

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may restrict data collection to a *survey universe* based on some available list, such list may be inaccurate and out of date.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 (PSU's). For example, in a national multistage household sample, PSU's 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 PSU's 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 Standard Statistical Establishment List (SSEL). The SSEL 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) *nonsampling 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 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 values. These same errors usually have a nonrandom component. 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 with in 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: Bureau of the Census, *Census of Population*

Tables: See tables citing *Census of Population* in section 1 and also in sections 2, 4, 6, 8, 13, 14, 21, 25, 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 inhabitants.

Type of Data Collection Operation: In 1970, 1980, and 1990 complete census for some items—age, sex, race, marital status, and relationship to household head. 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, 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 and 1990, 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 CV's 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 undercoverage amounted to about 1.5 to 2 percent of the total resident population.

Sources of Additional Material: U.S. Bureau of the Census, *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. Bureau of the Census, *Current Population Survey (CPS)*

Tables: See tables citing *Current Population Reports* primarily in section 1, but also in sections 2, 3, 4, 8, 12, 13, 14, 18, 23, 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: Multi-stage probability sample of about 60,000 households in 729 PSU's in 1993. Over-sampling in some States and the largest MSA's 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 CV's 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 CV's for family income and poverty rate for all persons in 1986 are 0.5% and 1.5%, respectively. CV's for sub-national 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. Four to 5.0% of sample households unavailable for interviews.

Sources of Additional Material: U.S. Bureau of the Census and Bureau of Labor Statistics, *Concepts and Methods Used in Labor Force Statistics from Current Population Survey* (Census series P-23, No. 62; BLS Report No. 463) and Bureau of the

Census, *Current Population Survey* (Tech. Paper 40) and Bureau of Labor Statistics, *Employment and Earnings*, monthly, Explanatory Notes and Estimates of Error, tables 1-A through 1-H and *BLS Handbook of Methods*, Chapter 1 (Bulletin 2414.)

SECTION 2. VITAL STATISTICS

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

Tables: See tables citing *Vital Statistics of the United States*; 319 in section 5; and 1318 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, CV's 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 *Monthly Vital Statistics Report*.

(See section 1 above for information pertaining to tables 102-107.)

SECTION 3. HEALTH AND NUTRITION

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

Tables: 180, 183, 197, 206, 207, 212, 217, 218, 221, and 228.

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: Multi-stage probability sample of 42,000 households (in 376 PSU's) selected in groups of about four adjacent households.

Data Collection and Imputation Procedures: Personal household interviews with extensive followup of nonrespondents. Data are adjusted for nonresponse by imputation procedure based on "average" characteristics of persons in interviewed households in the same geographic area.

Estimates of Sampling Error: Estimated CV's: 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 95.7% in 1985 for the NHIS.

Sources of Additional Material: U.S. National Center for Health Statistics, "Current Estimates from the National Health Interview Survey, U.S., 1983," Vital and Health Statistics, series 10.

(See section 13 for information pertaining to table 175, section 15 for table 169, and section 27 for table 174.)

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: 283, 284, 287 and 288.

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 (in alternate years) 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: Approximately 87% response rate.

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), Financial Statistics of Institutions of Higher Education*.

Tables: 234, 237, 287, and 290.

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 1990, a 87% response rate. Imputed expenditures amounted to about 2.8% of total expenditures.

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

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), Degrees and Other Formal Awards Conferred*.

Tables: 303-308.

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 (in alternate years).

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 1989-90, approximately 92.3% response rate.

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

(See sections 1 and 9 for information pertaining to the Bureau of the Census.)

SECTION 5. LAW ENFORCEMENT, COURTS, AND PRISONS

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

Tables: 313-318, 320-323, 328-330, and 332.

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 96% of the population (97.4% in MSA's, 90.6% in "other cities," and 87.8% 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*.

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

Tables: 324-327.

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 PSU's 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; 8,000 households are interviewed monthly. Personal interviews are used in the first and fifth interviews; the intervening interviews are conducted by telephone whenever possible.

Estimates of Sampling Error: CV's in 1994: 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 319.)

SECTION 7. PARKS AND RECREATION

(See section 27 for details on tables 408 and 431.)

SECTION 8. ELECTIONS

(See section 1 above for information pertaining to tables 462 and 463 and section 9 for information pertaining to tables 456.)

