

---

## Appendix III

# Limitations of the Data

---

**Introduction**—The data presented in this *Statistical Abstract* came from many sources. The sources include not only federal statistical bureaus and other organizations that collect and issue statistics as their principal activity, but also governmental administrative and regulatory agencies, private research bodies, trade associations, insurance companies, health associations, and private organizations such as the National Education Association and philanthropic foundations. Consequently, the data vary considerably as to reference periods, definitions of terms and, for ongoing series, the number and frequency of time periods for which data are available.

The statistics presented were obtained and tabulated by various means. Some statistics are based on complete enumerations or censuses while others are based on samples. Some information is extracted from records kept for administrative or regulatory purposes (school enrollment, hospital records, securities registration, financial accounts, social security records, income tax returns, etc.), while other information is obtained explicitly for statistical purposes through interviews or by mail. The estimation procedures used vary from highly sophisticated scientific techniques, to crude “informed guesses.”

Each set of data relates to a group of individuals or units of interest referred to as the *target universe* or *target population*, or simply as the *universe* or *population*. Prior to data collection the target universe should be clearly defined. For example, if data are to be collected for the universe of households in the United States, it is necessary to define a “household.” The target universe may not be completely tractable. Cost and other considerations may restrict data collection to a *survey universe* based on some available list, such list may be of date. This list is called a *survey frame* or *sampling frame*.

The data in many tables are based on data obtained for all population units, a *census*, or on data obtained for only a portion, or *sample*, of the population units. When the data presented are based on a sample, the sample is usually a scientifically selected *probability sample*. This is a sample selected from a list or sampling frame in such a way that every possible sample has a known chance of selection and usually each unit selected can be assigned a number, greater than zero and less than or equal to one, representing its likelihood or probability of selection.

For large-scale sample surveys, the probability sample of units is often selected as a multistage sample. The first stage of a multistage sample is the selection of a probability sample of large groups of population members, referred to as primary sampling units (PSUs). For example, in a national multistage household sample, PSUs are often counties or groups of counties. The second stage of a multistage sample is the selection, within each PSU selected at the first stage, of smaller groups of population units, referred to as secondary sampling units. In subsequent stages of selection, smaller and smaller nested groups are chosen until the ultimate sample of population units is obtained. To qualify a multistage sample as a probability sample, all stages of sampling must be carried out using probability sampling methods.

Prior to selection at each stage of a multistage (or a single stage) sample, a list of the sampling units or sampling frame for that stage must be obtained. For example, for the first stage of selection of a national household sample, a list of the counties and county groups that form the PSUs must be obtained. For the final stage of selection, lists of households, and sometimes persons within the households, have to be compiled in the field. For surveys of economic entities and for

the economic censuses the Bureau generally uses a frame constructed from the Bureau's Business Register. The Business Register contains all establishments with payroll in the United States including small single establishment firms as well as large multi-establishment firms.

Wherever the quantities in a table refer to an entire universe, but are constructed from data collected in a sample survey, the table quantities are referred to as *sample estimates*. In constructing a sample estimate, an attempt is made to come as close as is feasible to the corresponding universe quantity that would be obtained from a complete census of the universe. Estimates based on a sample will, however, generally differ from the hypothetical census figures. Two classifications of errors are associated with estimates based on sample surveys: (1) *sampling error*—the error arising from the use of a sample, rather than a census, to estimate population quantities and (2) *non-sampling error*—those errors arising from nonsampling sources. As discussed below, the magnitude of the sampling error for an estimate can usually be estimated from the sample data. However, the magnitude of the nonsampling error for an estimate can rarely be estimated. Consequently, actual error in an estimate exceeds the error that can be estimated.

The particular sample used in a survey is only one of a large number of possible samples of the same size which could have been selected using the same sampling procedure. Estimates derived from the different samples would, in general, differ from each other. The *standard error* (SE) is a measure of the variation among the estimates derived from all possible samples. The standard error is the most commonly used measure of the sampling error of an estimate. Valid estimates of the standard errors of survey estimates can usually be calculated from the data collected in a probability sample. For convenience, the standard error is sometimes expressed as a percent of the estimate and is called the relative standard error or *coefficient of variation* (CV). For example, an estimate of 200 units with an estimated standard error of 10 units has an estimated CV of 5 percent.

A sample estimate and an estimate of its standard error or CV can be used to construct interval estimates that have a prescribed confidence that the interval includes the average of the estimates derived from all possible samples with a known probability. To illustrate, if all possible samples were selected under essentially the same general conditions, and using the same sample design, and if an estimate and its estimated standard error were calculated from each sample, then: (1) Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average estimate derived from all possible samples; (2) approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average estimate derived from all possible samples; and (3) approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average estimate derived from all possible samples.

Thus, for a particular sample, one can say with the appropriate level of confidence (e.g., 90 percent or 95 percent) that the average of all possible samples is included in the constructed interval. Example of a confidence interval: An estimate is 200 units with a standard error of 10 units. An approximately 90 percent confidence interval (plus or minus 1.6 standard errors) is from 184 to 216.

All surveys and censuses are subject to nonsampling errors. Nonsampling errors are of two kinds—*random* and *nonrandom*. Random nonsampling errors arise because of the varying interpretation of questions (by respondents or interviewers) and varying actions of coders, keyers, and other processors. Some randomness is also introduced when respondents must estimate. Nonrandom nonsampling errors result from total nonresponse (no usable data obtained for a sampled unit), partial or item nonresponse (only a portion of a response may be usable), inability or unwillingness on the part of respondents to provide correct information, difficulty interpreting questions, mistakes

in recording or keying data, errors of collection or processing, and coverage problems (overcoverage and undercoverage of the target universe). Random nonresponse errors usually, but not always, result in an understatement of sampling errors and thus an overstatement of the precision of survey estimates. Estimating the magnitude of nonsampling errors would require special experiments or access to independent data and, consequently, the magnitudes are seldom available.

Nearly all types of nonsampling errors that affect surveys also occur in complete censuses. Since surveys can be conducted on a smaller scale than censuses, nonsampling errors can presumably be controlled more tightly. Relatively more funds and effort can perhaps be expended toward eliciting responses, detecting and correcting response error, and reducing processing errors. As a result, survey results can sometimes be more accurate than census results.

To compensate for suspected nonrandom errors, adjustments of the sample estimates are often made. For example, adjustments are frequently made for nonresponse, both total and partial. Adjustments made for either type of nonresponse are often referred to as *imputations*. Imputation for total nonresponse is usually made by substituting for the questionnaire responses of the nonrespondents the “average” questionnaire responses of the respondents. These imputations usually are made separately within various groups of sample members, formed by attempting to place respondents and nonrespondents together that have “similar” design or ancillary characteristics. Imputation for item nonresponse is usually made by substituting for a missing item the response to that item of a respondent having characteristics that are “similar” to those of the nonrespondent.

For an estimate calculated from a sample survey, the *total error* in the estimate is composed of the sampling error, which can usually be estimated from the sample, and the nonsampling error, which

usually cannot be estimated from the sample. The total error present in a population quantity obtained from a complete census is composed of only nonsampling errors. Ideally, estimates of the total error associated with data given in the *Statistical Abstract* tables should be given. However, due to the unavailability of estimates of nonsampling errors, only estimates of the levels of sampling errors, in terms of estimated standard errors or coefficients of variation, are available. To obtain estimates of the estimated standard errors from the sample of interest, obtain a copy of the referenced report which appears at the end of each table.

**Principal data bases.**—Beginning below are brief descriptions of 36 of the sample surveys and censuses that provide a substantial portion of the data contained in this *Abstract*.

#### SECTION 1. POPULATION

**Source and Title:** U.S. Census Bureau, *Census of Population*

**Tables:** See Tables citing *Census of Population* in Section 1 and also in Sections 4, 6, 7, 20, and 29.

**Universe, Frequency, and Types of Data:** Complete count of U.S. population conducted every 10 years since 1790. Data obtained on number and characteristics of people in the U.S.

