

S E C O N D E D I T I O N

THE METHODS AND MATERIALS OF DEMOGRAPHY

Edited by
JACOB S. SIEGEL
DAVID A. SWANSON



ELSEVIER
ACADEMIC
PRESS

Amsterdam • Boston • Heidelberg • London • New York • Oxford
Paris • San Diego • San Francisco • Singapore • Sydney • Tokyo

Collection and Processing of Demographic Data

THOMAS BRYAN AND ROBERT HEUSER

This chapter deals with the collection and processing of demographic data. This topic is closely related to that of the preceding chapter, which treated the important kinds of demographic statistics and their availability. The discussion covers censuses and surveys and also registration systems for the collection of vital statistics. Practices differ considerably from country to country, and it would not be practicable to cover in this chapter all the important differences in data collection methods. Instead, this subject is discussed mainly in terms of the norms as countries with a long history of censuses or registration systems recognize them and as they are presented in publications of the United Nations and other international organizations.

POPULATION CENSUSES AND SURVEYS

Since many of the procedures and problems of data collection are common to censuses and surveys, these two data sources are treated together. Some distinctions between censuses and surveys were mentioned in Chapter 2. The United Nations (UN) states, "Population and housing censuses are a primary means of collecting basic population and housing statistics as part of an integrated program of data collection and compilation aimed at providing a comprehensive source of statistical information for economic and social development planning, for administrative purposes, for assessing conditions in human settlements, for research and for commercial and other uses" (United Nations, 1998, pp. 4-5).

Essential Features of a Population Census

The essential features of a population census, as stated in a recent United Nations publication, are "individual

enumeration, universality within a defined territory, simultaneity, and defined periodicity" (United Nations, 1998, p. 3).

Individual Enumeration

The principle to be observed here is to list persons individually along with their specified characteristics. However, in some earlier types of censuses, the "group enumeration" method is employed, whereby the number of adult males, adult females, and children is tallied within each group or family. This procedure was widely practiced in most of the enumerations of the African populations during the colonial era. The first few censuses of the United States represented a variation of such group enumeration methods. The main disadvantage of this method is that no greater detail on characteristics can be provided in the tabulations than that contained in the tally cells themselves. Tabulation becomes a process of mere summation. It is impossible to cross-classify characteristics unless they were tallied in cross-classification during the enumeration.

Universality Within a Defined Territory

Ideally, a national census should cover the country's entire territory and all people resident or present (depending on whether the basis of enumeration is *de jure* or *de facto*). When these ideals cannot be achieved for some reason (e.g., enemy occupation of part of the country in wartime or civil strife), then the type of coverage attempted and achieved should be fully described in the census publications.

Simultaneity

Ideally, a census is taken as of a given day. The canvass itself need not be completed on that day, particularly in the

case of a *de jure* census. Often, the official time is midnight of the census day. The more protracted the period of the canvass, however, the more difficult it becomes to avoid omissions and duplications. Some of the topics in a census may refer not to status on the census day but to status at a specified date or period in the past, such as residence 5 years ago, labor force status in the week preceding the census day, and income in the preceding calendar year.

Defined Periodicity

The United Nations recommends, "Censuses should be taken at regular intervals so that comparable information is made available in a fixed sequence. A series of censuses makes it possible to appraise the past, accurately describe the present and estimate the future" (United Nations, 1998, p. 3). If the censuses are spaced exactly 5 or 10 years apart, cohort analysis can be carried out more readily and the results can be presented in more conventional terms. However, some countries may find that they need to conduct a census at an irregular interval because of rapid changes in their population characteristics or major geographic changes. In the interests of international comparability, the United Nations suggests that population censuses be taken as closely as feasible to the years ending in "0."

Periodicity is obviously not an intrinsic requirement of a census but sponsorship by a national government should be seen as such a requirement. The United Nations also emphasizes the importance of sponsorship of the census by the national government (United Nations, 1998, p. 4). A national census is conducted by the national government, perhaps with the active cooperation of state or provincial governments. While it is feasible to have a national sample survey conducted by a private survey organization or to have a small-scale census (for a limited area) conducted by a city government, university department, training center, or some other entity, only national governments have the resources to support the vast organization and large expenditures of a full-scale census.

