# Appendix A.—GENERAL INFORMATION CONCERNING THE DATA

PRESENTATION OF THE DATA .. App-1
Derived figures (percents, medians, etc.) ..... App-1
Symbols ..... App-1
Boundaries ..... App-1
DATA COLLECTION PROCEDURES ..... App-1
PROCESSING PROCEDURES .... App-2

## PRESENTATION OF THE DATA

**Derived figures (percents, medians, etc.).**—Percents, medians, and means, as well as certain rates and ratios are shown in Volume II reports. For all types of derived figures in this series of reports, the figure is not presented (but indicated by three dots "...") if the base is smaller than the minimum number prescribed for the sample on which the figure is based. The minimum bases for 1970 are 100 (persons, families, or households) for the 20-percent sample, 132 for the 15-percent, and 400 for the 5-percent.

The reader should exercise caution in the use and interpretation of data for very small subgroups of the population; these data are especially subject to the effects of sampling variability, misreporting, and processing errors.

Percents that round to less that 0.1 are not shown but indicated as zero (i.e., "–"). The median, which is a type of average, is the middle value in a distribution; i.e., the median divides the distribution into two equal parts—one-half of the cases fall below the median and one-half of the cases exceed the median. When the median falls in the lower terminal category of an open-end distribution, the method of presentation is to show the initial value of the next category followed by a minus sign; thus, for example, if the median falls in the category "Less than $250," it is shown as "$250–." When the median falls in the upper terminal category of an open-end distribution, the method of presentation is to show the initial value of the terminal category followed by a plus sign; thus, for example, if the median falls in the category "$10,000 and over," it is shown as "$10,000+." The mean is the arithmetic average derived by adding the values in a particular distribution and dividing by the number of units in the distribution.

**Symbols.**—A dash "–" signifies zero. Three dots "..." mean not applicable, or that the base for the derived figure is too small for it to be shown. The symbol "NA" means not available.

**Boundaries.**—The data shown for 1970 relate to boundaries as they existed on January 1, 1970. Information on boundary changes between 1960 and 1970 for certain types of areas is given in the Volume I report for each State.

## DATA COLLECTION PROCEDURES

The 1970 census was conducted primarily through self-enumeration. In 1960, self-enumeration was first introduced on a nationwide scale as a substitute for the traditional census direct interview.

A census questionnaire was delivered by postal carriers to every household several days before Census Day, April 1, 1970. This questionnaire contained certain explanatory information and was accompanied by an instruction sheet; in areas with comparatively large proportions of Spanish-speaking persons, a Spanish version of the instruction sheet was also enclosed. Facsimiles of the questionnaires and instructions to respondents are included in the Volume I reports.

In the larger metropolitan areas and some adjacent counties, altogether containing about three-fifths of the population of the United States, the householder was requested to fill out and mail back the form on Census Day. Approximately 87 percent of the householders returned their forms by mail. The mailed-back forms were reviewed by the census enumerator (or, in some localities, a census clerk) and if the form was determined to be incomplete or inconsistent, a followup was made. The bulk of these followups were made by telephone, the rest by personal visit. For the households which did not mail back their forms, a followup was also made, in almost all cases by personal visit and in the remainder by telephone.

For the remaining two-fifths of the population, the householder was requested to fill out the form and give it to the enumerator when he called; approximately 80 percent did so. Incomplete and unfilled forms were completed by interview during the enumerator's visit.

Three types of questionnaires were used throughout the country; 80 percent of the households answered a form containing a limited number of population and housing questions, and the remainder, split into 15-percent and 5-percent samples, answered forms which contained these questions as well as a number of additional questions. Some of the additional questions were the same on the 15-percent and 5-percent versions; others were different. A random procedure was used to determine which of the three forms any particular household answered.

In the metropolitan and adjacent areas, the designated type was sent to each household. In the remaining

areas, the questionnaire with a limited number of questions was distributed to all households, and the enumerators asked the additional questions in those households designated for the 15-percent and 5-percent samples.

## PROCESSING PROCEDURES

The 1970 census questionnaires were specially designed to be processed by FOSDIC (Film Optical Sensing Device for Input to Computer). For most items on the questionnaire, the information supplied by the respondent or obtained by the enumerator was indicated by marking the answers in predesignated positions that would be "read" by FOSDIC from a microfilm copy of the questionnaire onto computer magnetic tape with no intervening manual processing. Among the items covered in these reports, however, a number required reviewing written entries to determine the proper code. Consequently, the processing involved a manual coding and editing operation in which clerks determined the appropriate codes and marked the specified positions on the questionnaire; for example, the clerks applied a 2-digit numerical code for the State-of-birth entry. These marks as well as those made by the respondent and enumerator were read by FOSDIC onto magnetic tape.

The tape containing the information from the questionnaires was processed on the Census Bureau's computers through a number of editing and tabulation steps (see Appendix D, "Accuracy of the Data"). Figures shown in this series of reports may differ from comparable figures in Volume I because of differences in processing.

One of the end results of the tabulation operation was a computer tape from which the tables in this report were prepared on a cathode-ray-tube phototypesetting machine at the Government Printing Office. Another end result was the summary tapes which are available for purchase, as described in Appendix E, "Publication and Computer Summary Tape Program."

# Appendix B.—AREA CLASSIFICATIONS

USUAL PLACE OF RESIDENCE . . . . . App-3
URBAN AND RURAL RESIDENCE . . . . . App-3
Definition . . . . . App-3
Farm and nonfarm residence . . . . App-3
REGIONS AND GEOGRAPHIC DIVISIONS . . . . . App-3
COUNTIES . . . . . App-3
URBANIZED AREAS . . . . . App-3
STANDARD METROPOLITAN STATISTICAL AREAS . . . . . App-3

## USUAL PLACE OF RESIDENCE

In accordance with census practice dating back to 1790, each person enumerated in the 1970 census was counted as an inhabitant of his usual place of residence, which is generally construed to mean the place where he lives and sleeps most of the time. This place is not necessarily the same as his legal residence, voting residence, or domicile. In the vast majority of cases, however, the use of these different bases of classification would produce substantially the same statistics, although there may be appreciable differences for a few areas. (The Volume I reports of the 1970 Census of Population contain a more complete discussion of the residence rules for certain categories of persons whose usual place of residence is not immediately clear.)

## URBAN AND RURAL RESIDENCE

**Definition.**—The urban population comprises all persons living in urbanized areas and in places of 2,500 inhabitants or more outside urbanized areas. More specifically, the urban population consists of all persons living in (a) places of 2,500 inhabitants or more incorporated as cities, villages, boroughs (except Alaska), and towns (except in the New England States, New York, and Wisconsin), but excluding those persons living in the rural portions of extended cities;[1] (b) unincorporated places of 2,500 inhabitants or more; and (c) other territory, incorporated or unincorporated, included in urbanized areas. The population not classified as urban constitutes the rural population.

**Farm and nonfarm residence.**—The rural population is subdivided into the rural-farm population, which comprises all rural residents living on farms, and the rural-nonfarm population, which comprises the remaining rural population. As in the 1960 census, the farm population consists of persons living on places of 10 or more acres from which sales of farm products amounted to $50 or more in the preceding calendar year or on places of less than 10 acres from which sales of farm products amounted to $250 or more in the preceding year.

## REGIONS AND GEOGRAPHIC DIVISIONS

For statistical presentation, the United States is sometimes divided into regions. The regions may be subdivided into geographic divisions, which are combinations of States. Generally, four regions are shown—Northeast, North Central, South, and West—but the first two may be combined into one region, the North. In reports that present statistics for regions, a map following the contents page shows the States included in each region and division.

[1] Extended cities are so designated because they have one or more large portions (normally at the boundary of the city) with relatively low population density. These portions are classified as rural, and the residents, as a rule, are not included in the urban population.

## COUNTIES

The primary divisions of the States are, in general, termed counties, but in Louisiana these divisions are known as parishes. There are no counties in Alaska. In this State, statistical areas called census divisions were developed for general statistical purposes through the cooperation of the State and the Census Bureau and are treated as county equivalents. In four States (Maryland, Missouri, Nevada, and Virginia), there are one or more cities which are independent of any county organization and thus constitute primary divisions of their States.

## URBANIZED AREAS

An urbanized area, generally, consists of at least one city of 50,000 inhabitants or more in 1970 and the surrounding closely settled area that meets certain criteria of population density or land use. An urbanized area may be subdivided into the central city or cities, and the remainder of the area or "urban fringe." The central city portion, generally, consists of the population of the city or cities named in the title of the urbanized area.