**Type of Data Collection Operation:** In 1970, 1980, 1990 and 2000 complete census for some items—age, sex, race, and relationship to householder. In 1970, other items collected from a 5% and a 15% probability (systematic) sample of the population. In 1980, approximately 19% of the housing units were included in the sample; in 1990 and 2000, approximately 17%.

**Data Collection and Imputation Procedures:** In 1970, extensive use of mail questionnaires in urban areas; personal interviews in most rural areas. In 1980, 1990, and 2000, mail questionnaires were used in even more areas than in 1970, with personal interviews in the remainder. Extensive telephone and personal followup for nonrespondents was done in the censuses. Imputations were made for missing characteristics.

**Estimates of Sampling Error:** Sampling errors for data are estimated for all items collected by sample and vary by characteristic and geographic area. The CVs for national and state estimates are generally very small.

**Other (Nonsampling) Errors:** Since 1950, evaluation programs have been conducted to provide information on the magnitude of some sources of nonsampling errors such as response bias and undercoverage in each census. Results from the evaluation program for the 1990 census indicate that the net under coverage amounted to about 1.5 to 2 percent of the total resident population.

**Sources of Additional Material:** U.S. Census Bureau, *The Coverage of Population in the 1980 Census*, PHC80-E4; *Content Reinterview Study: Accuracy of Data for Selected Population and Housing Characteristics as Measured by Reinterview*, PHC80-E2; *1980 Census of Population*, Vol. 1., (PC80-1), Appendixes B, C, and D.

**Source and Title:** U.S. Census Bureau, *Current Population Survey (CPS)*

**Tables:** See Tables citing *Current Population Reports* primarily in Section 1, but also in Sections 2, 3, 4, 7, 11, 12, 13, and 29. Many Bureau of Labor Statistics' (BLS) Tables in Section 13 are CPS based.

**Universe, Frequency, and Types of**

**Data:** Nationwide monthly sample survey of civilian noninstitutional population, 15 years old or over, to obtain data on employment, unemployment, and a number of other characteristics.

**Type of Data Collection Operation:** Multistage probability sample of about 50,000 households in 754 PSUs in 1996. Over-sampling in some states and the largest MSAs to improve reliability for those areas of employment data on annual average basis. A continual sample rotation system is used. Households are in sample 4 months, out for 8 months, and in for 4 more. Month-to-month overlap is 75%; year-to-year overlap is 50%.

**Data Collection and Imputation Procedures:** For first and fifth months that a household is in sample, personal interviews; other months, approximately, 85% of the data collected by phone.

Imputation is done for both item and total nonresponse. Adjustment for total nonresponse is done by a predefined cluster of units, by MSA size and residence; for item nonresponse imputation varies by subject matter.

**Estimates of Sampling Error:** Estimated CVs on national annual averages for labor force, total employment, and nonagricultural employment, 0.2%; for total unemployment and agricultural employment, 1.0% to 2.5%. The estimated CVs for family income and poverty rate for all persons in 1986 are 0.5% and 1.5%, respectively. CVs for subnational areas, such as states, would be larger and would vary by area.

**Other (Nonsampling) Errors:** Estimates of response bias on unemployment are not available, but estimates of unemployment are usually 5% to 9% lower than estimates from reinterviews. Six to 7.0% of sample households unavailable for interviews.

**Sources of Additional Material:** U.S. Census Bureau and Bureau of Labor Statistics, *Current Population Survey; Design and Methodology*, (Tech. Paper 63), available on internet <[www.bls.census.gov/cps/tp/tp63.htm](http://www.bls.census.gov/cps/tp/tp63.htm)> and Bureau of Labor Statistics, *Employment and Earnings*, monthly, Explanatory Notes and Estimates of Error, Tables 1-A through 1-D and *BLS Handbook of Methods*, Chapter 1 (Bulletin 2490.)

## SECTION 2. VITAL STATISTICS

**Source and Title:** U.S. National Center for Health Statistics (NCHS), *National Vital Statistics System*

**Tables:** See Tables citing *Vital Statistics of the United States*; 297 in Section 5; and 1308 in Section 29.

**Universe, Frequency, and Types of Data:** Annual data on births and deaths in the United States.

**Type of Data Collection Operation:** Mortality data based on complete file of death records, except 1972, based on 50% sample. Natality statistics 1951-71, based on 50% sample of birth certificates, except a 20% to 50% in 1967, received by NCHS. Beginning 1972, data from some states received through Vital Statistics Cooperative Program (VSCP)

and complete file used; data from other states based on 50% sample. Beginning 1986, all reporting areas participated in the VSCP.

#### **Data Collection and Imputation**

**Procedures:** Reports based on records from registration offices of all states, District of Columbia, New York City, Puerto Rico, Virgin Islands, and Guam.

**Estimates of Sampling Error:** For recent years, CVs for births are small due to large portion of total file in sample (except for very small estimated totals).

**Other (Nonsampling) Errors:** Data on births and deaths believed to be at least 99% complete.

**Sources of Additional Material:** U.S. National Center for Health Statistics, *Vital Statistics of the United States*, Vol. I and Vol. II, annual, and *National Vital Statistics Report*.

(See Section 1 above for information pertaining to Tables 82-84.)

### SECTION 3. HEALTH AND NUTRITION

**Source and Title:** U.S. National Center for Health Statistics, *National Health Interview Survey (NHIS)*

**Tables:** 156, 178, 182, 191, 193, and 197.

#### **Universe, Frequency, and Types of**

**Data:** Continuous data collection covering the civilian noninstitutional population to obtain information on personal and demographic characteristics, illnesses, injuries, impairments, and other health topics.

#### **Type of Data Collection Operation:**

Multistage probability sample of 49,000 households (in 198 PSUs) from 1985 to 1994; 43,000 households (358 design PSUs) from 1995 on, selected in groups of about four adjacent households.

#### **Data Collection and Imputation Procedures:**

Some missing data items (e.g., age) are imputed using an average value. Unit nonresponse is compensated for by an adjustment to the survey weights

#### **Estimates of Sampling Error:**

Estimated CVs: For physician visits by males, 1.5%; for workdays lost by males, 3.5%; for persons injured at home, 4.7%.

**Other (Nonsampling) Errors:** Response rate was 93.8% in 1996 for the NHIS.

**Sources of Additional Material:** U.S. National Center for Health Statistics, "Current Estimates from the National Health Interview Survey, U.S., 1996," Vital and Health Statistics, Series 10 #200.

(See Section 1 for information regarding the Current Population Survey, Section 12 for information pertaining to Table 151, Section 14 for Table 129, Section 22 for Table 148, 149, and 150.)

### SECTION 4. EDUCATION

**Source and Title:** U.S. Department of Education, National Center for Education Statistics, *Higher Education General Information Survey (HEGIS), Fall Enrollment in Institutions of Higher Education*; beginning 1986, *Integrated Postsecondary Education Data Survey (IPEDS), Fall Enrollment*

**Tables:** 205, 207, 263, 264, 265, and 266.

#### **Universe, Frequency, and Types of**

**Data:** Annual survey of all institutions and branches listed in the *Directory* to obtain data on total enrollment by sex, level of enrollment, type of program, racial/ethnic characteristics (every other year prior to 1989, then annually) and attendance status of student, and on first-time students.

#### **Type of Data Collection Operation:**

Complete census.

#### **Data Collection and Imputation**

**Procedures:** Survey package is usually mailed in the spring with surveys due at varying dates in the summer and fall; mail and phone followup procedures for nonrespondents. Missing data are imputed by using data of similar institutions.

**Estimates of Sampling Error:** Not applicable.

**Other (Nonsampling) Errors:** For degree-granting institutions approximately 92% response rate in fall 1998.

**Sources of Additional Material:** U.S. Department of Education, National Center for Education Statistics, *Fall Enrollment in Higher Education*, annual.

