

Appendix B. Integration and Survey Methods

The U.S. Bureau of Labor Statistics (BLS or the Bureau) has been gathering information on the spending patterns and living costs of American consumers for more than a century—since the first such survey in 1888–91. Survey methods have been improved and refined over the years. A major methodological improvement, first used in the 1972–73 survey, was the introduction of two separate surveys—a quarterly interview survey and a weekly diary or recordkeeping survey—rather than a single interview survey, relying primarily on annual recall by survey participants. The Bureau added a further enhancement in 1980, when it began data collection for the survey on a continuing basis, rather than at intervals of about 10 years.

The Bureau designed the surveys so that each has its own questionnaire and sample. For the quarterly Interview Survey, an interviewer visits every consumer unit in the sample every 3 months over a 12-month period. It was designed to obtain data on the types of expenditures respondents can be expected to recall for a period of 3 months or longer. For the Diary Survey, consumer units are asked to complete a record of expenses for two consecutive 1-week periods. It was designed to obtain detailed data on frequently purchased small items, such as food and beverages (both at home and in eating places).

Integrating data from the Interview and Diary Surveys provides a complete accounting of expenditures and income, which neither survey component alone is designed to do. Expenditure levels and expenditure shares (the percent of the total spent on each category) shown in this report result from integrating the Diary and Interview Survey data.

Description of survey

BLS contracts with the U.S. Census Bureau to carry out data collection for both surveys. In the Interview Survey, a U.S. Census Bureau field representative meets with respondents and collects expenditure and income data via a computer-assisted personal interview questionnaire. In the Diary Survey, respondents are asked to report all expenditures made during their 2-week participation in the survey. Expenditures and related data are recorded in a self-reporting record of daily living expenses. All data collected in both surveys are subject to confidentiality requirements that prevent the disclosure of respondents' identities or such geographic identifiers that may lead to their identification.

In addition to the Interview Survey questionnaire and the Diary Survey record of daily expenses, information pertaining

to age, sex, race, marital status, and family composition is collected from each survey respondent in a Household Characteristics Questionnaire. The questionnaire also asks for information on work experience, occupation, industry, retirement status, and income. Income includes member earnings from wages and salaries, net income from a business or profession, net income from a farm, and income from all other sources. Data on household characteristics are collected to determine the eligibility of the family for inclusion in the population covered by the Consumer Price Index (CPI), to classify families by family type for purposes of analysis, and to adjust for nonresponse by families who do not cooperate in the survey. The data also provide the link between the Diary and Interview Surveys to permit the integration of the data by demographic characteristics.

Quarterly Interview Survey. The quarterly interview portion of the survey is designed to collect data on major items of expense, household characteristics, and income. The survey covers expenditures that one would expect respondents to recall for 3 months or longer, such as those for property, automobiles, and major appliances, and those that occur on a regular basis, such as rent, insurance premiums, and utilities. The survey includes detailed data on an estimated 60 to 70 percent of total household expenditures. In addition, global estimates, that is, expense patterns for a 3-month period, are obtained for food and other selected items, accounting for an additional 20 to 25 percent of total expenditures. Each sample household is interviewed once per quarter, for five consecutive quarters. Data collected in each quarter are estimated independently, so annual estimates do not depend upon the participation of a consumer unit for the full five quarters.

New panels are introduced into the interview sample on a monthly basis, as other panels complete their participation. For the Interview Survey as a whole, 20 percent of the sample is dropped, and a new group added each quarter. This rotating procedure allows panel estimates to reflect population changes; it also provides operational efficiency by distributing interviewer workload across time.

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 is also collected in this interview, with a 1-month recall. Expenditure information is used, along with the inventory information for bounding purposes to minimize telescoping errors. These double counting errors, common in

retrospective interviews, result from a tendency to report past events in the reference period of the survey.

The second through fifth interviews use uniform questionnaires to collect expenditure information in each quarter. In the second and fifth interviews, information also is obtained on income, such as wage and salary earnings, unemployment compensation, child support, and alimony, as well as information on the employment of each household member. For new consumer unit members and members who started work since the second interview, the interviewers ask for wage, salary, and other information on employment in the third and fourth interviews. If there is no new employment information, information is carried over from the second interview to the third and fourth. In the fifth interview, a supplement is used to collect changes in assets and liabilities.

Households that move away from the sample address between interviews are dropped from the survey. New households that move into the sample address are screened for eligibility and, if found qualified, are included in the survey.