## STANDARD METROPOLITAN STATISTICAL AREAS

Except in the New England States, a standard metropolitan statistical area is a county or group of contiguous counties which contains at least one city of 50,000 inhabitants or more, or "twin cities" with a combined population of at least 50,000. In addition to the county, or counties, containing

such a city or cities, contiguous counties are included in an SMSA if, according to certain criteria, they are socially and economically integrated with the central city or cities. In the New England States, SMSA's consist of towns and cities instead of counties. Each SMSA must include at least one central city, and the complete title of an SMSA identifies the central city or cities.

The population living in SMSA's is designated as the metropolitan population. The population living outside SMSA's constitutes the nonmetropolitan population. The metropolitan population may be subdivided into those living in the central city or cities and those living in the balance of the SMSA.

# Appendix C.—DEFINITIONS AND EXPLANATIONS OF SUBJECT CHARACTERISTICS


GENERAL ..... App-5

SOCIAL CHARACTERISTICS ... App-5

Age ..... App-5
Race ..... App-5
Nativity, place of birth, and parentage ..... App-5
Country of origin ..... App-6
Citizenship ..... App-6
Year of immigration ..... App-6
Mother tongue ..... App-6
Spanish heritage ..... App-6
Spanish origin ..... App-6
Year moved into present house ... App-7
Residence in 1965 ..... App-7
School enrollment ..... App-7
Years of school completed ..... App-7
Vocational training ..... App-8
Veteran status ..... App-8
Marital status ..... App-8
Marital history ..... App-8
Household ..... App-8
Relationship to head of household ..... App-8
Group quarters ..... App-9
Inmate of institution ..... App-9
Married couple ..... App-9
Family and subfamily ..... App-9
Own child and related child ..... App-9
Unrelated individual ..... App-10
Children ever born ..... App-10
Work disability ..... App-10

ECONOMIC CHARACTERISTICS App-10

Employment status ..... App-10
Hours worked ..... App-11
Weeks worked in 1969 ..... App-11
Year last worked ..... App-11
Place of work ..... App-11
Means of transportation to work .. App-11
Occupation and industry ..... App-11
Class of worker ..... App-11
Activity 5 years ago ..... App-11
Occupation 5 years ago ..... App-12
Income in 1969 ..... App-12
Poverty status in 1969 ..... App-12


## GENERAL

As stated in Appendix A, the 1970 census was conducted primarily through self-enumeration. The principal determinant for the responses was, therefore, the questionnaire and its accompanying instruction sheet. Furthermore, census takers were instructed, in their telephone and personal-visit interviews, to read the questions directly from the questionnaire. The definitions and explanations given below for each subject are largely drawn from various technical and procedural materials used in the collection of the data. This material helped the enumerative personnel to understand more fully the intent of each question and thus to resolve problem or unusual cases in a manner consistent with this intent. Also included is certain explanatory information to assist the user in the proper utilization of the statistics. Some terms that are briefly defined in this report are more fully explained in the Volume I reports. Volume I also contains facsimiles of the questions and instructions to respondents. More complete definitions of the terms relating to the central subject of this report are presented in the introduction.

## SOCIAL CHARACTERISTICS

**Age.**—The age classification is based on the age of the person in completed years as of April 1, 1970, and was determined from the reply to questions on age and on month and year of birth.

**Race.**—Information on race was obtained primarily through self-enumeration, and the data represent essentially self-classification by people according to the race with which they identify themselves. For persons of mixed parentage who were in doubt as to their classification, the race of the person's father was to be used.

The category "white" includes persons who indicated their race as white, as well as persons who did not classify themselves in one of the specific race categories on the questionnaire but entered Mexican, Puerto Rican, or a response suggesting Indo-European stock.

The category "Negro" includes persons who indicated their race as Negro or black, as well as persons who did not classify themselves in one of the specific race categories on the questionnaire but had such entries as Jamaican, Trinidadian, West Indian, Haitian, and Ethiopian. The term "Negro and other races" includes persons of all races other than white.

The category "Indian" includes persons who indicated their race as American Indian or who did not indicate a specific race category but reported the name of an Indian tribe.

Additional race categories shown separately in some reports are Japanese, Chinese, Filipino, Hawaiian, Korean, Eskimo, and Aleut.

The residual category "all other races" appearing in some tables includes all races for which data are not shown separately in the particular table.

**Nativity, place of birth, and parentage.**—The category "native" comprises persons born in the United States, in the Commonwealth of Puerto Rico, in an outlying area of the United States, or at sea. Also included in this category is the small number of persons who, although they were born in a foreign country, have at least one American parent. Persons not classified as native are classified as foreign

born. Native persons are classified according to their State or region of birth. Respondents were instructed to report place of birth in terms of the mother's usual State of residence at the time of the birth rather than in terms of the location of the hospital if the birth occurred in a hospital.

Information on birthplace of parents is used to classify the native population into two categories: "native of native parentage" and "native of foreign or mixed parentage." The category "native of native parentage" comprises native persons with both parents born in the United States. The category "native of foreign or mixed parentage" includes native persons with one or both parents foreign born. The rules for determining the nativity of parents are substantially the same as those for determining the nativity of the person himself.

The category "foreign stock" includes the foreign-born population and the native population of foreign or mixed parentage. Persons of foreign stock thus comprise all first- and second-generation Americans.

**Country of origin.**—Persons of foreign stock are classified according to their country of origin; that is, country of birth of the foreign born and birthplace of parents for native persons of foreign or mixed parentage. Respondents were instructed to report the country of birth according to international boundaries as recognized by the United States government on April 1, 1970. Native persons of foreign parentage whose parents were born in different foreign countries are classified according to the country of birth of the father.

**Citizenship.**—Citizens include foreign-born persons who reported that they had been naturalized and all persons classified as native (see definition above). Foreign-born persons who have not been naturalized are classified as alien.

**Year of immigration.**—The data on year of immigration relate to the foreign born and indicate the period covering the year when the person came to stay permanently in the United States.

**Mother tongue.**—Mother tongue refers to the language spoken in the person's home when he was a child. If both English and another mother tongue were reported, preference was given to the language other than English. The data on mother tongue may not reflect a person's current language skills since the vast majority of persons reporting a mother tongue other than English have learned to speak English during or after their childhood.

**Spanish heritage.**—The term "Spanish heritage" relates to statistics based on the 15-percent sample. The population of Spanish heritage is composed of population groups identified in various ways in different parts of the country, using information derived from the 15-percent sample. In 42 States and the District of Columbia, this population is identified as "persons of Spanish language"; in five Southwestern States, as "persons of Spanish language or Spanish surname"; and in the three Middle Atlantic States, as "persons of Puerto Rican birth or parentage." The specific definitions involved in identifying these population groups are given below.

*Spanish language.*—Persons of Spanish language comprise persons of Spanish mother tongue and all other persons in families in which the head or wife reported Spanish as his or her mother tongue.

*Spanish surname.*—In five Southwestern States (Arizona, California, Colorado, New Mexico, and Texas) persons with Spanish surnames were identified by means of a list of over 8,000 Spanish surnames originally compiled by the Immigration and Naturalization Service[1] (and later updated by the Bureau of the Census). The population of Spanish heritage in the five Southwestern States consists of persons of Spanish language combined with all other persons of Spanish surname.

*Puerto Rican birth or parentage.*—The population of Puerto Rican birth or parentage includes persons born in Puerto Rico and persons born in the United States or an outlying area with one or both parents born in Puerto Rico. The population of Spanish heritage in New York, New Jersey, and Pennsylvania consists of persons of Puerto Rican birth or parentage.

**Spanish origin.**—The term "Spanish origin" relates to statistics based on the 5-percent sample. A person is classified as being of Spanish origin if his or her reply to the question on origin or descent was any of the following: Mexican, Puerto Rican, Cuban, Central or South American, or other Spanish.

---

[1] U.S. Immigration and Naturalization Service, **Supplement to Manual of Immigration, Spanish—Spanish Personal Names,** selected by Inspector George Lockwood, New York, 1936.

**Year moved into present house.**—The data on year moved into present house refer to the most recent move the person made. Thus, a person who had moved back into the same house (or apartment) in which he had previously lived was asked to give the date at which he began the present occupancy. If a person had moved from one apartment to another in the same building, he was expected to give the year when he moved into the present apartment.

**Residence in 1965.**—Information on residence in 1965 was obtained from both the 15-percent and the 5-percent samples, but the information from the 5-percent sample relates to State of residence only. Residence on April 1, 1965, is the usual place of residence 5 years before enumeration. The information was used in conjunction with residence in 1970 to determine the extent of residential mobility of the population. The category "same house" includes all persons 5 years old and over who did not move during the 5 years, as well as those who had moved but by 1970 had returned to their 1965 residence.