**Source and Title:** U.S. Department of Education, National Center for Education

Statistics, *Higher Education General Information Survey (HEGIS)*, *Financial Statistics of Institutions of Higher Education*; beginning 1986, *Integrated Postsecondary Education Data Survey (IPEDS)*, *Finance*.

**Tables:** 206 and 264.

**Universe, Frequency, and Types of**

**Data:** Annual survey of all institutions and branches listed in the *Education Directory, Colleges, and Universities* to obtain data on financial status and operations, including current funds revenues, current funds expenditures, and physical plant assets.

**Type of Data Collection Operation:**

Complete census.

**Data Collection and Imputation**

**Procedures:** Survey package is usually mailed in the spring with surveys due at varying dates in the summer and fall; mail and phone followup procedures for nonrespondents. Missing data are imputed by using data of similar institutions.

**Estimates of Sampling Error:** Not applicable.

**Other (Nonsampling) Errors:** For 1997, 95% for degree-granting institutions.

**Sources of Additional Material:** U.S. Department of Education, National Center for Education Statistics, *Financial Statistics*.

**Source and Title:** U.S. Department of Education, National Center for Education Statistics, *Higher Education General Information Survey (HEGIS)*, *Degrees and Other Formal Awards Conferred*. Beginning 1986, *Integrated Postsecondary Education Data Survey (IPEDS)*, *Completions*.

**Tables:** 207, 284, 285, 286, 287, 288, and 289.

**Universe, Frequency, and Types of**

**Data:** Annual survey of all institutions and branches listed in the *Education Directory, Colleges, and Universities* to obtain data on earned degrees and other formal awards, conferred by field of study, level of degree, sex, and by racial/ethnic characteristics (every other year prior to 1989, then annually).

**Type of Data Collection Operation:**

Complete census.

**Data Collection and Imputation Procedures:** Survey package is usually mailed in the spring with surveys due at varying dates in the summer and fall; mail and phone followup procedures for nonrespondents. Missing data are imputed by using data of similar institutions.

**Estimates of Sampling Error:** Not applicable.

**Other (Nonsampling) Errors:** For 1996-98, approximately 92% response rate for degree-granting institutions.

**Sources of Additional Material:** U.S. Department of Education, National Center for Education Statistics, *Completions*, annual.

(See Sections 1 and 9 for information pertaining to the Census Bureau and Section 5 for information pertaining to Table 220.)

SECTION 5. LAW ENFORCEMENT, COURTS, AND PRISONS

**Source and Title:** U.S. Federal Bureau of Investigation, *Uniform Crime Reporting (UCR) Program*

**Tables:** 291, 292, 293, 294, 295, 296, 298, 299, 300, 301, 308, 309, and 311.

**Universe, Frequency, and Types of**

**Data:** Monthly reports on the number of criminal offenses that become known to law enforcement agencies. Data are collected on crimes cleared by arrest, by age, sex, and race of offender, and on assaults on law enforcement officers.

**Type of Data Collection Operation:**

Crime statistics are based on reports of crime data submitted either directly to the FBI by contributing law enforcement agencies or through cooperating state UCR programs.

**Data Collection and Imputation Procedures:**

States with UCR programs collect data directly from individual law enforcement agencies and forward reports, prepared in accordance with UCR standards, to FBI. Accuracy and consistency edits are performed by FBI.

**Estimates of Sampling Error:** Not applicable.

**Other (Nonsampling) Errors:** Coverage of 95% of the population (96% in MSAs, 92% in "other cities," and 90% in rural areas) by UCR program, though varying number of agencies report. Some error may be present through incorrect reporting.

**Sources of Additional Material:** U.S. Federal Bureau of Investigation, *Crime in the United States, Hate Crime Statistics, annual, Law Enforcement Officers Killed & Assaulted*, annual.

**Source and Title:** U.S. Bureau of Justice Statistics (BJS), *National Crime Victimization Survey*

**Tables:** 302, 303, 304, 305, 306, 307, and Table 220 in Section 4.

**Universe, Frequency, and Types of Data:** Monthly survey of individuals and households in the United States to obtain data on criminal victimization of those units for compilation of annual estimates.

**Type of Data Collection Operation:** National probability sample survey of about 50,000 interviewed households in 376 PSUs selected from a list of addresses from the 1980 census, supplemented by new construction permits and an area sample where permits are not required.

#### **Data Collection and Imputation**

**Procedures:** Interviews are conducted every 6 months for 3 years for each household in the sample; 8,300 households are interviewed monthly. Personal interviews are used in the first interview; the intervening interviews are conducted by telephone whenever possible.

**Estimates of Sampling Error:** CVs averaged over the period 1996-1999 are: 2.7% for crimes of violence; 8.5% for estimate of rape/sexual assault counts; 5.5% for robbery counts; 2.8% for assault counts; 8.0% for personal theft counts; 1.9% for property crimes; 3.0% for burglary counts; 2.0% for theft; and 4.4% for motor vehicle theft counts.

**Other (Nonsampling) Errors:** Respondent recall errors which may include reporting incidents for other than the reference period; interviewer coding and processing errors; and possible mistaken reporting or classifying of events. Adjustment is made for a household noninterview rate of about 4% and for a within-household noninterview rate of 7%.

**Sources of Additional Material:** U.S. Bureau of Justice Statistics, *Criminal Victimization in the United States*, annual. (See Section 2 for details on Table 297.)

## SECTION 7. ELECTIONS

(See Section 1 above for information pertaining to Tables 401 and 402.)

## SECTION 8. STATE AND LOCAL GOVERNMENT FINANCES AND EMPLOYMENT

**Source and Title:** U.S. Census Bureau, *Census of Governments*

**Tables:** See Tables in Section 8 citing *Census of Governments*.

**Universe, Frequency, and Types of Data:** Survey of all governmental units in the United States conducted every 5 years to obtain data on government revenue, expenditures, debt, assets, employment and employee retirement systems, property values, public school systems, and number, size, and structure of governments.

**Type of Data Collection Operation:** Complete census. List of units derived through classification of government units recently authorized in each state and identification, counting, and classification of existing local governments and public school systems.

**Data Collection and Imputation Procedures:** Data collected through field and office compilation of financial data from official records and reports for states and large local governments; mail canvass of selected data items, like state tax revenue and employee retirement systems; and collection of local government statistics through central collection arrangements with state governments.

**Estimates of Sampling Error:** Not applicable.

**Other (Nonsampling) Errors:** Some nonsampling errors may arise due to possible inaccuracies in classification, response, and processing.

**Sources of Additional Material:** U.S. Census Bureau, *Census of Governments, 1987*, various reports, and *State Government Finances in 1990*, GF 90, No. 3.

**Source and Title:** U.S. Census Bureau, *Annual Surveys of State and Local Government*

**Tables:** See Tables citing *Public Employment* and *Government Finances* in Section 8 and Table 241 in Section 4; and Table 531 in Section 11.

## **Universe, Frequency, and Types of**

**Data:** Sample survey conducted annually to obtain data on revenue, expenditure, debt, and employment of state and local governments. Universe is all governmental units in the United States (about 87,000).

## **Type of Data Collection Operation:**

Sample of about 22,000 units includes all state governments, county governments with 50,000+ population, municipalities and townships with 25,000+ population, all school districts with 10,000+ enrollment in October 1986, and other governments meeting certain criteria; probability sample for remaining units.

## **Data Collection and Imputation Procedures:**

Field and office compilation of data from official records and reports for states and large local governments; central collection of local governmental financial data through cooperative agreements with a number of state governments; mail canvass of other units with mail and telephone followups of nonrespondents. Data for nonresponses are imputed from previous year data or obtained from secondary sources, if available.

**Estimates of Sampling Error:** CVs for estimates of major employment and financial items are generally less than 2% for most states and less than 1.2% for the majority of states.