Diary Survey. The Diary portion of the survey collects expenditure data for small items purchased on a daily or weekly basis, such as food, beverages, food consumed away from home, housekeeping supplies, nonprescription drugs and medical supplies, and personal care products and services. However, participants are asked to record all purchases made each day for two consecutive 1-week periods. Respondents receive each weekly diary during a separate visit by a Census Bureau interviewer. Data collected in each week are used independently in the annual estimates, so participation of a consumer unit for both weeks is not required. However, most respondents participate for both weeks.

Beginning in 2004, the Consumer Expenditure Survey (CE) implemented multiple imputation of income to provide estimated values for missing income data. The process is applied to both Interview and Diary Survey data. Data from 2004 onward are not strictly comparable to data from previous years. For an explanation of income imputation, see “Changes to the 2004 and 2005 Consumer Expenditure Survey published tables and selected highlights” of the 2004–2005 CE biennial report.

Source selection

The Diary and Interview Surveys overlap considerably in their coverage of household expenditures. As mentioned above, the Interview Survey is designed to obtain data on major expenditures such as automobiles and major appliances that respondents can be expected to recall for 3 months or longer, while the Diary Survey is designed to obtain data on small frequently purchased items such as food and beverages. When data for an item category are collected in both surveys, BLS uses the data from only one of the surveys in its published expenditure estimates. The survey with more statistically reliable data is chosen. The process of choosing the survey with more statistically reliable data is called “source selection.”

For source selection, the item categories are divided into three groups:

- Those that are collected solely in the Diary Survey, such as detailed food expenditures, personal care products, postage, housekeeping supplies, and nonprescription drugs.
- Those that are collected solely in the Interview Survey, such as expenditures made on trips (overnight and longer), and information on reimbursed expenditures, such as for medical care and automobile repairs.
- Those that are collected in both the Diary and Interview Surveys, and where a sufficient amount of detail is collected in both of them. For this group of items, the survey with more statistically reliable data must be determined.

In the past, the source selection procedure was run periodically every few years; since 2007, it will be run every 2 years. The source selection procedure is run separately for each item category in the third group. The item categories are called “universal classification codes” (UCCs), which are members of a 6-digit coding scheme that classifies expenditures. These are the lowest level at which CE expenditures are tabulated. The Consumer Expenditure Survey has approximately 900 expenditure UCCs, of which 275 are in the first group, 375 are in the second group, and 250 are in the third group.

For the third group, it is first determined whether there is a statistically significant difference between the two surveys. If there is, then the survey with the larger mean annualized expenditure per consumer unit is selected as the source. If there is not, the survey currently used as the source continues to be used.

The procedure was most recently run in 2007, using data from 2004 to 2006. It started by counting the number of expenditures in the database for each survey to determine whether credible expenditure estimates could be produced from them. The threshold for credibility used in the procedure was 60 expenditures per year. If there were 60 or more expenditures per year in both surveys, and one of the surveys had a significantly higher mean annualized expenditure, then the survey with the higher mean was selected as the source of data for the published expenditure estimates. If only one of the surveys had 60 or more expenditures per year, then the survey with 60 or more expenditures was selected as the source. Otherwise, the survey currently used as the source was selected again.

The level of statistical significance was determined from the z-score

$$z = \frac{\bar{x}_I - \bar{x}_D}{\sqrt{SE_I^2 + SE_D^2}},$$

where \bar{x}_I and \bar{x}_D are the mean annualized expenditures per consumer unit in the Interview and Diary surveys, and where SE_I and SE_D are their standard errors. The z-scores were computed separately for each of the 3 years, and a weighted average of them was computed, with more weight given to the more recent years. The Interview survey was chosen when the weighted z-score was greater than 1.645 or when the smallest

of the three yearly z-scores was greater than 1.000. The Diary survey was chosen when the weighted z-score was less than -1.645 or when the largest of the three yearly z-scores was less than -1.000.

Sample design

The Consumer Expenditure Survey is a nationwide household survey designed to measure the expenditures of the U.S. civilian noninstitutional population, which comprises about 98 percent of the total U.S. population. CE includes people living in houses, condominiums, apartments, and group quarters, such as college dormitories. It excludes military personnel living on base, nursing home residents, and people in prisons.

The selection of households for the survey begins with the definition and selection of geographic areas called “primary sampling units” (PSUs). PSUs are small groups of counties that have been grouped together into geographic entities called “core-based statistical areas” (CBSAs). The CBSAs are categorized into three groups based on the population of the largest urbanized area inside them: Metropolitan Statistical Areas, which are CBSAs whose largest urbanized area has more than 50,000 people; Micropolitan Statistical Areas, which are CBSAs whose largest urbanized area has between 10,000 and 50,000 people; and non-CBSA areas, which are all remaining areas.