SECTION 9. STATE AND LOCAL GOVERNMENT FINANCES AND EMPLOYMENT

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

Tables: See tables in section 9 citing *Census of Governments* and table 456 in section 8.

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. Bureau of the Census, *Census of Governments, 1992*, various reports, and *State Government Finances in 1992*, GF 90, No. 3.

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

Tables: See tables citing *Public Employment and Government Finances* in section 9 and table 261 in section 4; and table 591 in section 12.

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 83,000).

Type of Data Collection Operation: Sample of about 13,000 units includes all State governments 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 nonrespondents are imputed from previous year data or obtained from secondary sources, if available.

Estimates of Sampling Error: CV's 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. Bureau of the Census, *Public Employment in 1992*, GE 90, No. 1, *Governmental Finances in 1991-92*, GF 90, No. 5, and *Census of Governments, 1992*, various reports.

SECTION 10. FEDERAL GOVERNMENT

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

Tables: 529-533.

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 CV's for tax year 1987: Adjusted gross income less deficit .13%; salaries and wages .20%; and tax-exempt interest received 4.51%. (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 12. SOCIAL INSURANCE AND HUMAN SERVICES

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

Tables: 587 and 588.

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: 602-605.

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 3 for information pertaining to the National Center for Health Statistics, and section 9 for information pertaining to Annual Surveys of State and Local Government.)

SECTION 13. LABOR FORCE, EMPLOYMENT, AND EARNINGS

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

Tables: 660-662, 667; in section 3, table 175; in section 21, table 991; and in section 22, table 1059.

Universe, Frequency, and Types of Data:

Monthly survey covering about 6 million nonagricultural establishments to obtain data on employment, hours, and earnings, by industry.

Type of Data Collection Operation: Sample survey of over 390,000 establishments in March 1993.

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: Estimated CV's for employment, 0.1%, for average weekly hours paid, 0.1% and for average hourly earnings, 0.2%.

Other (nonsampling) Errors: Estimates of employment adjusted annually to reflect complete universe. Average adjustment is 0.2%.

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-G.

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

SECTION 14. INCOME, EXPENDITURES, AND WEALTH

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

Tables: 747 and 748, and in section 16, 778, 780, 782 and 800.

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 forty-eight 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 Bureau of the Census.)

SECTION 15. PRICES

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

Tables: 751-754, 767, 768, and in section 3, table 169.

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 56 urban areas; thereafter, in 85 PSU's, 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 57,000 housing units, and 19,000 other reporters in 85 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 five largest geographic areas and every other month in others.

Estimates of Sampling Error: Estimates of standard errors are not available at present.

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, *The Consumer Price Index: 1987 Revision*, Report 736, and *BLS Handbook of Methods*, Chapter 19, Bulletin 2414.

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

Tables: 751, 757-759, and in section 21, table 1042; in section 24, table 1139; and in section 25, table 1177.

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 3,100 about 75,000 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 based on total net selling value of all commodities in 1982.

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 16, Bulletin 2414.

SECTION 16.

(See section 14 for information pertaining to tables 778, 780-782, and 800 and section 17 for information pertaining to table 773.)

SECTION 17. BUSINESS ENTERPRISE

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

Tables: 844-847, and in section 16, table 773; in section 27 tables 1263 and 1283; and in section 29 table 1327.

Universe, Frequency, and Types of Data: Annual tabulation of basic data items extracted from the Standard Statistical Establishment List, a file of all known single and multiestablishment companies maintained and updated by the Bureau of the

Census. 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 multi-establishment 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. Bureau of the Census, *General Explanation of County Business Patterns*.

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

Tables: 833-836.

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 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 CV's 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: 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, Sole Proprietorship Returns* (for years through 1980) and *Statistics of Income Bulletin*, vol. 10, No. 1 (summer 1990).

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

Tables: 833-835, 837, and 838.

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 Procedures: 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: Estimated CV's 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: 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, Partnership Returns and Statistics of Income Bulletin*, vol. 10, No. 1 (summer 1990).

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

Tables: 833-835 and 841-843.

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 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 and to achieve statistical definitions.

Estimates of Sampling Error: Estimated CV's 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.

SECTION 18. COMMUNICATIONS

(See section 26 for information pertaining to table 905.)