Persons who changed residence from 1965 to 1970 are classified into groups according to type of move. The category "different house in United States" includes persons who, on April 1, 1965, lived in the United States in a different house from the one they occupied on April 1, 1970, and for whom sufficient information concerning the 1965 residence was collected. This group (also referred to as "movers") is divided into three subgroups according to their 1965 residence: "different house, same county," "different county, same State," and "different State." The second and third groups comprise the population classified as "migrants." The category "abroad" includes those with residence in a foreign country or an outlying area of the United States in 1965, including Armed Forces stationed overseas. Persons 5 years old and over who indicated they had moved into their present residence after April 1, 1965, but for whom sufficiently complete and consistent information regarding residence on April 1, 1965, was not collected, are included in the group "moved, 1965 residence not reported."

**School enrollment.**—Persons were included as enrolled in school if they reported attending a "regular" school or college at any time between February 1, 1970, and the time of enumeration. Regular schooling is that which may advance a person toward an elementary school certificate or high school diploma, or a college, university, or professional degree. Schooling that was not obtained in a regular school and schooling from a tutor or through correspondence courses were counted only if the credits obtained were regarded as transferable to a school in the regular school system. Persons were included as enrolled in nursery school only if the school included instruction as an important and integral phase of its program.

Schooling which is generally regarded as *not* "regular" includes that given in nursery schools which simply provide custodial day care; in specialized vocational, trade, or business schools; in on-the-job training; and through correspondence courses.

*Level and year of school in which enrolled.*—The levels of school which may be separately identified in census statistics are nursery school, kindergarten, elementary school, high school, and college. Children in "Head Start" or similar programs are counted under "Nursery" or "kindergarten" as appropriate. Elementary school, as defined here, includes grades 1 to 8, and high school includes grades 9 to 12. If a person was attending a junior high school, the equivalent in terms of 8 years of elementary school and 4 years of high school was obtained. The term "college" includes junior or community colleges, regular 4-year colleges, and graduate or professional schools.

The term "mode for age" relates to relative progress in shcool, in terms of whether the grade in which a person is enrolled is below, at, or above the modal, or typical grade for a person that age. To avoid a bimodal effect which resulted when modes were based on age as of April 1, 1970, the data were adjusted to age as of October 1, 1969. The resulting modes are grade 1 of elementary school for age 6 and one grade higher for each additional year of age up to age 17.

*Public, parochial, or other private school.*—In general, a "public" school is defined as any school which is controlled and supported primarily by a local, State, or Federal government agency. A "parochial" school is defined here as a private school which is controlled or supported primarily by a religious organization. An "other private" school is defined as a school controlled or supported primarily by private groups other than religious organizations.

**Years of school completed.**—The information on years of school completed

was based on responses to two questions. The first called for the highest grade attended, regardless of "skipped" or "repeated" grades. Persons whose highest grade of attendance was in a foreign school system, or in an ungraded school whose highest level of schooling was measured by "readers," or whose training was received through a tutor were instructed to report the approximate equivalent grade in the regular United States school system. The second question on educational attainment asked whether or not the highest grade attended had been finished.

The number in each category of highest grade of school completed represents the combination of (a) persons who reported that they had attended the indicated grade and had finished it, (b) those who had attended the next higher grade but had not finished it, and (c) those still attending the next higher grade. Persons who have not completed the first year of elementary school are classified as having no years of school completed.

**Vocational training.**—Included as "vocational training" are formal vocational training programs completed in high school, through an apprenticeship program, in a school of business, in a nursing school or trade school, in a technical institute, in the Armed Forces, or in Job Corps Training. Excluded from "vocational training" programs are single courses which were not part of an organized program of study, on-the-job training, training in company schools, training by correspondence, and basic training in the Armed Forces. Persons who reported having completed a vocational training program were asked to designate their main field of vocational training.

**Veteran status.**—As defined in this report, a "veteran" is a male 16 years old or over who has served but is not now serving in the Armed Forces of the United States. All other civilian males 16 years old and over are classified as nonveterans. In the classification by period of service, veterans with more than one period of service are generally classified according to the most recent wartime period; however, veterans with both Korean and World War II service are recognized in a separate category.

**Marital status.**—The marital status classification refers to the status at the time of enumeration. Persons classified as "married" consist of those who have been married only once and those who remarried after having been widowed or divorced. Persons reported as separated (either legally separated or otherwise absent from their spouse because of marital discord) are classified as a subcategory of married persons. Persons in common-law marriages are classified as married, and persons whose only marriage was annulled are classified as never married. Persons classified as never married are shown as "single."

Married persons with "spouse present" are men or women whose wife, or husband, was enumerated as a member of the same household even though he or she may have been temporarily absent on business or vacation, visiting, in a hospital, etc., at the time of enumeration. Married persons with "spouse absent" include both those who are separated because of marital discord and those whose usual place of residence is different from that of their spouse for other reasons, such as service in the Armed Forces or employment at a considerable distance from home.

**Marital history.**—For all persons in the 5-percent sample reported as married (including separated), widowed, or divorced at the time of the enumeration, data were obtained on the date of the first marriage. From this information and from current age, data on age at first marriage were derived. For each person who had been married more than once, information was obtained on the reason for termination of the first marriage. Persons shown as "known to have been widowed" include widowed persons and those currently married or divorced persons whose first marriage ended in widowhood. Persons shown as "known to have been divorced" include divorced persons and those currently married or widowed persons whose first marriage did not end in widowhood.

**Household.**—A household includes all persons who occupy a group of rooms or a single room which constitutes a housing unit. A group of rooms or single room is regarded as a housing unit when it is occupied as separate living quarters; that is, when the occupants do not live and eat with any other persons in the structure, and when there is either (1) direct access from the outside of the building or through a common hall or (2) complete kitchen facilities for the exclusive use of the occupants of the household.

**Relationship to head of household.**—Several categories of relationship to head of household are recognized in census reports.

*Head of household.*— One person in each household is designated as the "head," that is, the person who is regarded as the head by the members of the household. However, if

a married woman living with her husband was reported as the head, her husband was considered the head for the purpose of simplifying the tabulations.

Two types of household head are distinguished—the head of a family and a primary individual. A family head is a household head living with one or more persons related to him by blood, marriage, or adoption. A primary individual is a household head living alone or with nonrelatives only.

*Wife of head.*—A woman married to and living with a household head, including women in common-law marriages as well as women in formal marriages.

*Child of head.*—A son, daughter, stepchild, or adopted child of the head of the household of which he is a member, regardless of the child's age or marital status. The category excludes sons-in-law and daughters-in-law. (See definition of "own child" below.)

*Other relative of head.*—Any person related to the household head by blood, marriage, or adoption, who is not included in one of the specific relationship categories shown in the particular table.

*Not related to head.*—All persons in the household not related to the head by blood, marriage, or adoption. Roomers, boarders, lodgers, partners, resident employees, wards, and foster children are included in this category.

**Group quarters.**—All persons not living in households are classified by the Bureau of the Census as living in group quarters. Group quarters are located most frequently in institutions, rooming houses, military barracks, college dormitories, fraternity and sorority houses, hospitals, monasteries, convents, and ships. A house or apartment is classified as group quarters if there are five or more persons unrelated to the head; or, when no head is designated, if six or more unrelated persons share the unit. Certain types of living quarters, e.g., college dormitories and military barracks, are classified as group quarters regardless of the number of persons in the unit. Two general categories of persons in group quarters are recognized: inmates of institutions and all other persons in group quarters.

**Inmate of institution.**—Persons under care or custody in institutions at the time of enumeration are classified as "inmate of institution" regardless of their length of stay in that place and regardless of the number of people in that place. Institutions are a subcategory of group quarters and include homes, schools, hospitals, or wards for juveniles, for the physically handicapped, or for the mentally handicapped; hospitals for mental, tubercular, or chronic disease patients; homes for unwed mothers; nursing, convalescent and rest homes; homes for the aged and dependent; and correctional institutions.

**Married couple.**—A married couple is defined as a husband and his wife enumerated as members of the same household. The number of married couples, as shown in this report, is identical to the number of married men with wife present.

A "married couple with own household" is a married couple in which the husband is a household head; the number of such married couples is the same as the number of "husband-wife families." Married couples without own household are, by census definition, the same as married couples in subfamilies.