**Other (Nonsampling) Errors:** Nonresponse rate is less than 15% for number of units. Other possible errors may result from undetected inaccuracies in classification, response, and processing.

**Sources of Additional Material:** U.S. Census Bureau, *Public Employment in 1990*, GE 90, No. 1, *Governmental Finances in 1989-90*, GF 90, No. 5, and *Census of Governments, 1987*, various reports.

## SECTION 9. FEDERAL GOVERNMENT

**Source and Title:** U.S. Internal Revenue Service, *Statistics of Income, Individual Income Tax Returns*

**Tables:** 473, 474, 475, 476, and 477.

## **Universe, Frequency, and Types of**

**Data:** Annual study of unaudited individual income tax returns, forms 1040, 1040A, and 1040EZ, filed by U.S. citizens and residents. Data provided on various financial characteristics by size

of adjusted gross income, marital status, and by taxable and nontaxable returns. Data by state, based on 100% file, also include returns from 1040NR, filed by nonresident aliens plus certain self-employment tax returns.

## **Type of Data Collection Operation:**

Annual stratified probability sample of approximately 125,000 returns broken into sample strata based on the larger of total income or total loss amounts as well as the size of business plus farm receipts. Sampling rates for sample strata varied from .025% to 100%.

## **Data Collection and Imputation Procedures:**

Computer selection of sample of tax return records. Data adjusted during editing for incorrect, missing, or inconsistent entries to ensure consistency with other entries on return.

**Estimates of Sampling Error:** Estimated CVs for tax year 1997: Adjusted gross income less deficit .15%; salaries and wages .27%; and tax-exempt interest received 2.15%. (State data not subject to sampling error.)

**Other (Nonsampling) Errors:** Processing errors and errors arising from the use of tolerance checks for the data.

**Sources of Additional Material:** U.S. Internal Revenue Service, *Statistics of Income, Individual Income Tax Returns*, annual.

## SECTION 11. SOCIAL INSURANCE AND HUMAN SERVICES

**Source and Title:** U.S. Social Security Administration, *Benefit Data*

**Tables:** 527 and 528.

## **Universe, Frequency, and Types of**

**Data:** All persons receiving monthly benefits under Title II of Social Security Act. Data on number and amount of benefits paid by type and state.

## **Type of Data Collection Operation:**

Data based on administrative records. Data based on 100% files, as well as 10% and 1% sample files.

## **Data Collection and Imputation Procedures:**

Records used consist of actions pursuant to applications dated by subsequent post-entitlement actions.

**Estimates of Sampling Error:** Varies by size of estimate and sample file size.

**Other (Nonsampling) Errors:** Processing errors, which are believed to be small.

**Sources of Additional Material:** U.S. Social Security Administration, *Annual Statistical Supplement to the Social Security Bulletin*.

**Source and Title:** U.S. Social Security Administration, *Supplemental Security Income (SSI) Program*

**Tables:** 542 and 543.

**Universe, Frequency, and Types of**

**Data:** All eligible aged, blind, or disabled persons receiving SSI benefit payments under SSI program. Data include number of persons receiving federally administered SSI, amounts paid, and state administered supplementation.

**Type of Data Collection Operation:** Data based on administrative records.

**Data Collection and Imputation Procedures:** Data adjusted to reflect returned checks and overpayment refunds. For federally administered payments, actual adjusted amounts are used.

**Estimates of Sampling Error:** Not applicable.

**Other (Nonsampling) Errors:** Processing errors, which are believed to be small.

**Sources of Additional Material:** U.S. Social Security Administration, *Annual Statistical Supplement to the Social Security Bulletin*.

(See Section 1 above for information pertaining to the Current Population Survey, Section 8 for information pertaining to Annual Surveys of State and Local Government, and Section 22 for the 1997 Economic Census and the Service Annual Survey.)

SECTION 12. LABOR FORCE, EMPLOYMENT, AND EARNINGS

**Source and Title:** U.S. Bureau of Labor Statistics (BLS), *Current Employment Statistics (CES) Program*

**Tables:** 607, 608, 609, 616, in Section 3, Table 151, and in Section 23, Table 1055.

**Universe, Frequency, and Types of**

**Data:** Monthly survey covering about 7 million establishments to obtain data on employment, hours, and earnings, by industry.

**Type of Data Collection Operation:** Sample survey of about 350,000 establishments in June 2001.

**Data Collection and Imputation Procedures:** Cooperating state agencies mail questionnaires to sample establishments to develop state and local estimates; information is forwarded to BLS where national estimates are prepared.

**Estimates of Sampling Error:** Not available until survey redesign is completed.

**Other (Nonsampling) Errors:** Estimates of employment adjusted annually to reflect complete universe. Average adjustment is 0.3% over the last decade.

**Sources of Additional Material:** U.S. Bureau of Labor Statistics, *Employment and Earnings*, monthly, Explanatory Notes and Estimates of Error, Tables 2-A through 2-H.

(See Section 1 above for information pertaining to the Current Population Survey.)

SECTION 13. INCOME, EXPENDITURES, AND WEALTH

**Source and Title:** Board of Governors of the Federal Reserve System, *Survey of Consumer Finances*

**Tables:** 687 and 688, and in Section 25, 1167, 1169, 1170, 1171, 1192, and 1209.

**Universe, Frequency, and Types of**

**Data:** Periodic sample survey of families. In this survey a given household is divided into a primary economic unit and other economic units. The primary economic unit, which may be a single individual, is generally chosen as the unit that contains the person who either holds the title to the home or is the first person listed on the lease. The primary unit is used as the reference family. The survey collects detailed data on the composition of family balance sheets, the terms of loans, and relationships with financial institutions. It also gathered information on the employment history and pension rights of the survey respondent and the spouse or partner of the respondent.

**Type of Data Collection Operation:** The survey employs a two-part strategy for sampling families. Some families were selected by standard multistage area-probability sampling methods from the 48 contiguous states. The remaining families in the survey were selected

using tax data under the strict rules governing confidentiality and the rights of potential respondents to refuse participation.

**Data Collection and Imputation Procedures:** The Survey Research Center at the University of Michigan collected the 1989 survey data between August 1989 and March 1990. Adjustments for nonresponse errors are made through systematic imputation of unanswered questions and through weighting adjustments based on data used in the sample design for families that refused participation.

**Estimates of Sampling Error:** Because of the complex design of the survey, the estimation of potential sampling errors is not straightforward.

**Other (Nonsampling) Errors:** The achieved sample of 3,143 families represents a response rate of about 69 percent in the area-probability sample and a rate of about 34 percent in the tax-data sample. Proper training of interviewers and careful design of questionnaires were used to control inaccurate survey responses.

**Sources of Additional Material:** Board of Governors of the Federal Reserve System, "Changes in Family Finances from 1983 to 1989: Evidence from the Survey of Consumer Finances", *Federal Reserve Bulletin*, January 1992.

(See Section 1 above for information pertaining to the Census Bureau.)

## SECTION 14. PRICES

**Source and Title:** U.S. Bureau of Labor Statistics (BLS), *Consumer Price Index (CPI)*

**Tables:** 691, 692, 693, 694, 706, and in Section 3, Table 129.

**Universe, Frequency, and Types of Data:** Monthly survey of price changes of all types of consumer goods and services purchased by urban wage earners and clerical workers prior to 1978, and urban consumers thereafter. Both indexes continue to be published.

**Type of Data Collection Operation:** Prior to 1978, sample of various consumer items in 87 urban areas; thereafter, in 85 PSUs, except from January 1987 through March 1988, when 91 areas were sampled.

**Data Collection and Imputation Procedures:** Prices of consumer items are obtained from about 50,000 housing units, and 23,000 other reporters in 87 areas. Prices of food, fuel, and a few other items are obtained monthly; prices of most other commodities and services are collected every month in the three largest geographic areas and every other month in others.

**Estimates of Sampling Error:** Estimates of standard errors are available.

**Other (Nonsampling) Errors:** Errors result from inaccurate reporting, difficulties in defining concepts and their operational implementation, and introduction of product quality changes and new products.