SECTION 19. ENERGY

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

Tables: 928, 929 and table 1207 in section 25.

Universe, Frequency, and Types of Data: Triennial 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 7,183 eligible units in 129 PSU's. 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 CV's for household averages: For consumption, 1.3%; for expenditures, 1.0%; for various fuels, values ranged from 1.4% for electricity to 5.9% for LPG.

Other (nonsampling) Errors: Household response rate of 86.7%. Nonconsumption data were mostly imputed for mail respondents (5.2% of eligible units). Usable responses from fuel suppliers for various fuels ranged from 82.8% for electricity to 55.7% for fuel oil.

Sources of Additional Material: U.S.

Energy Information Administration, *Household Energy Consumption and Expenditures, 1990 and Housing Characteristics, 1990*.

SECTION 21.

TRANSPORTATION—LAND

(See section 1 for information pertaining to table 1014, section 13 for table 991, and section 15 for table 1042 and section 27 for table 1033.)

SECTION 22. TRANSPORTATION—AIR AND WATER

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

Tables: See Bureau of the Census citations for export statistics in source notes in sections 22 and 28 and also table 1326 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 Bureau of the Census 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. Bureau of the Census, *U.S. Merchandise Trade: Exports, General Imports, and Imports for Consumption, SITC, Commodity by Country*, FT 925.

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

Tables: See Bureau of the Census citations for import statistics in source notes in sections 22 and 28 and also table 1326 in section 29.

Universe, Frequency, and Types of Data: The import entry documents collected by Customs 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 Bureau of the Census 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 Procedures: 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. Bureau of the Census, *U.S. Merchandise Trade: Exports, General Imports and Imports for Consumption, SITC, Commodity by Country*, FT 925.

(See section 13 for information pertaining to table 1059.)

SECTION 23. AGRICULTURE

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

Tables: 1080-1087.

Universe, Frequency, and Types of Data: Complete count of U.S. farms and ranches conducted once every five 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 Procedures: 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 CV's 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. Bureau of the Census, *1992 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 Procedures:

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 CV's 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 CV's 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 samples 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: CV's 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 24. NATURAL RESOURCES

(See section 15 for table 1139.)

SECTION 25. CONSTRUCTION AND HOUSING

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

Tables: 1185 and 1187-1189.

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 PSU's. 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. Bureau of the Census, *Construction Reports*, series C20, *Housing Starts*; C22, *Housing Completions*; and C25, *New One-Family Houses Sold and For Sale*.

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

Tables: 1179-1181.

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 buildings, 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 for all other private construction (nonbuilding) are obtained by phasing F. W. Dodge contract award data; 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 building 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. Bureau of the Census, *Construction Reports*, series C30, *Value of New Construction Put in Place*.

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

Tables: See tables citing *Census of Housing* in source notes in section 25.

Universe, Frequency, and Types of Data: Census of all occupied and vacant housing, excluding group quarters, conducted every 10 years as part of the decennial census (see section 1 above) to determine characteristics of U.S. housing.

Type of Data Collection Operation: For 1970, 1980, and 1990, a complete count of some housing items (e.g. Owned or rented, and value). In 1970, other items collected from 5% and 15% probability samples selected from two sets of detailed questions on housing (these two sets having some common items). In 1980, approximately 19% of the housing units were included in the sample; in 1990, approximately 17%.

Data Collection and Imputation Procedures: In 1970, a self-enumeration census using a mail-out/mail-back procedure was used in most areas. In 1980 and 1990, mail questionnaires were used in even more areas than in 1970, with personal interviews

in the remainder. Followup for nonrespondents and identification of vacant units done by phone and personal visit.

Estimates of Sampling Error: Sampling errors for data are estimated for all items collected by sample and vary by characteristic and geographic area.

Other (nonsampling) Errors: Evaluation studies for 1980 estimated the underenumeration of occupied housing units at 1.5%. The missed rate in 1980 for all units was 2.6% or approximately 2.3 million units, 1 million of which were vacant housing units.

Sources of Additional Material: U.S. Bureau of the Census, *1980 Census of Population and Housing, The Coverage of Housing in the 1980 Census*, PHC80-E1, July 1985.

Source and Title: U.S. Bureau of the Census, *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 103 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 51,300 units eligible for interview in 1987. Sample units, selected within 394 PSU's, were surveyed over a 5-month period.