**Family and subfamily.**—According to 1970 census definitions, a family consists of a household head and one or more other persons living in the same household who are related to the head by blood, marriage, or adoption; all persons in a household who are related to the head are regarded as family members. A "husband-wife family" is a family in which the head and his wife are enumerated as members of the same household. Not all households contain families, because the household head may be living alone or with nonrelatives only. The mean size of family is derived by dividing the number of persons in families by the total number of families.

A subfamily is a married couple with or without children, or one parent with one or more single children under 18 years old, living in a household and related to, but not including, the head of the household or his wife. The most common example of a subfamily is a young married couple sharing the home of the husband's or wife's parents. Members of a subfamily are also included among the members of a family. The number of subfamilies, therefore, is not included in the number of families.

**Own child and related child.**—Statistics on the presence of "own" children may be shown for married couples, families, subfamilies, or women ever married. An "own child" is a child under 18 years of age reported as a single (never married) son, daughter, stepchild, or adopted child of the

specified person(s). The term "sons/daughters" is used when the designated group includes children 18 years old and over; the term includes all sons and daughters regardless of marital status or age. The number of "children living with both parents" includes stepchildren and adopted children as well as sons and daughters born to the couple.

The term "related children" includes not only "own" children, as defined above, but also all other family members, regardless of marital status, who are under 18 years old (except the head or wife).

**Unrelated individual.**—An unrelated individual may be (1) a household head living alone or with nonrelatives only, (2) a household member who is not related to the head, or (3) a person living in group quarters who is not an inmate of an institution. Unrelated individuals who are household heads are called "primary individuals"; those who are not household heads are called "secondary individuals." In tables on income and poverty status, unrelated individuals are limited to those 14 years of age or older. In statistics on poverty status of unrelated individuals, college students living in dormitories and members of the Armed Forces living in barracks are excluded.

**Children ever born.**—In reporting statistics on children ever born, respondents were instructed to include children born to the woman before her present marriage, children no longer living, and children away from home, as well as children born to the woman who were still living in the home.

**Work disability.**—The information on disability was used to identify persons hindered because of a health or physical condition from working at a job. The question refers to a serious illness that has lasted (or is likely to last) for a relatively long time, or a serious physical or mental impairment, defect, or handicap. Also determined was whether such persons were able to work at all and how long each person had been limited in his working ability.

## ECONOMIC CHARACTERISTICS

**Employment status.**—According to current definition, the labor force and the various employment status categories relate to the population 16 years old and over. However, in this series of reports, data are sometimes shown for persons 14 and 15 years old as well.

*Reference week.*—Data on employment status refer to the calendar week prior to the date on which respondents completed their questionnaires or were interviewed by enumerators. Since the week of enumeration was not the same for all persons, the reference week for employment data is not entirely uniform.

*Employed.*—Employed persons comprise all civilians 16 years old and over who were either (a) "at work"—those who did any work at all as paid employees or in their own business or profession, or on their own farm, or who worked 15 hours or more as unpaid workers on a family farm or in a family business; or (b) were "with a job but not at work"—those who did not work during the reference week but had jobs or businesses from which they were temporarily absent due to illness, bad weather, industrial dispute, vacation, or other personal reasons. Excluded from the employed are persons whose only activity consisted of work around the house or volunteer work for religious, charitable, and similar organizations.

*Unemployed.*—Persons are classified as unemployed if they were civilians 16 years old and over and: (a) were neither "at work" nor "with a job, but not at work" during the reference week, (b) were looking for work during the past 4 weeks, and (c) were available to accept a job.

Also included as unemployed are persons who did not work at all during the reference week and were waiting to be called back to a job from which they had been laid off.

*Experienced unemployed.*—Unemployed persons who have worked at any time in the past are classified as the "experienced unemployed."

*Civilian labor force.*—The civilian labor force consists of persons classified as employed or unemployed in accordance with the criteria described above.

*Experienced civilian labor force.*—The "experienced civilian labor force" comprises the employed and experienced unemployed.

*Labor force.*—The labor force includes all persons classified in the civilian labor force plus members of the Armed Forces (persons on active duty with the United States Army, Air Force, Navy, Marine Corps, or Coast Guard).

*Not in labor force.*—All persons 16 years old and over who are not classified as members of the labor

force are defined as "not in labor force." This category consists mainly of students, housewives, retired workers, seasonal workers enumerated in an "off" season who were not looking for work, inmates of institutions, disabled persons, and persons doing only incidental unpaid family work (less than 15 hours during the reference week).

**Hours worked.**—The statistics on hours worked relate to persons "at work" during the reference week and represent the number of hours actually worked at all jobs during that week. They do not necessarily reflect the number of hours usually worked or the scheduled number of hours. The category "at work, full-time" includes persons who worked 35 hours or more during the reference week, and the category "at work, part-time" includes persons who worked 1 to 34 hours during the reference week.

**Weeks worked in 1969.**—Data on weeks worked pertain to the number of weeks during 1969 in which a person did any work for pay or profit (including paid vacation and sick leave) or worked without pay on a family farm or in a family business. Weeks of active service in the Armed Forces are also included.

**Year last worked.**—The "year last worked" pertains to the most recent year in which a person did any work for pay or profit, or worked without pay on a family farm or in a family business, or was on active duty in the Armed Forces.

**Place of work.**—Place of work refers to the geographic location at which civilians and Armed Forces personnel not on leave carried out their occupational or job activities during the calendar week prior to the date on which respondents completed their questionnaires or were interviewed by enumerators.

**Means of transportation to work.**—Means of transportation to work refers to the chief means (i.e., covering the greatest distance) of travel or type of conveyance used in traveling to and from work on the last day the respondent worked at the address given as his or her place of work.

**Occupation and industry.**—For employed persons the data on occupation and industry refer to the job held during the reference week. For persons employed at two or more jobs, the data refer to the job at which the person worked the greatest number of hours. For the unemployed with work experience and for persons not currently in the labor force who worked at some time after 1959, these data refer to the last job the person held. The occupation and industry statistics presented here are based on the detailed systems developed for the 1970 census: see 1970 Census of Population, **Classified Index of Industries and Occupations,** U.S. Government Printing Office, Washington, D.C., 1971.

The major occupation groups are sometimes combined into four divisions, as follows:

*White collar.*—Professional, technical, and kindred workers; managers and administrators, except farm; sales workers; and clerical and kindred workers.

*Blue collar.*—Craftsmen and kindred workers; operatives, except transport; transport equipment operatives; and laborers, except farm.

*Farm workers.*—Farmers and farm managers, farm laborers and farm foremen.

*Service workers.*—Service workers including private households.

The sequence in which these four divisions appear is not intended to imply that any division has a higher social or skill level than another.

**Class of worker.**—The information on class of worker refers to the same job as the information on occupation and industry. There are two major class-of-worker categories:

1. Wage and salary workers are persons who work for an employer for wages, salary, commission, tips, pay in kind, or pay at piece rates. This category may be subdivided into government or private. Government workers include employees of any Federal, State, or local governmental unit, regardless of the activity of the particular agency.

2. Self-employed workers are those who work for profit or fees in their own unincorporated business, profession, or trade, or who operate a farm. Also included in this group are the small number of "unpaid family workers," i.e., persons working without pay in a business operated by a person to whom they are related by blood, marriage, or adoption.

**Activity 5 years ago.**— A person is classified as a "worker in 1965" if he answered "yes" to either (a) working or (b) in the Armed Forces. All other persons are classified as nonworkers in 1965. The number of workers in 1965 may be understated, since there is a tendency for respondents to forget

intermittent or short periods of employment.

**Occupation 5 years ago.**— Classification by occupation 5 years ago was based on the **Classified Index of Industries and Occupations**, referred to above. There are indications that, because of failure of the respondent to recall accurately, the responses on occupation in 1965 are subject to greater error than those on current occupation. Therefore, the occupational classification in 1965 is restricted to the major occupation group.

**Income in 1969.**—Information on money income received in the calendar year 1969 was requested from all persons 14 years old and over. "Total income" is the algebraic sum of amounts received as wage or salary income, self-employment income, and income other than earnings. Wage or salary income is defined as the total money earnings received for work performed as an employee. It represents the amount received before deductions for personal income taxes, Social Security, bond purchases, union dues, etc. Self-employment income is defined as net money income (gross receipts minus operating expenses) from a business, farm, or professional enterprise in which the person was engaged on his own account. Self-employment income may be subdivided into farm and nonfarm self-employment income. Earnings is the algebraic sum of wage or salary income and self-employment income. Income other than earnings includes money income received from such sources as interest, dividends, net rental income, Social Security benefits, pensions, veterans' payments, unemployment insurance, and public assistance or other governmental payments, and periodic receipts from insurance policies or annuities. Not included as income are money received from the sale of property (unless the recipient was engaged in the business of selling such property), the value of income "in kind," withdrawals of bank deposits, money borrowed, tax refunds, and gifts and lump-sum inheritances or insurance payments.