**Sources of Additional Material:** U.S. Bureau of Labor Statistics, Internet site and *BLS Handbook of Methods*, Chapter 17, Bulletin 2490.

**Source and Title:** U.S. Bureau of Labor Statistics, *Producer Price Index (PPI)*

**Tables:** 691, 697, 698, 699, and in Section 18, Table 855; and in Section 20, Table 927; and in Section 23, Table 1115.

**Universe, Frequency, and Types of Data:** Monthly survey of producing companies to determine price changes of all commodities produced in the United States for sale in commercial transactions. Data on agriculture, forestry, fishing, manufacturing, mining, gas, electricity, public utilities, and a few services.

**Type of Data Collection Operation:** Probability sample of approximately 30,000 about 100,000 price quotations per month.

**Data Collection and Imputation Procedures:** Data are collected by mail. If transaction prices are not supplied, list prices are used. Some prices are obtained from trade publications, organized exchanges, and government agencies. To calculate index, price changes are multiplied by their relative weights taken from 1992 shipment values from the Census of Manufactures.

**Estimates of Sampling Error:** Not applicable.

**Other (Nonsampling) Errors:** Not available at present.

**Sources of Additional Material:** U.S. Bureau of Labor Statistics, *BLS Handbook of Methods*, Chapter 14, Bulletin 2490.

## SECTION 15. BUSINESS ENTERPRISE

**Source and Title:** U.S. Census Bureau, *County Business Patterns*

**Tables:** 723 and 724, and in Section 18, Table 847; in Section 19, Table 890; in Section 20, Table 925; in Section 21, Tables 972 and 973; in Section 22, Table 1017 and 1049; in Section 23, Table 1053; in Section 24, Table 1121; in Section 25, Table 1162 and 1228; in Section 26, Table 1232; and in Section 27, Tables 1263, 1264, and 1280.

### **Universe, Frequency, and Types of**

**Data:** Annual tabulation of basic data items extracted from the Business Register, a file of all known single and multi-establishment companies maintained and updated by the Census Bureau. Data include number of establishments, number of employees, first-quarter and annual payrolls, and number of establishments by employment size class. Data are excluded for self-employed persons, domestic service workers, railroad employees, agricultural production workers, and most government employees.

### **Type of Data Collection Operation:**

The annual Company Organization Survey provides individual establishment data for multiestablishment companies. Data for single establishment companies are obtained from various Census Bureau programs, such as the Annual Survey of Manufactures and Current Business Surveys, as well as from administrative records of the Internal Revenue Service and the Social Security Administration.

**Estimates of Sampling Error:** Not applicable.

**Other (Nonsampling) Error:** Response rates of greater than 90% for the 1993 Company Organization Survey.

**Sources of Additional Materials:** U.S. Census Bureau, *General Explanation of County Business Patterns*.

**Source and Title:** U.S. Census Bureau, *Surveys of Minority- and Women-Owned Business Enterprises* (SMOBE/SWOBE).

**Tables:** 728, 729, 730, 731, 732, 733, 734, 735, and 736.

### **Universe, Frequency, and Types of**

**Data:** The surveys provide basic economic data on businesses owned by Blacks, Hispanics, Asians, Pacific Islanders, Alaska Natives, American Indians, and Women. All firms operating during 1997, except those classified as agricultural, are represented. The lists of all firms (or sample frames) are compiled from a combination of business tax returns and data collected on other economic census reports. The published data include the number of firms, gross receipts, number of paid employees, and annual payroll. The data are presented by geographic area, industry, size of firm, and legal form of organization of firm.

### **Type of Data Collection Operation:**

The surveys are based on a stratified probability sample of approximately 2.5 million firms from a universe of approximately 20.8 million firms. There were approximately 5.3 million firms with paid employees and 15.5 million firms with no paid employees. The data are based on the entire firm rather than on individual locations of a firm.

### **Data Collection and Imputation Procedures:**

Data were collected through a mailout/mailback operation. Compensation for missing data is addressed through reweighting, edit correction, and standard statistical imputation methods.

**Estimates of Sampling Error:** Variability in the estimates is due to the sample selection and estimation for items collected by SMOBE/SWOBE. CVs are applicable to only published cells in which sample cases are tabulated. The CVs for number of firms and receipts at the national level range from 1 to 4 percent.

**Other (Nonsampling) Error:** Nonsampling errors are attributed to many sources: inability to obtain information for all cases in the universe, adjustments to the weights of respondents to compensate for nonrespondents, imputation for missing data, data errors and biases, mistakes in recording or keying data, errors in collection or processing, and coverage problems.

**Sources of Additional Materials:** U.S. Census Bureau, *Guide to the 1997 Economic Census and Related Statistics*.

**Source and Title:** U.S. Internal Revenue Service, *Statistics of Income, Sole Proprietorship Returns* and *Statistics of Income Bulletin*

**Tables:** 710, 711, 721, and 713.

**Universe, Frequency, and Types of**

**Data:** Annual study of unaudited income tax returns of nonfarm sole proprietorships, form 1040 with business schedules. Data provided on various financial characteristics by industry.

**Type of Data Collection Operation:**

Stratified probability sample of approximately 31,000 sole proprietorships for tax year 1990. The sample is classified based on presence or absence of certain business schedules; the larger of total income or loss; and size of business plus farm receipts. Sampling rates vary from .043% to 100%.

**Data Collection and Imputation Proce-**

**dures:** Computer selection of sample of tax return records. Data adjusted during editing for incorrect, missing, or inconsistent entries to ensure consistency with other entries on return.

**Estimates of Sampling Error:** Esti-

mated CVs for tax year 1990 are not available; for 1987 (the latest available): For sole proprietorships, business receipts, 1.66%; net income, (less loss), 1.33%; depreciation 2.17%; interest expense 2.80%; and employee benefit programs 7.55%.

**Other (Nonsampling) Errors:** Process-

ing errors and errors arising from the use of tolerance checks for the data.

**Sources of Additional Material:** U.S.

Internal Revenue Service, *Statistics of Income, Sole Proprietorship Returns* (for years through 1980) and *Statistics of Income Bulletin*, Vol. 19, No. 1 (summer 1999).

**Source and Title:** U.S. Internal Revenue Service, *Statistics of Income, Partnership Returns* and *Statistics of Income Bulletin*

**Tables:** 710, 712, 714, and 715.

**Universe, Frequency, and Types of**

**Data:** Annual study of unaudited income tax returns of partnerships, Form 1065. Data provided on various financial characteristics by industry.

**Type of Data Collection Operation:**

Stratified probability sample of approximately 28,000 partnership returns from a population of 1,660,000 filed during

calendar year 1990. The sample is classified based on combinations of gross receipts, net income or loss, and total assets, and on industry. Sampling rates vary from .04% to 100%.

**Data Collection and Imputation Proce-**

**dures:** Computer selection of sample of tax return records. Data are adjusted during editing for incorrect, missing, or inconsistent entries to ensure consistency with other entries on return. Data not available due to regulations are not imputed.

**Estimates of Sampling Error:** Esti-

mated CVs for tax year 1988 (latest available): For number of partnerships, .51%; business receipts, .78%; net income, 3.03%; net loss, 2.21% and total assets, 1.22%.

**Other (Nonsampling) Errors:** Process-

ing errors and errors arising from the use of tolerance checks for the data.

**Sources of Additional Material:** U.S.

Internal Revenue Service, *Statistics of Income, Partnership Returns* and *Statistics of Income Bulletin*, Vol. 19, No. 2 (fall 1999).

**Source and Title:** U.S. Internal Revenue Service, *Corporation Income Tax Returns*

**Tables:** 710, 711, 712, 718, 719, and 720.

**Universe, Frequency, and Types of**

**Data:** Annual study of unaudited corporation income tax returns, Forms 1120 and 1120 (A, F, L, PC, REIT, RIC, and S), filed by corporations or businesses legally defined as corporations. Data provided on various financial characteristics by industry and size of total assets, and business receipts.

