the data for public use.

In the Interview survey, data imputation routines are carried out when the respondent has indicated the occurrence of an expenditure, but has failed to specify a dollar amount. The procedures are performed quarterly on the data and are intended to improve the estimates derived from the survey. Imputation in the Interview survey is done at the cell level, with cells defined by variables such as income class, family size, region, and so on. The methods used—hot deck, weighting class, and percent distribution—depend on the types of expenditures. In addition, allocation routines are applied to the Interview data in a fashion similar to that used for the Diary data.

Sample Design

Selection of households

The Consumer Expenditure Survey is a national probability sample of households designed to represent the total U.S. civilian noninstitutional population. The selection of households begins with the definition and selection of primary sampling units (PSUs), which consist of counties (or parts thereof), groups of counties, or independent cities. The sample of PSUs used for the survey consists of 105 areas, of which 87 urban areas also have been selected for the BLS Consumer Price Index program.

The 105 PSUs are classified into four categories:

- 31 *"A" PSUs*, which are Metropolitan Statistical Areas (MSAs) with a population of 1.5 million or greater
- 46 *"B" PSUs*, which are MSAs with a population less than 1.5 million
- 10 *"C" PSUs*, which are nonmetropolitan areas used in the CPI
- 18 *"D" PSUs*, which are nonmetropolitan areas not used in the CPI

Within these PSUs, the sampling frame (that is, the list from which housing units are chosen) for the Consumer Expenditure Survey is now generated from the 1990 census 100-percent detail file. The frame is augmented by a sample drawn from new construction permits and extra housing units identified through improvements in coverage techniques. Beginning in 2005, the 2000 census address file will be used.

The population represented by the survey is the total U.S. civilian noninstitutional population, including persons living in houses, condominiums, apartments, and group quarters such as college dormitories. Excluded are military personnel living on base and nursing home residents, as well as people in prisons or other institutions.

The Census Bureau selects a sample of approximately 12,500 addresses for participation in the Diary survey each year. Usable diaries are obtained from approximately 7,700 households at these addresses. Some diaries are not completed due to refusals, vacancies, ineligibility, or the nonexistence of a housing unit at the selected address. The actual workload of diaries is spaced over 52 weeks of the year.

The Interview survey is a rotating panel survey in which approximately 15,000 addresses are contacted in each calendar quarter of the year. One-fifth of the addresses contacted each quarter are new to the survey and provide the bounding interviews that afford baseline data, but are not used to compute the survey's published expenditure estimates. Excluding these bounding interviews, as well as interviews not completed due to refusals, vacancies, ineligibility, or the nonexistence of a housing unit at the selected address, usable interviews are obtained from approximately 7,600 households each quarter. After a housing unit has been in the sample for five consecutive quarters, it is dropped from the panel and a new housing unit is selected to replace it.

Cooperation levels

Data on responses to the 2001 Consumer Expenditure Survey are shown in table 1. For the Interview survey, the totals refer to housing units in the second through fifth quarters of the survey (the nonbounding interviews), with each unique housing unit providing up to four usable interviews. For the Diary survey, the totals refer to housing units, with each unique housing unit affording up to two usable interviews during the two consecutive 1-week periods. Most Diary respondents participate for both weeks.

There are three general categories of nonresponses:

- *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

Table 1. Analysis of responses in the Consumer Expenditure Survey, 2001

Sample unit	Interview survey	Diary survey
Housing units designated for the survey	47,998	25,600
Less Type B or C nonresponses	9,158	5,037
Equals eligible units	38,840	20,563
Less Type A nonresponses	8,527	5,159
Equals Interview units	30,313	15,404
Percent of eligible units interviewed	78.0	74.9

- *Type C nonresponses* are destroyed or abandoned housing units, and housing units converted to nonresidential use

Response rates are defined to be the percentage of eligible housing units (that is, the designated sample, less Type B and Type C nonresponses) from which usable interviews are collected. In the 2001 Interview survey, there were 38,840 eligible units, from which 30,313 usable interviews were collected—a response rate of 78.0 percent. In the 2001 Diary survey, there were 20,563 eligible units, from which 15,404 usable interviews were collected—a response rate of 74.9 percent.

Estimation methodology

The estimation of population quantities of interest, such as the average expenditure per consumer unit on a particular item, is achieved through the use of weights. Each consumer unit included in the survey is assigned a weight, which is interpreted as the number of similar consumer units in the U.S. civilian noninstitutional population that the sampled consumer unit represents. With the use of these weights, the average expenditure per consumer unit on a particular item category is estimated by

$$\bar{y} = \frac{\sum_{i \in s} w_i y_i}{\sum_{i \in s} w_i},$$

where

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

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

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

s = sample of consumer units that participate in the survey.