Family income is computed by summing the income of all family members 14 years old and over; whereas, the income of unrelated individuals or persons 14 years old and over includes only their own income.

The income data presented in census reports are not directly comparable with estimates of aggregate personal income or per capita personal income prepared by the Bureau of Economic Analysis (BEA), formerly the Office of Business Economics, nor with the distribution of families and individuals by personal income brackets previously published by that agency. For a detailed explanation of the difference between the Census and BEA income definitions and coverage, see U.S. Bureau of the Census, **Current Population Reports**, Series P-60, No. 75, "Income in 1969 of Families and Persons in the United States" or the 1970 Census of Population Volume I reports.

**Poverty status in 1969.**—Families and unrelated individuals are classified as being above or below the poverty level, using the poverty index adopted by a Federal Interagency Committee in 1969. This index provides a range of income cutoffs or "poverty thresholds" adjusted to take into account such factors as family size, sex and age of the family head, the number of children, and farm-nonfarm residence. The poverty cutoffs for farm families have been set at 85 percent of the nonfarm levels. These income cutoffs are updated every year to reflect the changes in the Consumer Price Index. The poverty threshold for a nonfarm family of four was $3,743 in 1969. For a detailed explanation of the poverty definition, see the Volume I reports of the 1970 Census of Population.

# Appendix D.—ACCURACY OF THE DATA

SOURCES OF ERROR ..... App-13
EDITING OF UNACCEPTABLE DATA ..... App-13
ALLOCATION TABLES ..... App-13
SAMPLE DESIGN ..... App-14
RATIO ESTIMATION ..... App-14
SAMPLING VARIABILITY ..... App-16

## SOURCES OF ERROR

Human and mechanical errors occur in any mass statistical operation such as a decennial census. Errors during the data collection phase can include failure to obtain required information from respondents, obtaining incorrect or inconsistent information, and recording information in the wrong place or incorrectly. Errors can also occur during the field review of the enumerator's work, the clerical handling of the questionnaires, and the various stages of the electronic processing of the material. Careful efforts are made in every census to keep the errors in each step at an acceptably low level. Quality control and check measures are utilized throughout the census operation. As was done for the 1950 and 1960 censuses, evaluative material on many aspects of the 1970 census will be published as soon as the appropriate data are accumulated and analyzed.

## EDITING OF UNACCEPTABLE DATA

The objective of the processing operation is to produce a set of statistics that describes the population as accurately and clearly as possible. To meet this objective, certain unacceptable entries were edited.

In the field, questionnaires were reviewed for completeness by a census clerk or enumerator, and a followup was made for missing information. The major review occurred in the central processing office, where the editing and coding operation provided an opportunity to correct obvious errors in the respondents' entries for those items which required manual processing.

For a few of the items, the respondents' entries were reviewed for reasonableness or consistency on the basis of other information on the questionnaire. As a rule, however, editing was performed by hand only when it could not be done effectively by machine.

Because of the limitations of computer capacity and other resources, a number of complicated editing steps were not introduced when the effect on the final data was considered to be small. Thus, there may be a small number of cases having unlikely combinations of characteristics.

As one of the first steps in mechanical editing, the configuration of marks on the questionnaire was scanned electronically to determine whether it contained information for a person or merely spurious marks. If the questionnaire contained entries for at least two of the basic characteristics (relationship, sex, race, age, marital status), or for at least two relevant sample characteristics, the inference was made that the marks represented a person. Names were not used as a criterion of the presence of a person because the electronic scanning was unable to distinguish between a name and any other entry in the name space.

If any characteristics for a person were missing, they were, in most cases, supplied by allocation. Allocations, or assignments of acceptable codes in place of unacceptable entries, were needed most often where an entry for a given item was lacking or where the information reported for a person on that item was inconsistent with other information for the person. As in earlier censuses, the general procedure for changing unacceptable entries was to assign an entry that was consistent with entries for other persons with similiar characteristics. Thus, a person who was reported as a 20-year-old son of the household head, but for whom marital status was not reported, was assigned the same marital status as that of the last son processed in the same age group. The assignment of acceptable codes in place of blanks or unacceptable entries, it is believed, enhances the usefulness of the data.

The editing process also includes another type of correction; namely, the assignment of a full set of characteristics for a person. When there was indication that a housing unit was occupied but the questionnaire contained no information for any person, a previously processed household was selected as a substitute and the full set of characteristics for each substitute person was duplicated.

Specific tolerances were established for the number of computer allocations and substitutions that would be permitted. If the number of corrections was beyond tolerance, the questionnaires in which the errors occurred were clerically reviewed. If it was found that the errors resulted from damaged questionnaires, from improper microfilming, from faulty reading by FOSDIC of undamaged questionnaires, or from other types of machine failure, the questionnaires were reprocessed.

## ALLOCATION TABLES

Some of the Volume II reports contain tables presenting allocation rates for the principal subject of the report and

showing the extent of allocation of that characteristic cross-classified by other selected characteristics. When such tables are included, they follow the numbered tables and are discussed in the introduction of the report. Additional allocation tables presented in Volume I, Part 1, **Characteristics of the Population, United States Summary** (tables C-1 to C-4) include rates for most of the subjects covered in the census. The information shown in these tables includes (1) the percent of the total population enumerated in the sample, (2) the allocation rate for each subject, and (3) the distributions for each subject before allocation. For a given subject the distribution before allocation may be compared with the corresponding statistics shown in the detailed tables to measure the net effect of allocation. Further explanation of the allocation tables is included in Appendix C of the Volume I reports.

## SAMPLE DESIGN

For persons living in housing units at the time of the 1970 census, the housing unit, including all its occupants, was the sampling unit; for persons in group quarters identified in advance of the census, the sampling unit was the person. In nonmail areas, the enumerator canvassed his assigned area and listed all housing units in an address register sequentially in the order in which he first visited the units whether or not he completed the interview. Every fifth line of the address register was designated as a sample line, and the housing units listed on these lines were included in the sample. Each enumerator was given a random line on which he was to start listing and the order of canvassing was indicated in advance, although the instructions allowed some latitude in the order of visiting addresses. In mail areas, the list of housing units was prepared prior to Census Day either by employing commercial mailing lists corrected through the cooperation of the post office or by listing the units in a process similar to that used in nonmail areas. As in other areas, every fifth housing unit of these lists was designated to be in the sample. In group quarters, all persons were listed and every fifth person was selected for the sample.

This 20-percent sample was subdivided into a 15-percent and a 5-percent sample by designating every fourth 20-percent sample unit as a member of the 5-percent sample. The remaining sample units became the 15-percent sample. Two types of sample questionnaires were used, one for the 5-percent and one for the 15-percent sample units. Some questions were included on both the 5-percent and 15-percent sample forms and therefore appear for a sample of 20 percent of the units in the census.

Other items appeared on either the 15-percent or the 5-percent questionnaires. An item collected on a 20-percent basis may have been tabulated on a 15- or 5-percent basis. The sample rate for tabulation is shown in the headnote of the table if the entire table is based on the same sample. For tables having figures based on different samples, the sample rates are explained under the heading "Sample size" in the introduction.

Although the sampling procedure did not automatically insure an exact 20-percent sample of persons or housing units in each locality, the sample design was unbiased if carried through according to instructions; generally for larger areas the deviation from 20 percent was found to be quite small. Biases may have arisen, however, when the enumerator failed to follow his listing and sampling instructions exactly. Quality control procedures were used throughout the census process, however, and where there was clear evidence that the sampling procedures were not properly followed, some enumerators' assignments were returned to the field for resampling. For the United States as a whole 19.4 percent of the population and 19.6 percent of the housing units were enumerated in the sample. The bases for these percentages included several classes of the population for which no attempt at sampling was made. These were the relatively small numbers of persons and housing units (in most States, less than one percent) added to the enumeration from the post-census post office check, the various supplemental forms, and the special check of vacant units. (If these classes are excluded from the bases, the respective proportions become 19.6 and 19.7 percent.) The ratio estimation procedure described below adjusts the sample data to reflect these classes of population and housing units.