**Type of Data Collection Operation:**

Stratified probability sample of approximately 85,000 returns for 1987, distributed by sample classes generally based on type of return, size of total assets, size of net income or deficit, and selected business activity. Sampling rates for sample strata varied from .25% to 100%.

**Data Collection and Imputation Proce-**

**dures:** Computer selection of sample of tax return records. Data adjusted during editing for incorrect, missing, or inconsistent entries to ensure consistency with other entries on return and to achieve statistical definitions.

**Estimates of Sampling Error:** Estimated CVs for 1988: Number of returns in subgroups ranged from 1.4% with assets under \$100,000, to 0% with assets over \$100 mil.; for amount of net income and amount of income tax, .18%

**Other (Nonsampling) Errors:** Processing errors and errors arising from the use of tolerance checks for the data.

**Sources of Additional Material:** U.S. Internal Revenue Service, *Statistics of Income, Corporation Income Tax Returns*, annual.

(See Section 15 for Table 721)

## SECTION 17. AGRICULTURE

**Source and Title:** U.S. Department of Agriculture, National Agriculture Statistics Service (NASS), *Census of Agriculture*.

**Tables:** 796, 797, 798, 799, 800, 801, and 802.

### **Universe, Frequency, and Types of**

**Data:** Complete count of U.S. farms and ranches conducted once every 5 years with data at the national, state, and county level. Data published on farm numbers and related items/ characteristics.

### **Type of Data Collection Operation:**

Complete census for—number of farms; land in farms; estimated market value of land and buildings, machinery and equipment, agriculture products sold; total cropland; irrigated land; total farm production expenses; farm operator characteristics; livestock and poultry inventory and sales; and selected crops harvested.

### **Data Collection and Imputation Proce-**

**dures:** Data collection is by mailing questionnaires to all farmers and ranchers. Nonrespondents are conducted by telephone and correspondence followups. Imputations were made for all nonresponse item/characteristics.

**Estimates of Sampling Error:** Variability in the estimates is due to the sample selection and estimation for items collected by sample and census nonresponse estimation procedures. The CVs for national and state estimates are generally very small. Approximately 85% response rate.

**Other (Nonsampling) Errors:** Nonsampling errors are due to incompleteness of the census mailing list, duplications

on the list, respondent reporting errors, errors in editing reported data, and in imputation for missing data. Evaluation studies are conducted to measure certain nonsampling errors such as list coverage and classification error. Results from the evaluation program for the 1987 census indicate the net under coverage amounted to about 7.2% of the nations total farms.

**Sources of Additional Material:** U.S. Department of Agriculture (NASS), *1997 Census of Agriculture*, Volume 2, Subject Series—Part 1, *Agriculture Atlas of the U.S.*, Part 2, *Coverage Evaluation*, Part 3, *Rankings of States and Counties*, Part 4, *History*, Part 5, *ZIP Code Tabulation of Selected Items*, ; and Volume 4 *Farm and Ranch Irrigation Survey*.

**Source and Title:** U.S. Department of Agriculture, National Agricultural Statistics Service (NASS), *Basic Area Frame Sample*

**Tables:** See Tables citing NASS in source notes in Section 23, which pertain to this or the following two surveys.

### **Universe, Frequency, and Types of**

**Data:** Two annual area sample surveys of U.S. farm operators: June agricultural survey collects data on planted acreage and livestock inventories; and a February Farm Costs and Returns survey that collects data on total farm production, expenses and specific commodity costs of production.

### **Type of Data Collection Operation:**

Stratified probability sample of about 16,000 land area units of about 1 sq. mile (range from 0.1 sq. mile in cities to several sq. miles in open grazing areas). Sample includes 60,000 parcels of agricultural land. About 20% of the sample replaced annually.

### **Data Collection and Imputation Proce-**

**dures:** Data collection is by personal enumeration. Imputation is based on enumerator observation or data reported by respondents having similar agricultural characteristics.

**Estimates of Sampling Error:** Estimated CVs range from 1% to 2% for regional estimates to 3% to 6% for state estimates of livestock inventories.

**Other (Nonsampling) Errors:** Minimized through rigid quality controls on the collection process and careful review of all reported data.

**Sources of Additional Material:** U.S. Department of Agriculture, SRS, *Scope and Methods of the Statistical Reporting Service*, (name changed to National Agricultural Statistics Service), Miscellaneous Publication No. 1308, September 1983 (revised).

**Source and Title:** U.S. Department of Agriculture National Agricultural Statistics Service (NASS), *Multiple Frame Surveys*

**Tables:** See Tables citing NASS in source notes in Section 23, which pertain to this or the following survey.

**Universe, Frequency, and Types of Data:** Surveys of U.S. farm operators to obtain data on major livestock inventories, selected crop acreages and production, grain stocks, and farm labor characteristics; and to obtain farm economic data for price indexing.

**Type of Data Collection Operation:** Primary frame is obtained from general or special purpose lists, supplemented by a probability sample of land areas used to estimate for list incompleteness.

**Data Collection and Imputation Procedures:** Mail, telephone, or personal interviews used for initial data collection. Mail nonrespondent followup by phone and personal interviews. Imputation based on average of respondents.

**Estimates of Sampling Error:** Estimated CV for number of hired farm workers is about 3%. Estimated CVs range from 1% to 2% for regional estimates to 3% to 6% for state estimates of livestock inventories.

**Other (Nonsampling) Errors:** In addition to above, replicated sampling procedures used to monitor effects of changes in survey procedures.

**Sources of Additional Material:** U.S. Department of Agriculture, SRS, *Scope and Methods of the Statistical Reporting Service*, (name changed to National Agricultural Statistics Service), Miscellaneous Publication No. 1308, September 1983 (revised).

**Source and Title:** U.S. Department of Agriculture, National Agricultural Statistics Service (NASS), *Objective Yield Surveys*

**Tables:** See Tables citing NASS in source notes in Section 23, which pertain to this or the preceding survey.

**Universe, Frequency, and Types of Data:** Surveys for data on corn, cotton, potatoes, soybeans, wheat, and rice to forecast and estimate yields.

**Type of Data Collection Operation:** Random location of plots in probability sample *s* selected in June from Basic Area Frame Sample (see above).

**Data Collection and Imputation Procedures:** Enumerators count and measure plant characteristics in sample fields. Production measured from plots at harvest. Harvest loss measured from post harvest gleanings.

**Estimates of Sampling Error:** CVs for national estimates of production are about 2-3%.

**Other (Nonsampling) Errors:** In addition to above, replicated sampling procedures used to monitor effects of changes in survey procedures.

**Sources of Additional Material:** U.S. Department of Agriculture, SRS, *Scope and Methods of the Statistical Reporting Service*, (name changed to National Agricultural Statistics Service), Miscellaneous Publication No. 1308, September 1983 (revised).

(See Section 1 above for information pertaining to the Census of Population and Current Population Survey.)

## SECTION 18. NATURAL RESOURCES

(See Section 14 for Table 855; Section 15 for Table 847; and Section 22 for Tables 864 and 865.)

## SECTION 19. ENERGY AND UTILITIES

**Source and Title:** U.S. Energy Information Administration, *Residential Energy Consumption Survey*

**Tables:** 898 and Table 964 in Section 20.

**Universe, Frequency, and Types of Data:** Quadriennial survey of households and fuel suppliers. Data are obtained on energy-related household characteristics, housing unit characteristics, use of fuels, and energy consumption and expenditures by fuel type.

**Type of Data Collection Operation:** Probability sample of 5,900 eligible units in 116 PSUs. For responding units, fuel consumption and expenditure data obtained from fuel suppliers to those households.

**Data Collection and Imputation Procedures:** Personal interviews. Extensive followup of nonrespondents including

mail questionnaires for some households. Adjustments for nonrespondents were made in weighting for respondents. Most item nonresponses were imputed.