Data Collection and Imputation Procedures: For 1987, 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 B of the 1987 report. As an example, the estimated CV is about 0.5% 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 1985 national sample, approximately 6% of the total housing inventory was not adequately represented by the AHS sample.

Sources of Additional Material: U.S. Bureau of the Census, *Current Housing Reports*, series H-150 and H-170, *American Housing Survey*.

(See section 1 above for information pertaining to the Census of Population, section 15 pertaining to table 1177, and section 19 for table 1207.)

Section 26. MANUFACTURES

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

Tables: See tables citing *Census of Manufactures* in source notes in section 26 and also table 905 in section 18 and table 1330 in section 29.

Universe, Frequency, and Types of Data: Conducted every 5 years to obtain information on labor, materials, capital input and output characteristics, plant location, and legal form of organization for all plants in the United States with one or more paid employees. Universe was 350,000 manufacturing establishments in 1987.

Type of Data Collection Operation: Complete enumeration of data items obtained from 200,000 firms. Administrative records from Internal Revenue Service and Social Security Administration are used for 150,000 smaller single-location firms, which were determined by various cutoffs based on size and industry.

Data Collection and Imputation Procedures: Five mail and telephone followups for larger nonrespondents. Data for small single-location firms (generally those with fewer than 10 employees) not mailed census questionnaires were estimated from administrative records of IRS and SSA. Data for nonrespondents were imputed from related responses or administrative records from IRS and SSA. Approximately 8% of total value of shipments was represented by fully imputed records in 1987.

Estimates of Sampling Error: Not applicable.

Other (nonsampling) Errors: Based on evaluation studies, estimates of nonsampling errors for 1972 were about 1.3% for estimated total payroll; 2% for total employment; and 1% for value of shipments. Estimates for later years are not available.

Sources of Additional Material: U.S. Bureau of the Census, *1987 Census of Manufactures, Industry Series, Geographic Area Series, and Subject Series*.

Source and Title: U.S. Bureau of the Census, *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 inter-censal 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. Bureau of the Census, *Annual Survey of Manufactures*, and Technical Paper 24.

Section 27. DOMESTIC TRADE AND SERVICES

Source and Title: U.S. Bureau of the Census, *Census of Wholesale Trade, Census of Retail Trade, Census of Service Industries*

Tables: See tables citing the above censuses in source notes in section 27 and table 1330 in section 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

1987, the universe was all employer establishments primarily engaged in wholesale trade, and employer and nonemployer establishments in retail trade or service industries.

Type of Data Collection Operation: All wholesale firms with paid employees surveyed; all retail and service large employer firms surveyed (i.e. all employer firms above the payroll size cutoff established to separate large from small employers) plus a 10-percent sample of smaller employer firms. Firms with no employees were not required to file a census return.

Data Collection and Imputation Procedures: Mail questionnaire is utilized with both mail and telephone followups for nonrespondents. Data for nonrespondents and all employer firms in retail trade and service industries are obtained from administrative records of IRS and the Social Security Administration.

Estimates of Sampling Error: Not applicable.

Other (nonsampling) Errors: Response rate in 1987 of 80% for single establishment firms; 83% for multi-establishment firms. Item response ranged from 60% to 90% with higher rates for less detailed questions.

Sources of Additional Material: U.S. Bureau of the Census, Appendix A of *Census of Retail Trade; Census of Service Industries; Census of Wholesale Trade; and History of the 1987 Economic Censuses*, April 1992.

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

Tables: 1266-1268, 1282, 1288, 1291, and table 174 in section 3, tables 408 and 431 in section 7, and table 1033 in section 21.

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 Proce-

dures: 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, CV's are about 0.6% for estimated total retail sales, 1.7% for wholesale sales, 1.3% for wholesale inventories. For dollar volume of receipts, CV's 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.

Bureau of the Census, *Current Business Reports, Monthly Retail Trade, Monthly Wholesale Trade, and Service Annual Survey*.

(See section 17 for information pertaining to tables 1263 and 1283.)

Section 29. OUTLYING AREAS

(See section 1 for information pertaining to tables 1319 and 1320, section 2 for table 1318, section 17 for table 1327, section 22 for table 1326, sections 26 and 27 for table 1330.)