## RATIO ESTIMATION

The statistics based on 1970 census sample data are estimates made through the use of ratio estimation procedures, which were applied separately for population and for housing data for each of the 5-, 15-, and 20-percent samples. The first step in carrying through the ratio estimates was to establish the areas within which separate ratios were to be prepared. These are referred to as "weighting areas." For the 15- and 20-percent samples, the weighting areas contained a minimum population size of 2,500. The weighting areas used for the

5-percent ratio estimate were larger areas, having a minimum population size of 25,000 and comprising combinations of the weighting areas used for the 15- and 20-percent samples. Weighting areas were established by a mechanical operation on the computer and were defined to conform, as nearly as possible, to areas for which tabulations are produced. Where these areas do not agree (primarily for smaller areas), there may be some differences between complete counts and sample estimates.

The ratio estimation process for population operated in three stages. The first stage employed 19 household-type groups (the first of which was empty by definition). The second stage used two groups, head of household and not head of household, and the third stage used 24 age-sex-race groups.

*Group*

STAGE I

*Male head with own children under 18*

1 — 1-person household
2 — 2-person household
3 — 3-person household
. .
. .
6 — 6-or-more-person household

*Male head without own children under 18*

7-12 — 1-person to 6-or-more-person households

*Female head*

13-18 — 1-person to 6-or-more-person households

19 — *Group quarters persons*

STAGE II

20 — *Head of household*

21 — *Not head of household (including persons in group quarters)*

STAGE III

*Male Negro*

22 — Age under 5 years
23 — 5-13
24 — 14-24
25 — 25-44
26 — 45-64
27 — 65 and older

*Male, not Negro*

28-33 — Same age groups as for Male Negro

*Female Negro*

34-39 — Same age groups as for Male Negro

*Female, not Negro*

40-45 — Same age groups as for Male Negro

At each stage, for each of the groups, the ratio of the complete count to the weighted sample count of the population in the group was computed and applied to the weight of each sample person in the group. This operation was performed for each of the 19 groups in the first stage, then for the two groups in the second stage and finally for the 24 groups in the third stage. As a rule, the weighted sample counts within each of the 24 groups in the third stage should agree with the complete counts for the weighting areas. Close, although not exact, consistency can be expected for the two groups in the second stage and the 19 groups in the first stage.

There are some exceptions to this general rule, however. As indicated above, there may be differences between the complete counts and sample estimates when the tabulation area is not made up of whole weighting areas. Furthermore, in order to increase the reliability, a separate ratio was not computed in a group whenever certain criteria pertaining to the complete count of persons and the magnitude of the weight were not met. For example, for the 20-percent sample the complete count of persons in a group had to exceed 85 persons and the ratio of the complete count to the unweighted sample count could not exceed 20. Where these criteria were not met, groups were combined in a specific order until the conditions were met. Where this occurred, consistency between the weighted sample and the complete counts would apply as indicated above for the combined group but not necessarily for each of the groups in the combination.

Each sample person was assigned an integral weight to avoid the complications involved in rounding in the final tables. If, for example, the final weight for a 20-percent group was 5.2, one-fifth of the persons in the group (selected at random) were assigned a weight of 6 and the remaining four-fifths a weight of 5.

The estimates realize some of the gains in sampling efficiency that would have resulted had the population been stratified into the groups before sampling. The net effect is a reduction in both the sampling error and possible bias of most statistics below what would be obtained by weighting the results of the sample by a uniform factor (e.g., by weighting the 20-percent sample results by a uniform factor of 5). The reduction in sampling error will be trivial for some items and substantial for others. A byproduct of this estimation procedure is that estimates for this sample are, in general, consistent with the complete count for the population groups used in the estimation procedure. A more com-

plete discussion of the technical aspects of these ratio estimates will be presented in a separate report.

## SAMPLING VARIABILITY

The estimates from the 20-, 15-, and 5-percent sample tabulations are subject to sampling variability. The standard errors of these estimates can be approximated by using the data in tables A through C. The chances are about 2 out of 3 that the difference (due to sampling variability) between the sample estimate and the figure that would have been obtained from a complete count of the population is less than the standard error. The chances are about 19 out of 20 that the difference is less than twice the standard error and about 99 out of 100 that it is less than 2½ times the standard error. The amount by which the estimated standard error must be multiplied to obtain other odds deemed more appropriate can be found in most statistical textbooks. The sampling errors may be obtained by using the factors shown in table C in conjunction with table A for absolute numbers and in conjunction with table B for percentages. These tables reflect the effect of simple response variance, but not of bias arising in the collection, processing, and estimation steps nor of the correlated errors enumerators introduce; estimates of the magnitude of some of these factors in the total error are being evaluated and will be published at a later date.

Table A shows approximate standard errors of estimated numbers for most statistics based on the 20-percent sample. In determining the figures for this table, some aspects of the sample design, the estimation process, and the population of the area over which the data have been compiled are ignored. Table B shows standard errors of most percentages based on the 20-percent sample. Linear interpolation in tables A and B will provide approximate results that are satisfactory for most purposes. Table C provides a factor by which the standard errors shown in tables A or B should be multiplied to adjust for the combined effect of the sample size (i.e., whether a 20-percent, 15-percent, or 5-percent sample), the sample design, and the estimation procedure.

To estimate the standard error for a given characteristic based on the 15- or 5-percent sample, or for a more precise estimate for the 20-percent sample, locate in table C the factor applying to the characteristic and sample size used to tabulate the data and multiply this factor by the standard error found in table A or B. If the estimate is not identified in table C, use the factor shown for "all other." Where data are shown as cross-classifications of two characteristics, locate and use the larger factor. Similarly, if an item, although collected on one sample basis, has been tabulated for a smaller sample, use the factor appropriate for the smaller sample.

The standard errors estimated from these tables are not directly applicable to differences between two sample estimates. In order to estimate the standard error of a difference, the tables are to be used somewhat differently in the three following situations:

1. For a difference between the sample figure and one based on a complete count (e.g., arising from comparisons between 1970 sample statistics and complete-count statistics for 1960 or 1950), the standard error is identical with the standard error of the 1970 estimate alone.

2. For a difference between two sample figures (that is, one from 1970 and the other from 1960, or both from the same census year), the standard error is approximately the square root of the sum of the squares of the standard errors of each estimate considered separately. This formula will represent the actual standard error quite accurately for the difference between estimates of the same characteristics in two different areas, or for the difference between separate and uncorrelated characteristics in the same area. If, however, there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error. The approximate standard error for the 1970 sample figure is derived directly from tables A through C. The standard error of a 25-percent 1960 sample figure may be obtained from the relevant 1960 census report or an approximate value may be obtained by multiplying the appropriate value in table A or B by 0.9.

3. For a difference between two sample estimates, one of which represents a subclass of the other, the tables can be used directly with the difference considered as the sample estimate.

The sampling variability of the medians presented in certain tables (median age, median years of school completed, and median income) depends on the size of the base and on the distribution on which the median is based. An approximate method for measuring the reliability of an estimated median is to determine an interval about the estimated median

such that there is a stated degree of confidence the true median lies within the interval. As the first step in estimating the upper and lower limits of the interval (that is, the confidence limits) about the median, compute one-half the number on which the median is based (designated $\frac{N}{2}$). From table A, following the method outlined in other parts of this section, compute the standard error of an estimated number equal to $\frac{N}{2}$. Subtract this standard error from $\frac{N}{2}$. Cumulate the frequencies (in the table on which the median is based) until the total first exceeds the difference between $\frac{N}{2}$ and its standard error and by linear interpolation obtain a value corresponding to this number. In a corresponding manner, add the standard error to $\frac{N}{2}$, cumulate the frequencies in the table, and obtain a value in the table on which the median is based corresponding to the sum of $\frac{N}{2}$ and its standard error. The chances are about 2 out of 3 that the median would lie between these two values. The range for 19 chances out of 20 and for 99 in 100 can be computed in a similar manner by multiplying the standard error by the appropriate factors before subtracting from and adding to one-half the number reporting the characteristic. Interpolation to obtain the values corresponding to these numbers gives the confidence limits for the median.

The sampling variability of a mean, such as the number of children ever born per 1,000 women or mean income, presented in certain reports, depends on the variability of the distribution on which the mean is based, the size of the sample, the sample design (for example, the use of households as the sampling unit), and the use of ratio estimates.

An approximation to the variability of the mean may be obtained as follows: compute the standard deviation of the distribution on which the mean is based; divide this figure by the square root of one-fifth of the total units in the distribution; multiply this quotient by the factor from table C appropriate to the statistic and the actual sample rate on which the mean is based. If the distribution is not published in the detailed tables, calculate the standard deviation from a comparable distribution for a larger area or for a similar population group; divide by the square root of one-fifth of the units on which the mean of interest is based; multiply the quotient by the factor from table C.