**Estimates of Sampling Error:** Estimated CVs for household averages: For consumption, 1.3%; for expenditures, 1.0%; for various fuels, values ranged from 2.0% for electricity to 7.0% for LPG.

**Other (Nonsampling) Errors:** Household response rate of 81.0%. Nonconsumption data were mostly imputed for mail respondents (2.5% of eligible units). Usable responses from fuel suppliers for various fuels ranged from 80.7% for electricity to 56.6% for fuel oil.

**Sources of Additional Material:** U.S. Energy Information Administration, A look at Residential Energy Consumption in 1997.

(See Section 15 for Table 890 and Section 22 for Table 889)

## SECTION 20. CONSTRUCTION AND HOUSING

**Source and Title:** U.S. Census Bureau, *Monthly Survey of Construction*

**Tables:** 936 and 938, 939, 940, and 941.

### **Universe, Frequency, and Types of**

**Data:** Survey conducted monthly of newly constructed housing units (excluding mobile homes). Data are collected on the start, completion, and sale of housing. (Annual figures are aggregates of monthly estimates.)

### **Type of Data Collection Operation:**

Probability sample of housing units obtained from building permits selected from 17,000 places. For nonpermit places, multistage probability sample of new housing units selected in 169 PSUs. In those areas, all roads are canvassed in selected enumeration districts.

**Data Collection and Imputation Procedures:** Data are obtained by telephone inquiry and field visit.

**Estimates of Sampling Error:** Estimated CV of 3% to 4% for estimates of national totals, but are as high as 20% for estimated totals of more detailed characteristics, such as housing units in multiunit structures.

**Other (Nonsampling) Errors:** Response rate is over 90% for most items. Nonsampling errors are attributed to definitional problems, differences in

interpretation of questions, incorrect reporting, inability to obtain information about all cases in the sample, and processing errors.

**Sources of Additional Material:** U.S. Census Bureau, *Construction Reports, Series C20, Housing Starts; C22, Housing Completions; and C25, New One-Family Houses Sold and For Sale.*

**Source and Title:** U.S. Census Bureau, *Value of Construction Put in Place*

**Tables:** 929, 930, 931, and 932.

### **Universe, Frequency, and Types of**

**Data:** Survey conducted monthly on total value of all construction put in place in the current month, both public and private projects. Construction values include costs of materials and labor, contractors' profits, overhead costs, cost of architectural and engineering work, and miscellaneous project costs. (Annual figures are aggregates of monthly estimates.)

### **Type of Data Collection Operation:**

Varies by type of activity: Total cost of private one-family houses started each month is distributed into value put in place using fixed patterns of monthly construction progress; using a multistage probability sample, data for private multifamily housing are obtained by mail from owners of multiunit projects. Data for residential additions and alterations are obtained in a quarterly survey measuring expenditures; monthly estimates are interpolated from quarterly data. Estimates of value of private nonresidential construction, and state and local government construction are obtained by mail from owners (or agents) for a probability sample of projects. Estimates of farm nonresidential construction expenditures are based on U.S. Department of Agriculture annual estimates of construction; public utility estimates are obtained from reports submitted to federal regulatory agencies and from private utility companies; estimates of federal construction are based on monthly data supplied by federal agencies.

### **Data Collection and Imputation Procedures:**

See "Type of Data Collection Operation." Imputation accounts for approximately 20% of estimated value of construction each month.

**Estimates of Sampling Error:** CV estimates for private nonresidential construction range from 3% for estimated value of industrial buildings to 9% for religious buildings. CV is approximately 2% for total new private nonresidential buildings.

**Other (Nonsampling) Errors:** For directly measured data series based on samples, some nonsampling errors may arise from processing errors, imputations, and misunderstanding of questions. Indirect data series are dependent on the validity of the underlying assumptions and procedures.

**Sources of Additional Material:** U.S. Census Bureau, *Construction Reports*, Series C30, *Value of Construction Put in Place*.

**Source and Title:** U.S. Census Bureau, *American Housing Survey*

**Tables:** See Tables citing *American Housing Survey* in source notes.

**Universe, Frequency, and Types of**

**Data:** Conducted nationally in the fall in odd numbered years to obtain data on the approximately 112 million occupied or vacant housing units in the United States (group quarters are excluded). Data include characteristics of occupied housing units, vacant units, new housing and mobile home units, financial characteristics, recent mover households, housing and neighborhood quality indicators, and energy characteristics.

**Type of Data Collection Operation:**

The national sample was a multistage probability sample with about 53,000 units eligible for interview in 1997. Sample units, selected within 394 PSUs, were surveyed over a 5-month period.

**Data Collection and Imputation Procedures:** For 1997, the survey was conducted by personal interviews. The interviewers obtained the information from the occupants or, if the unit was vacant, from informed persons such as landlords, rental agents, or knowledgeable neighbors.

**Estimates of Sampling Error:** For the national sample, illustrations of the S.E. of the estimates are provided in the Appendix D of the 1997 report. As an example, the estimated CV is about 0.2% for the estimated percentage of owner occupied units with two persons.

**Other (Nonsampling) Errors:** Response rate was about 97%. Nonsampling errors may result from incorrect or incomplete responses, errors in coding and recording, and processing errors. For the 1997 national sample, approximately 1.7% of the total housing inventory was not adequately represented by the AHS sample.

**Sources of Additional Material:**

U.S. Census Bureau, *Current Housing Reports*, Series H-150 and H-170, *American Housing Survey*.

(See Section 1 above for information pertaining to Tables 950 and 951, Section 14 pertaining to Table 927, Section 15 for Table 925; Section 19 for Table 964; and Section 22 for Table 926.)

SECTION 21. MANUFACTURES

**Source and Title:** U.S. Census Bureau, *Annual Survey of Manufactures*

**Tables:** See Tables citing *Annual Survey of Manufactures* in source notes.

**Universe, Frequency, and Types of**

**Data:** Conducted annually to provide basic measures of manufacturing activity for intercensal years for all manufacturing establishments having one or more paid employees.

**Type of Data Collection Operation:**

Sampling frame is 350,000 establishments in the 1987 Census of Manufactures (see above), supplemented by Social Security Administration lists of new manufacturers and new manufacturing establishments of multi-establishment companies identified annually by the Census Bureau's Company Organization Survey. A probability sample of about 55,000 establishments is selected. All establishments of companies with more than \$500 million of manufacturing shipments in 1987 are included with certainty. All establishments with 250+ employees are also included with certainty along with a probability sample of smaller establishments.

**Data Collection and Imputation Procedures:**

Survey is conducted by mail with phone and mail followups of nonrespondents. Imputation (for all nonresponse items) is based on previous year reports, or for new establishments in survey, on industry averages.

**Estimates of Sampling Error:** Estimated standard errors for number of employees, new expenditure, and for value added totals are given in annual publications. For U.S. level industry statistics, most estimated standard errors are 2% or less, but vary considerably for detailed characteristics.

**Other (Nonsampling) Errors:** Response rate is about 85%. Nonsampling errors include those due to collection, reporting, and transcription errors, many of which are corrected through computer and clerical checks.

**Sources of Additional Material:** U.S. Census Bureau, *Annual Survey of Manufactures*, and Technical Paper 24.

(See Section 15 for Tables 972 and 973)

## SECTION 22. DOMESTIC TRADE

**Source and Title:** U.S. Census Bureau, *1997 Economic Census Geographic Area Series and Subject Series* (for NAICS sectors 22, 42, 44-45, 48-49, and 51-81).

**Tables:** See tables citing the above census in source notes in Sections 3, 11, 15, 18, 20, 22, 23, 24, 25, 26, 27, and 29.