The sample of PSUs used in the 2006–2007 survey consists of 91 PSUs, of which 75 urban PSUs are also used by the Consumer Price Index program. The 91 PSUs are classified into four categories:

- 21 “A” PSUs, which are metropolitan CBSAs with a population over 2.7 million people.
- 38 “X” PSUs, which are metropolitan CBSAs with a population under 2.7 million people.
- 16 “Y” PSUs, which are micropolitan CBSAs.
- 16 “Z” PSUs, which are non-CBSA areas, and are often referred to as “rural” PSUs.

Within these 91 PSUs, the sampling frame (the list of addresses from which the sample is drawn) is generated from the 2000 Census 100-percent detail file. The sampling frame is augmented by new construction permits and extra housing units identified through coverage improvement techniques to compensate for recognized deficiencies in the file.

From this sampling frame, the U.S. Census Bureau selects a representative sample of approximately 12,000 addresses per year to participate in the Diary Survey. Usable diaries are obtained from approximately 7,000 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 actual placement of diaries is spread equally over all 52 weeks of the year.

For the Interview Survey, approximately 14,000 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 “bounding” interviews that provide baseline data, which are not used in the survey’s published expenditure estimates. Excluding these bounding interviews and 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 7,000 households each quarter. After a housing unit has been in the sample for five consecutive quarters, it is dropped from the survey and a new housing unit is selected to replace it.

Response rates

Response data for the 2006–2007 Interview and Diary Surveys are shown in tables B-1 and B-2. For the Interview Survey, totals refer to housing units in the second through fifth quarters of the survey (the non-bounding interviews), with each unique housing unit providing up to four usable interviews. For the Diary Survey, the totals refer to housing units in weeks 1 and 2 of the survey, with each unique housing unit providing up to two usable interviews. Most Diary respondents participate for 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 nonexistent housing units, destroyed or abandoned housing units, and housing units converted to nonresidential use.

Type A nonresponses are considered to be in-scope and eligible units, because they were able to participate in the survey but either chose not to do so or could not be contacted. Type B nonresponses are considered to be in-scope but ineligible units, because the addresses are vacant. Type C nonresponses are considered to be out-of-scope units.

Response rates are defined to be the percent of eligible housing units (i.e., the designated sample less Type B and Type C nonresponses) from which usable interviews are collected. In the 2007 Interview Survey, there were 37,016 eligible housing units, from which 27,335 usable interviews were collected, resulting in a response rate of 73.8 percent. In the 2007 Diary Survey, there were 19,595 eligible housing units, from which 13,747 usable interviews were collected, resulting in a response rate of 70.2 percent.

Table B-1. **Analysis of response in the 2006–07 Interview Survey**

Sample unit	2006	2007
Housing units designated for the survey	46,789	45,996
Less: Type B and C nonresponses ...	9,080	8,980
Equals: Eligible units	37,709	37,016
Less: Type A nonresponses.....	8,842	9,681
Equals: Interview units.....	28,867	27,335
Percent of eligible units interviewed	76.6	73.8

Table B-2. **Analysis of response in the 2006–07 Diary Survey**

Sample unit	2006	2007
Housing units designated for the survey	24,320	24,642
Less: Type B and C nonresponses...	4,844	5,047
Equals: Eligible units	19,476	19,595
Less: Type A nonresponses.....	5,021	5,848
Equals: Interview units.....	14,455	13,747
Percent of eligible units interviewed	74.2	70.2

Weighting

Each consumer unit (CU) in the Consumer Expenditure Survey represents a given number of similar CUs in the U.S. civilian noninstitutional population. The translation of sample CUs into the universe of CUs is known as “weighting.” Several factors are involved in computing the weight of each CU for which a usable interview is obtained. Each CU is initially assigned a base weight, which is equal to the inverse of its probability of being selected for the sample. Base weights in the Consumer Expenditure Survey are typically around 10,000, which means that a CU in the sample represents 10,000 CUs in the U.S. civilian noninstitutional population—itsself plus 9,999 other CUs that were not selected for the sample. The base weight is then adjusted by the following factors:

Weighting control factor. This adjusts for subsampling in the field. Subsampling occurs when a field representative 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 CU’s refusal to participate in the survey or the fact that no one is home. This adjustment is based on the region of the country (Northeast, Midwest, South, West), household tenure (owner/renter), CU size, and race of the CU’s reference person.

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, region of the country, and urban/rural. The population counts come from the Current Population Survey, and they are updated

quarterly. Each CU is given its own unique calibration factor. There are infinitely many sets of calibration factors that make the weights add up to the 24 known population counts, and the set that is used minimizes the amount of change made to the “initial weights” (initial weight = base weight × weighting control factor × noninterview adjustment factor). The same population controls are used for both the Interview and Diary Surveys, hence both surveys have the same demographic estimates for the controlled variables.