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

To estimate the proportion of consumer units that purchased butter during a given period, the same formula is applied, where y_i is set equal to 1 if the ith consumer unit purchased butter during the period and 0 if it did not. When this 1/0 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 butter during the given period.

Several factors are involved in computing the weight of each sampled consumer unit for which a usable report is received. Each such unit is initially assigned a *base weight,* which is equal to the inverse of the consumer unit's probability of being selected for the sample. The base weight is then adjusted by the following factors in order to correct for certain nonsampling errors:

Weighting control factor. This factor 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 factor adjusts for interviews that cannot be conducted in occupied housing units due to a consumer unit's refusal to participate in the survey or because no one is home. The adjustment is based on region of the country, household tenure (owner or renter), size of the consumer unit, and race of the reference person.

Calibration factor. This factor both adjusts the weights to known population counts in order to account for frame undercoverage and reduces variances through poststratification. The weights are calculated with the use of a model-assisted, design-based regression estimator. The adjustment is made by means of the following auxiliary variables in the regression model: Age, race, region, urban areas within region, and household tenure. The same auxiliary variables and an intercept are used in both the Diary and the Interview surveys, resulting in equal counts of consumer units between the Diary and the Interview in those domains and in the total population. Extremely large weights and extremely small weights are moderated by constraints placed on the allowed departure of the final weight from the initial weight (initial weight = base weight × weighting control factor × noninterview adjustment factor) of each consumer unit.

Precision of the estimates

The precision of the estimator $\bar{y}$ is measured by its standard error. Standard errors measure the sampling variability of the Consumer Expenditure Survey estimates. (See table 2.) That is, they 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 data collected from every consumer unit in the Nation.

Standard errors are estimated by using the balanced repeated replication method of variance estimation. In this method, the sampled PSUs are divided into 40 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 by randomly selecting 1 of the 2 half samples from each stratum. All 44 of these replicate estimates of $\bar{y}$ use different combinations of half samples from the 40 strata. The standard error of $\bar{y}$ is estimated by the equation

$$\mathrm{SE}(\bar{y}) = \sqrt{\frac{1}{44}\sum_{r=1}^{44}(\bar{y}_r - \bar{y})^2} \quad ,$$

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

The coefficient of variation is a related measure of the variability of a sample. This factor, which measures the variability of the survey estimate relative to the mean, is defined by the equation

$$\mathrm{CV}(\bar{y}) = \frac{\mathrm{SE}(\bar{y})}{\bar{y}} \times 100$$

and usually is expressed as a percent.

Presentation

Information from the Consumer Expenditure Survey is available in press releases, reports, and analytical papers and on CD-ROMs. Data also are available from the BLS Web site on the Internet (**http://www.bls.gov/cex**) and from the BLS fax-on-demand service. Publications may be obtained through the BLS Office of Publications and Special Studies or the BLS Division of Consumer Expenditure Surveys.

Publications from the Consumer Expenditure Survey generally include tabulations of average expenditures and income, arrayed by consumer unit characteristics, such as size, age of reference person, or income. Tabulations by two variables (cross tabulations) are available for selected characteristics, such as age by income or size by income. Integrated Diary and Interview Survey data are published on an annual basis, and tabulations are available on the BLS Web site back to 1984.

The Diary and Interview Survey microdata that are available on CD-ROM contain files of the expenditure and income reports of each consumer unit. In order to protect the identities of respondents, selected geographic detail is eliminated and selected income and expenditure variables may be topcoded. The Interview files contain expenditure data in two different 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 questionnaire in which they are collected. Expenditure values on the EXPN files cover different periods, depending on the specific questions asked, and the files also contain relevant nonexpenditure information not found on the MTAB files. The Interview and Diary microdata files are available on CD-ROM back to 1990 and for selected earlier years.

Articles that include analyses of Consumer Expenditure Survey data are published in the *Monthly Labor Review* (*MLR*) and as *Issues in Labor Statistics*. Copies of the most recent articles are posted on the BLS Web site. Other survey information also is available on the Internet, including answers to frequently asked questions, a glossary, and order forms for survey products.