TABLE A. **Approximate Standard Error of Estimated Number Based on 20-Percent Sample**

(Range of 2 chances out of 3; for factors to be applied see table C and text)

| Estimated number[1] | Number of persons, families, or households in area[2] | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1,000 | 10,000 | 25,000 | 100,000 | 250,000 | 1,000,000 | 3,000,000 | 5,000,000 | 20,000,000 |
| 50 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 |
| 100 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| 250 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 |
| 500 | 30 | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 |
| 1,000 | ... | 60 | 60 | 65 | 65 | 65 | 65 | 65 | 65 |
| 2,500 | ... | 90 | 95 | 100 | 100 | 100 | 100 | 100 | 100 |
| 5,000 | ... | 100 | 130 | 140 | 140 | 140 | 140 | 140 | 140 |
| 10,000 | ... | ... | 150 | 190 | 200 | 200 | 200 | 200 | 200 |
| 15,000 | ... | ... | 150 | 230 | 240 | 240 | 240 | 240 | 240 |
| 25,000 | ... | ... | ... | 270 | 300 | 310 | 310 | 320 | 320 |
| 50,000 | ... | ... | ... | 320 | 400 | 440 | 440 | 440 | 450 |
| 75,000 | ... | ... | ... | 270 | 450 | 520 | 540 | 540 | 540 |
| 100,000 | ... | ... | ... | ... | 490 | 600 | 620 | 630 | 630 |

[1] For estimated numbers larger than 100,000, the relative errors are somewhat smaller than for 100,000.

[2] An area is the smallest complete geographic unit to which the estimate under consideration pertains. Thus, the area may be the State, city, county, standard metropolitan statistical area, urbanized area, or the urban or rural portion of the State or county. The rural farm or rural nonfarm population in the State or county, the Negro population, etc., do not represent complete areas. Total persons, families, or households for an area, if not available in this report, may be found in Volume I, **Characteristics of the Population.**

TABLE B. **Approximate Standard Error of Estimated Percentage Based on 20-Percent Sample**

(Range of 2 chances out of 3; for factors to be applied see table C and text)

| Estimated percentage | Base of percentage | | | | | | |
|---|---|---|---|---|---|---|---|
| | 500 | 1,000 | 2,500 | 10,000 | 25,000 | 100,000 | 250,000 |
| 2 or 98 | 1.3 | 0.9 | 0.6 | 0.3 | 0.2 | 0.1 | 0.1 |
| 5 or 95 | 2.0 | 1.4 | 0.9 | 0.4 | 0.3 | 0.1 | 0.1 |
| 10 or 90 | 2.7 | 1.9 | 1.2 | 0.6 | 0.4 | 0.2 | 0.1 |
| 25 or 75 | 3.9 | 2.7 | 1.7 | 0.9 | 0.5 | 0.3 | 0.2 |
| 50 | 4.5 | 3.2 | 2.0 | 1.0 | 0.6 | 0.3 | 0.2 |

TABLE C. **Factor to be Applied to Standard Error**

(For cross-classifications of two or more items, use the factor for the item having the largest factor for the given sample rate)

| SUBJECT | Factor if sample rate is— | | |
|---|---|---|---|
| | 20 percent | 15 percent | 5 percent |
| Race | 1.1 | 1.4 | 2.5 |
| Age | 0.8 | 1.0 | 1.8 |
| Household relationship | 0.5 | 0.6 | 1.1 |
| Families and subfamilies[1] | 0.6 | 0.7 | 1.3 |
| Unrelated individuals | 1.3 | 1.5 | 2.8 |
| Type of group quarters | 0.6 | 0.7 | 1.2 |
| Marital status | 0.6 | 0.7 | 1.4 |
| Marital history | – | – | 2.0 |
| State of birth | 1.3 | 1.6 | 2.9 |
| Country of origin | | | |
| Foreign born | – | 1.6 | 2.9 |
| Foreign stock | – | 1.6 | – |
| Spanish origin[2] | – | – | 2.9 |
| Spanish heritage[2] | – | 1.6 | – |
| Nativity and parentage | – | 1.7 | 3.1 |
| Mother tongue | – | 1.8 | – |
| Citizenship | – | – | 3.3 |
| Year of immigration | – | – | 3.3 |
| Year moved into present house | – | 1.9 | – |
| Residence in 1965 | – | 2.0 | 3.7 |
| Rural farm and nonfarm residence | 1.7 | 2.0 | 3.7 |
| School enrollment | – | 1.0 | 1.8 |
| Years of school completed | 1.0 | 1.2 | 2.3 |
| Vocational training | – | – | 1.7 |
| Veteran status | – | 0.9 | – |
| Disability | – | – | 2.4 |
| Labor force status or participation | 0.8 | 0.9 | 1.6 |
| Unemployed | 1.1 | 1.3 | 2.4 |
| Hours worked | 0.8 | 0.9 | 1.6 |
| Weeks worked in 1969 | 0.8 | 0.9 | 1.6 |
| Year last worked | 0.8 | 0.9 | 1.6 |
| Activity 5 years ago | 0.8 | 0.9 | 1.6 |
| Place of work | – | 1.3 | – |
| Means of transportation to work | – | 1.3 | – |
| Occupation | 1.1 | 1.3 | 2.3 |
| Industry | 1.1 | 1.3 | 2.3 |
| Class of worker | 1.1 | 1.3 | 2.3 |
| Income in 1969 | | | |
| Persons | 1.0 | 1.2 | 2.2 |
| Families[1] | 1.0 | 1.2 | 2.3 |
| Poverty status in 1969 | | | |
| Persons | 1.8 | 2.2 | 4.0 |
| Families[1] | 1.1 | 1.2 | 2.3 |
| All other | 1.0 | 1.2 | 2.2 |

[1] When determining the standard error of a number that relates to families or households rather than to persons, use the number of families or households in the area rather than the population of the area for selecting the appropriate column in table A.

[2] Tabulations of characteristics for persons of Spanish heritage (Puerto Rican birth or parentage, Spanish language or Spanish surname) are based on the 15-percent sample, and the appropriate factor is found in the 15-percent column for the subject or for Spanish heritage, whichever is the larger. Tabulations of characteristics for persons of Spanish origin are based on the 5-percent sample, and the appropriate factor is found in the 5-percent column for the subject or for Spanish origin whichever is the larger.

# Appendix E.—PUBLICATION AND COMPUTER SUMMARY TAPE PROGRAM

The results of the 1970 Census of Population and Housing are being issued in the form of printed reports, microfiche copies of the printed reports, computer summary tapes, computer printouts, and microfilm. Listed below are short descriptions of the final report series and computer tapes, as currently planned. More detailed information on this program can be obtained by writing to the Publications Distribution Section, Social and Economic Statistics Administration, Washington, D.C. 20233.

## Population Census Reports

**Volume I.**
**CHARACTERISTICS OF THE POPULATION**
This volume will consist of 58 "parts"—number 1 for the United States, numbers 2 through 52 for the 50 States and the District of Columbia in alphabetical order, and numbers 53 through 58 for Puerto Rico, Guam, Virgin Islands, American Samoa, Canal Zone, and Trust Territory of the Pacific Islands, respectively. Each part, which will be a separate clothbound book, will contain four chapters designated as A, B, C, and D. Each chapter (for each of the 58 areas) will first be issued as an individual paperbound report in four series designated as PC(1)-A, B, C, and D, respectively. The 58 PC(1)-A reports have been specially assembled and issued in a clothbound book, designated as Part A.

- Series PC(1)-A.
  **NUMBER OF INHABITANTS**

Final official population counts are presented for States, counties by urban and rural residence, standard metropolitan statistical areas (SMSA's), urbanized areas, county subdivisions, all incorporated places, and unincorporated places of 1,000 inhabitants or more.

- Series PC(1)-B.
  **GENERAL POPULATION CHARACTERISTICS**

Statistics on age, sex, race, marital status, and relationship to head of household are presented for States, counties by urban and rural residence, SMSA's, urbanized areas, county subdivisions, and places of 1,000 inhabitants or more.

- Series PC(1)-C.
  **GENERAL SOCIAL AND ECONOMIC CHARACTERISTICS**

Statistics are presented on nativity and parentage, State or country of birth, Spanish origin, mother tongue, residence 5 years ago, year moved into present house, school enrollment (public or private), years of school completed, vocational training, number of children ever born, family composition, disability, veteran status, employment status, place of work, means of transportation to work, occupation group, industry group, class of worker, and income (by type) in 1969 of families and individuals. Each subject is shown for some or all of the following areas: States, counties (by urban, rural-nonfarm, and rural-farm residence), SMSA's, urbanized areas, and places of 2,500 inhabitants or more.