**Universe, Frequency, and Types of Data:** Conducted every 5 years to obtain data on number of establishments, number of employees, total payroll size, total sales, and other industry-specific statistics. In 1997, the universe was all employer and nonemployer establishments primarily engaged in wholesale, retail, utilities, finance & insurance, real estate, transportation & warehousing, and other service industries.

**Type of Data Collection Operation:** All large employer firms were surveyed (i.e. all employer firms above the payroll size cutoff established to separate large from small employers) plus a 5% to 25% sample of the small employer firms. Firms with no employees were not required to file a census return.

**Data Collection and Imputation Procedures:** Mail questionnaires were used with both mail and telephone followups for nonrespondents. Data for nonrespondents and for small employer firms not mailed a questionnaire were obtained from administrative records of

the IRS and Social Security Administration or imputed. Nonemployer data were obtained exclusively from IRS 1997 income tax returns.

**Estimates of Sampling Error:** Not applicable for basic data such as sales, revenue, payroll, etc.

**Other (Nonsampling) Errors:** Trade area level unit response rates in 1997 ranged from 85% to 99%. Item response rates ranged from 60% to 90% with lower rates for the more detailed questions. Nonsampling errors may occur during the collection, reporting, and keying of data, and industry misclassification.

**Sources of Additional Material:** U.S. Census Bureau, *1997 Economic Census: Geographic Area Series and Subject Series* (by NAICS sector), Appendix C and <[www.census.gov/epcd/www/econ97.html](http://www.census.gov/epcd/www/econ97.html)>

**Source and Title:** U.S. Census Bureau, *Current Business Surveys*

**Tables:** 1019, 1020, 1021, 1022, and 1047 and tables citing the *Service Annual Survey* in Sections 3, 11, 23, 24, 25, 26, and 27.

**Universe, Frequency, and Types of Data:** Provides monthly estimates of retail sales by kind of business and geographic area, and end-of-month inventories of retail stores; wholesale sales and end-of-month inventories; and annual receipts of selected service industries.

**Type of Data Collection Operation:** Probability sample of all firms from a list frame and, additionally, for retail and service an area frame. The list frame is the Bureau's Standard Statistical Establishment List (SSEL) updated quarterly for recent birth Employer Identification (EI) Numbers issued by the Internal Revenue Service and assigned a kind-of-business code by the Social Security Administration. The largest firms are included monthly; a sample of others is included every three months on a rotating basis. The area frame covers businesses not subjected to sampling on the list frame.

**Data Collection and Imputation Procedures:** Data are collected by mail questionnaire with telephone followups for nonrespondents. Imputation made for each nonresponse item and each item failing edit checks.

**Estimates of Sampling Error:** For the 1989 monthly surveys, CVs are about 0.6% for estimated total retail sales, 1.7% for wholesale sales, 1.3% for wholesale inventories. For dollar volume of receipts, CVs from the *Service Annual Survey* vary by kind of business and range between 1.5% to 15.0%. Sampling errors are shown in monthly publications.

**Other (Nonsampling) Errors:** Imputation rates are about 18% to 23% for monthly retail sales, 20% to 25% for wholesale sales, about 25% to 30% for monthly wholesale inventories and 14% for the *Service Annual Survey*.

**Sources of Additional Material:** U.S. Census Bureau, *Current Business Reports*, *Monthly Retail Trade*, *Monthly Wholesale Trade*, and *Service Annual Survey*.

(See Section 15 for information pertaining to Tables 1017 and 1049.)

#### SECTION 23. TRANSPORTATION

(See Section 15 for Table 1053; and Section 12 for Table 1055; Section 14 for Table 1115; and Section 22 for Tables 1054 1076, 1077, and 1110.)

#### SECTION 24. INFORMATION AND COMMUNICATIONS

(See Section 15 for Table 1121 and Section 22 for Tables 1119, 1120, 1129, 1137, 1139, 1140, 1143, 1145, 1146, 1147, and 1151)

#### SECTION 25. BANKING, FINANCE AND INSURANCE

(See Section 13 for information pertaining to Tables 1167, 1169, 1170, 1171, 1192 and 1209; Section 15 for information pertaining to Table 1162 and 1228; and Section 22 for Tables 1161, 1221, and 1227.)

#### SECTION 26. ARTS, ENTERTAINMENT, AND RECREATION

(See Section 15 for Table 1232 and Section 22 for Tables 1230 and 1231.)

#### SECTION 27. ACCOMMODATION, FOOD SERVICES, AND OTHER SERVICES

(See Section 15 for Tables 1263, 1264, 1267, and 1280; and Section 22 for Tables 1265, 1269, 1270, 1273, 1274, 1275, 1276, 1277, and 1278)

#### SECTION 28. FOREIGN COMMERCE AND AID

**Source and Title:** U.S. Census Bureau, *Foreign Trade—Export Statistics*

**Tables:** See Census Bureau citations for export statistics in source notes in Section 28 and also Table 1317 in Section 29.

**Universe, Frequency, and Types of Data:** The export declarations collected by Customs are processed each month to obtain data on the movement of U.S. merchandise exports to foreign countries. Data obtained include value, quantity, and shipping weight of exports by commodity, country of destination, Customs district of exportation, and mode of transportation.

**Type of Data Collection Operation:** Shipper's Export Declarations are required to be filed for the exportation of merchandise valued over \$1,500. Customs officials collect and transmit the documents to the Census Bureau on a flow basis for data compilation. Value data for shipments valued under \$1,501 are estimated, based on established percentages of individual country totals.

**Data Collection and Imputation Procedures:** Statistical copies of Shipper's Export Declarations are received on a daily basis from Customs ports throughout the country and subjected to a monthly processing cycle. They are fully processed to the extent they reflect items valued over \$1,500. Estimates for shipments valued at \$1,500 or less are made, based on established percentages of individual country totals.

**Estimates of Sampling Error:** Not applicable.

**Other (Nonsampling) Errors:** Clerical and complex computer checks intercept most processing errors and minimize otherwise significant reporting errors; other nonsampling errors are caused by undercounting of exports to Canada due to the nonreceipt of some Shipper's Export Declarations.

**Sources of Additional Material:** U.S. Census Bureau, *U.S. Merchandise Trade: Exports, General Imports, and Imports for Consumption, SITC, Commodity by Country*, FT 925 (discontinued after 1996).

**Source and Title:** U.S. Census Bureau, *Foreign Trade—Import Statistics*

**Tables:** See Census Bureau citations for import statistics in source notes in Section 28 and also Table 1317 in Section 29.

**Universe, Frequency, and Types of**

**Data:** The import entry documents coll are processed each month to obtain data on the movement of merchandise imported into the United States. Data obtained include value, quantity, and shipping weight by commodity, country of origin, Customs district of entry, and mode of transportation.

**Type of Data Collection Operation:**

Import entry documents are required to be filed for the importation of goods into the United States valued over \$1,000 or for articles which must be reported on formal entries. Customs officials collect and transmit statistical copies of the documents to the Census Bureau on a flow basis for data compilation. Estimates for shipments valued under \$1,001 and not reported on formal entries are based on established percentages of individual country totals.

**Data Collection and Imputation Proce-**

**dures:** Statistical copies of import entry documents, received on a daily basis from Customs ports of entry throughout the country, are subjected to a monthly

processing cycle. They are fully processed to the extent they reflect items valued at \$1,001 and over or items which must be reported on formal entries.

**Estimates of Sampling Error:** Not applicable.

**Other (Nonsampling) Errors:** Verification of statistical data reporting by customs officials prior to transmittal and a subsequent program of clerical and computer checks are utilized to hold nonsampling errors arising from reporting and/or processing errors to a minimum.

**Sources of Additional Material:** U.S. Census Bureau, *U.S. Merchandise Trade: Exports, General Imports and Imports for Consumption, SITC, Commodity by Country*, FT 925 (discontinued after 1996).

## SECTION 29. OUTLYING AREAS

(See Section 1 for information pertaining to Tables 1309, 1311, and 1323; Section 2 for Table 1308; and Section 28 for Table 1317.)