Data collection and processing

Due to differences in format and design, the Interview Survey and the Diary Survey are collected and processed separately. The U.S. Census Bureau, under contract with BLS, carries out data collection for both. 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.

Quarterly Interview Survey. Beginning April 2003, Census Field Representatives (FRs) began collecting the Interview 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 FR performs some coding of expenses—by selecting from a predetermined list—for vehicle make and model, trip destination, and job types for alterations, maintenance, and repair.

Data are electronically transferred from the FR’s laptop at completion of the interview to the Census Master Control System. The Census Bureau Demographics Surveys Division then reformats the data into datasets and does special processing for output to BLS (such as converting missing values to special characters and merging data records into the required BLS output structure.) Some data, like vehicle and mortgage records, are copied into an input file that is loaded on the laptops for subsequent interviews the next quarter. This way, a few fields are updated each quarter, rather than recollecting the entire data record.

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 missing demographic variables, calculate weights, and adjust data to include sales tax and to exclude business expenses or reimbursed expenditures.

Monthly data files are then combined into quarterly databases, and a more extensive data review is carried out. This step includes a review of the following: counts and means by region, family relationship coding inconsistencies, and selected extreme values for expenditure and income categories. Other adjustments convert mortgage and vehicle 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. Cases of questionable data values or relationships are investigated, and errors are corrected prior to 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 affect all fields in the database, except assets. Missing or invalid attributes or expenditures are imputed. 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 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 the data adjustment routines to analyze the results.

The interview and diary surveys implemented multiple imputations of income data, beginning with the publication of the 2004 data. Prior to that, only income data collected from complete income reporters were published. However, even complete income reporters did not provide 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. A limitation was that bracketed data only provide 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 used in the published data.

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 this time, the interviewer also leaves the Diary questionnaire—or daily expenditure record—with the consumer unit, to record expenditures for the week.

Respondents record all expenses incurred during their participation in the survey in the diary questionnaire, a self-reporting, product-oriented diary. The diary is divided by day of purchase and by a broad classification of goods and services.

At the end of the first week, the interviewer collects the diary, reviews the entries, answers any questions, and leaves a second diary. The interviewer picks up the second diary at the end of the second week 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. Each week of a consumer unit's participation in the survey is treated as a separate occurrence.

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. All notes are reviewed, so expenditure data can be transcribed to the questionnaire for keying. In addition, item codes are assigned to reported expenditure items, house-

hold and consumer unit codes are assigned to each household member, and industry and occupation codes are entered for each working member. After an initial clerical screening, data are key-entered into electronic formats and a computer file of the database containing these data is produced and transmitted monthly to BLS, along with image files of questionnaires.

Data then are processed to calculate population weights based on BLS specifications, impute demographic characteristics for missing or inconsistent demographic data, impute values for weeks worked when nonresponse is encountered, and apply appropriate sales taxes to the expenditure items.

Using three monthly diary data files, BLS creates a quarterly database and screens it 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 package type, needed for mapping Diary expenditures. Income data from the Diary Survey are processed in the same way as in the Interview Survey.

Reliability of the data

Sample surveys are subject to two types of errors, sampling and nonsampling. Sampling error is the uncertainty in the Consumer Expenditure Survey estimates caused by the fact that data are collected from a sample of consumer units across the United States, instead of collecting data from every consumer unit. (The United States has approximately 120 million consumer units.) In 2007, usable data were collected in 27,335 quarterly interviews in the Interview Survey, and in 13,747 weekly diaries in the Diary Survey. Non-sampling error is the rest of the error. Non-sampling error includes things like incorrect information given by respondents, data processing errors, and so on. Non-sampling error occurs, regardless of whether data are collected from a sample of consumer units, or from the complete universe of consumer units.

The most common measure of the variability in a survey's estimates caused by sampling error is the standard error of the estimates—the square root of the variance. The standard error of the Consumer Expenditure Survey estimates can be used to construct confidence intervals to test various statistical hypotheses. Tables showing the standard errors of the Consumer Expenditure Survey estimates are available on the Internet at the Consumer Expenditure Survey webpage: www.bls.gov/cex.

The Bureau of Labor Statistics is constantly working to reduce error in the Consumer Expenditure Survey. Sampling error is reduced by using a sample of consumer units that is as large as possible given resource constraints. Non-sampling error is reduced through a series of computerized and professional data reviews, through continuous survey process improvements, as well as through theoretical research.