- Series PC(1)-D.
  **DETAILED CHARACTERISTICS**

These reports cover most of the subjects shown in Series PC(1)-C, above, presenting the data in considerable detail and cross-classified by age, race, and other characteristics. Each subject is shown for some or all of the following areas: States (by urban, rural-nonfarm, and rural-farm residence), SMSA's, and large cities.

**Volume II.**
**SUBJECT REPORTS**
Each report in this volume, also designated as Series PC(2), will concentrate on a particular subject. Detailed information and cross-relationships will generally be provided on a national and regional level; in some reports, data for States or SMSA's will also be shown. Among the characteristics to be covered are national origin and race, fertility, families, marital status, migration, education, unemployment, occupation, industry, and income.

## Housing Census Reports

**Volume I.**
**HOUSING CHARACTERISTICS FOR STATES, CITIES, AND COUNTIES**
This volume consists of 58 "parts"—number 1 for the United States, numbers 2 through 52 for the 50 States and the District of Columbia in alphabetical order, and numbers 53 through 58 for Puerto Rico, Guam, Virgin Islands, American Samoa, Canal Zone, and Trust Territory of the Pacific Islands, respectively. Each part, which is a separate clothbound book, contains two chapters designated as A and B. Each chapter (for each of the 58 areas) is issued as an individual paperbound report in two series designated as HC(1)-A and B, respectively.

- Series HC(1)-A.
  **GENERAL HOUSING CHARACTERISTICS**

Statistics on tenure, kitchen facilities, plumbing facilities, number of rooms, persons per room, units in structure, mobile home, telephone, value, contract rent, and vacancy status are presented for some or all of the following areas: States (by urban and rural residence), SMSA's urbanized areas, places of 1,000 inhabitants or more, and counties.

- Series HC(1)-B.
  **DETAILED HOUSING CHARACTERISTICS**

Statistics are presented on a more detailed basis for the subjects included in the Series HC(1)-A reports, as well as on such additional subjects as year moved into unit, year structure built, basement, heating equipment, fuels, air conditioning, water and sewage, appliances, gross rent, and ownership of second home. Each subject is shown for some or all of the following areas: States (by urban, rural-nonfarm, and rural-farm residence), SMSA's, urbanized areas, places of 2,500 inhabitants or more, and counties (by rural and rural-farm residence).

**Volume II.**
**METROPOLITAN HOUSING CHARACTERISTICS**
These reports, also designated as Series HC(2), cover most of the 1970 census housing subjects in considerable detail and cross-classification. There is one report for each SMSA, presenting data for the SMSA and its central cities and places of 50,000 inhabitants or more, as well as a national summary report.

**Volume III.**
**BLOCK STATISTICS**
One report, under the designation Series HC(3), is issued for each urbanized area showing data for individual blocks on selected housing and population subjects. The series also includes reports for the communities outside urbanized areas which have contracted with the Census Bureau to provide block statistics from the 1970 census.

**Volume IV.**
**COMPONENTS OF INVENTORY CHANGE**
This volume will contain data on the disposition of the 1960 inventory and the source of the 1970 inventory, such as new construction, conversions, mergers, demolitions, and other additions and losses. Cross-tabulations of 1970 and 1960 characteristics for units that have not changed and characteristics of the present and previous residence of recent movers will also be provided. Statistics will be shown for 15 selected SMSA's and for the United States and regions.

**Volume V.**
**RESIDENTIAL FINANCE**
This volume will present data regarding the financing of privately owned nonfarm residential properties. Statistics will be shown on amount of outstanding mortgage debt, manner of acquisition of property, homeowner expenses, and other owner, property, and mortgage characteristics for the United States and regions.

**Volume VI.**
**PLUMBING FACILITIES AND ESTIMATES OF DILAPIDATED HOUSING**
This volume will present counts of housing units lacking some or all plumbing facilities and estimates of "dilapidated" units with all plumbing facilities. Comparative 1960 and 1970 data will be shown for the United States and regions, States, SMSA's, counties, and places of 10,000 inhabitants or more.

**Volume VII.**
**SUBJECT REPORTS**
Each report in this volume will concentrate on a particular subject. Detailed information and cross-classifications will generally be provided on a national and regional level; in some reports, data for States or SMSA's may also be shown. Among the subjects to be covered are housing characteristics by household composition, housing of minority groups and senior citizens, and households in mobile homes.

## Joint Population-Housing Reports

**Series PHC(1).**
**CENSUS TRACT REPORTS**
This series contains one report for each SMSA, showing data for most of the population and housing subjects included in the 1970 census.

**Series PHC(2).**
**GENERAL DEMOGRAPHIC TRENDS FOR METROPOLITAN AREAS, 1960 TO 1970**
This series consists of one report for each State and the District of Columbia, as well as a national summary report, presenting statistics for the State and for SMSA's and their central cities and constituent counties. Comparative 1960 and 1970 data are shown on population counts by age and race and on such housing subjects as tenure, plumbing facilities, value, and contract rent.

**Series PHC(3).**
**EMPLOYMENT PROFILES OF SELECTED LOW-INCOME AREAS**
This series consists of 76 reports, each presenting statistics on the social and economic characteristics of the residents of a particular low-income area. The data relate to low-income neighborhoods in 51 cities and seven rural poverty areas. Each report provides statistics on employment and unemployment, education, vocational training, availability for work, job history, and income, as well as on value or rent and number of rooms in the housing unit.

## Additional Reports

**Series PHC(E).**
**EVALUATION REPORTS**
This open series will present the results of the extensive evaluation program conducted as an integral part of the 1970 census program, and relating to such matters as completeness of enumeration and quality of the data on characteristics.

**Series PHC(R).**
**PROCEDURAL REPORTS**
This open series presents information on various administrative and methodological aspects of the 1970 census, and will include a comprehensive procedural history of the 1970 census. The first report issued focuses on the forms and procedures used in the data collection phase of the census.

## Computer Summary Tapes

The major portion of the results of the 1970 census are produced in a set of six tabulation counts. To help meet the needs of census users, these counts are designed to provide data with much greater subject and geographic detail than it is feasible or desirable to publish in printed reports. The data so tabulated are generally available—subject to suppression of certain detail where necessary to protect confidentiality—on magnetic computer tape, printouts, and microfilm, at the cost of preparing the copy.

**First Count**—source of the PC(1)-A reports; contains about 400 cells of data on the subjects covered in the PC(1)-B and HC(1)-A reports and tabulated for each of the approximately 250,000 enumeration districts in the United States.

**Second Count**—source of the PC(1)-B, HC(1)-A, and part of the PHC(1) reports; contains about 3,500 cells of data covering the subjects in these reports and tabulated for the approximately 35,000 tracts and 35,000 county subdivisions in the United States.

**Third Count**—source of the HC(3) reports; contains about 250 cells of data on the subjects covered in the PC(1)-B and HC(1)-A reports and tabulated for approximately 1,500,000 blocks in the United States.

**Fourth Count**—source of the PC(1)-C, HC(1)-B, and part of the PHC(1) reports; contains about 13,000 cells of data covering the subjects in these reports and tabulated for the approximately 35,000 tracts and 35,000 county subdivisions in the United States; also contains about 30,000 cells of data for each county.

**Fifth Count**—contains approximately 800 cells of population and housing data for 5-digit ZIP code areas in SMSA's and 3-digit ZIP code areas outside SMSA's; the ZIP code data are available only on tape.

**Sixth Count**—source of the PC(1)-D and HC(2) reports; contains about 260,000 cells of data covering the subjects in these reports and tabulated for States, SMSA's, and large cities.

The tapes are generally organized on a State basis. To use the First Count and Third Count tapes, it is necessary to purchase the appropriate enumeration district and block maps.

The term "cells" used herein to indicate the scope of subject content of the several counts refers to each figure or statistic in the tabulation for a specific geographic area. For example, in the Third Count, there are six cells for a cross-classification of race by sex: three categories of race (white, Negro, other race) by two categories of sex (male, female).

In addition to the above-mentioned summary tapes, the Census Bureau makes available for purchase certain sample tape files containing population and housing characteristics as shown on individual census records. These files contain no names or addresses, and the geographic identification is sufficiently broad to protect confidentiality. There are six files, each containing a 1-percent national sample of persons and housing units. Three of the files are drawn from the population covered by the census 15-percent sample and three from the population in the census 5-percent sample. Each of these three files provides a different type of geographic information: One identifies individual large SMSA's and, for the rest of the country, groups of counties; the second identifies individual States and, where they are sufficiently large, provides urban-rural and metropolitan-nonmetropolitan detail; and the third identifies State groups and size of place, with each individual record showing selected characteristics of the person's neighborhood.