Evaluation Research

The Bureau continuously evaluates the surveys by comparing Consumer Expenditure Survey results with other data and by performing internal statistical analysis. In order to improve the expenditure estimates, research related to the

Table 2. Precision of the Consumer Expenditure Survey expenditure estimates, integrated Diary and Interview survey data, 2000

Item category	Average annual expenditure per consumer unit, $\bar{y}$	Standard error SE ($\bar{y}$)	Coefficient of variation, CV ($\bar{y}$) (in percent)
Food	$5,158	$66	1.28
Housing	12,318	149	1.21
Apparel	1,853	39	2.10
Transportation	7,417	101	1.36
Health care	2,066	30	1.46
Entertainment	1,864	36	1.91
Personal care	564	8	1.41
Cash contributions	1,192	97	8.12
Personal insurance and pensions	3,365	55	1.63
Other	2,244	—	—
All items	**$38,041**	**$337**	**0.89**

data collection instruments and interview procedures began in the mid-1980s. During that time, the Bureau received funding from Congress to create a Survey Design Research Center. Shortly thereafter, in January 1987, a Questionnaire Design Advisory Conference was convened at the Bureau. The recommendations received by the Bureau at this conference, combined with results from the underreporting studies, led to an increase in research into the cognitive aspects of the data collection process. Recent research has placed an emphasis on the use of cognitive techniques for investigating respondents' thought processes in response to survey questions and for developing alternative formats for questionnaires and phrasing for questions. Attention also has been focused on the demands placed upon respondents' time.

A new Diary form with more categories and an expanded use of cues for respondents was introduced in 1991 on the basis of results from earlier field and laboratory studies. In 1998, the Food Away from Home section of the Diary was revised in order to classify prepared-food expenditures according to the types of establishments where such foods are purchased.

Recent research has focused on a variety of issues aimed at reducing the burden on respondents of answering questions and at improving the quality of survey data. These issues include examining the relative domains of the two Consumer Expenditure Survey instruments, finding ways to streamline the Interview survey questionnaire, and making the Diary more "user friendly."

Beginning in April 2003, data collection for the Interview Survey was converted from paper-and-pencil mode to CAPI. The questionnaires are administered by the interviewer, who, using a laptop computer, is able to perform basic consistency checks and range edits as the data are entered. For interviews 2 through 5, previously reported inventoried items, such as vehicles, properties, and insurance policies, are displayed for updating as needed. Because of the complexity of the skip patterns in the paper questionnaire, CAPI has increased accuracy and reduced the amount of postcollection data processing.

Uses and Limitations

The importance of the Consumer Expenditure Survey is that it allows data users to relate the expenditures and income of consumers to the characteristics of those consumers. The 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. The survey data are used by economic policymakers who are interested in the effects of policy changes on levels of living among diverse socioeconomic groups. Econometricians find the data useful in constructing economic models. Market researchers find them valuable in analyzing the demand for groups of goods and services. The U.S. Department of Commerce uses the survey data as a source of information for revising its benchmark estimates of some of the personal consumption expenditure components of the gross national product.

As in the past, the revision of the CPI remains a primary reason for undertaking such an extensive survey. Output from the Consumer Expenditure Survey has been used to select new market baskets of goods and services for the CPI, to determine the relative importance of the index's components, and to derive new cost weights for the baskets. In August 2002, BLS published a new index called the "Chained Consumer Price Index for All Consumers" (C-CPI-U), which is a supplement to the existing consumer price indexes. The use of expenditure data from different periods distinguishes the C-CPI-U from the existing CPI measures, which use only a single expenditure base period to compute changes in price over time. The new index is designed to measure the change in the "cost of living," in contrast to the existing indexes, which are designed to measure the change in the price of a fixed market basket of goods and services in retail outlets. The C-CPI-U uses expenditure data from different 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.

Sample surveys are subject to two types of error, sampling error and nonsampling error. *Sampling error* is the uncertainty caused by the fact that observations are taken from a sample and not from the entire population. *Nonsampling error* is the rest of the error, and can be attributed to many sources, such as differences in the interpretation of questions, an inability or unwillingness on the part of respondents to provide correct information, data-processing errors, and so on. Nonsampling error arises regardless of whether data are collected from a sample or from the entire population.

Another way of analyzing error is to divide it into variance and bias. The *variance* is a measure of how close different estimates would be to each other if it were possible to repeat the survey over and over, using different samples. While it is not feasible to repeat the survey in such a manner, statistical theory allows the variance to be estimated anyway. A small variance indicates that multiple independent samples would produce values that are consistently very close to each other. *Bias* is the difference between the "expected" value of an estimate and its "true" value. A statistic may have a small variance, but a large bias, or it may have a large variance, but a small bias. For an estimate to be considered accurate, both its variance and its bias must be small.

The Bureau is constantly trying to reduce the error in the Consumer Expenditure Survey estimates. Variance and sampling error are reduced by using a sample of respondents that is as large as possible, given resource constraints. Improving the accuracy of the estimates was the primary reason for the significant expansion in the sample for both the Interview and Diary surveys that occurred in 1999. The Bureau reduces nonsampling error by means of a series of computerized and professional data reviews, as well as through continuous improvements in the survey process and through theoretical research.