Census Strategic Objectives

The development or substantial improvement of a census involves a considerable amount of work. The task should be undertaken with the goal of fulfilling specific strategic objectives. These objectives should include, but are not limited to, census content and cost-effectiveness, census impact on the public, and the production of results.

The content of the census should be examined to ensure that it meets the demonstrated requirements of the users, particularly national government agencies, within the constraints of a budget. While the "requirements" of users may be endless, they must be assigned priorities so that the legally mandated and most important data are gathered

before less essential data are sought. Not only must data priorities be established, but efficiencies and economies of scale in collecting, organizing, and disseminating results must be established as well. The impact on the public of conducting a census can be measured by the burden it creates, its compliance with legal and ethical standards, and its ability to protect confidentiality. Obviously, the impact can vary widely, but in most cases the results of the census are used for distribution of political representation and of public funds and as the backbone of a national data system. The aim of producing census results must be to deliver mandated products and services that meet established standards of quality and are released according to a reasonable timetable. This includes producing standardized outputs with a minimum of error for widely recognized and agreed-upon geographic areas (United Nations, 1998, p. 4).

Advantages and Uses of Sample Surveys

As vehicles for the collection of demographic data, sample surveys have certain advantages and disadvantages, and their purposes and applications differ somewhat from those of censuses. Generally, surveys are not nearly as large and expensive, nor do they have the legal mandates and implications of censuses. Yates (1981, 321) wrote, "surveys fall into two main classes: those which have as their object the assessment of the characteristics of the population or different parts of it and those that are investigational in character." In the census type of survey, estimates of the characteristics, quantitative and qualitative, of the whole population and usually also of various previously defined subdivisions of it are required. In the investigational type of survey, we are more concerned with the study of relationships between different variates.

Since surveys of either type rarely have the regimented, standardized requirements of censuses, one resulting advantage is the possibility of experimenting with new questions. The fact that a new question is not altogether successful is less critical in the case of a sample survey than in that of a census, where the investment is much larger and where failure cannot be remedied until after the lapse of 5 or 10 years. In a continuing survey, new features can be introduced not only in the questions proper but also in the instructions to the canvassers, the coding, the editing, and the tabulations. Since a national population census is a multi-purpose statistical project, a fairly large number of different topics must be investigated, and no one of them can be explored in any great depth. In a survey, even when there is a nucleus of items that have to be included on the form every time, it is feasible in supplements, or occasional rounds, to probe a particular topic with a "battery" of related questions at relatively moderate additional cost.

In some instances, the data from a regular survey program may be superior in some respects to those from a

census. The field staff for surveys is often retained from month to month or year to year. The smaller size of the survey operation makes it possible to do the work with a smaller, select staff and to maintain closer surveillance and control of procedures.

The shorter time interval between surveys makes them more suitable for studying those population characteristics that change frequently in some countries, such as household formation, fertility, and employment status. With observations taken more frequently, it is much more feasible to analyze trends over time in the statistics. The analyst can delineate seasonal movements if the survey is conducted monthly or quarterly. Even when the survey data are available only annually, cyclical movements can be delineated more precisely than from censuses, and turning points in trends are more accurately located. The response of demographic phenomena to economic changes and to political events can also be studied more satisfactorily.

Among disadvantages of surveys, sampling error is the major one. This disadvantage is offset to some extent by the ability to compute the sampling error for estimates of various sizes and thus describe the limits of reliability. On the other hand, the magnitude of nonsampling error in surveys is oftentimes undetermined and the size of the survey samples is usually such that reliable statistics can be shown only in very limited geographic detail and for relatively broad cross-tabulations. For the latter reason, the census is the principal source of data for small areas and detailed cross-classifications of population characteristics. There is also usually some sampling bias arising from the design of the survey or from failure to carry out the design precisely. For example, it may not be practical to sample the entire population and coverage may not be extended to certain population subgroups, such as nomadic or tribal populations or persons living in group quarters. Moreover, the public may not cooperate as well in a sample survey as in a national census, which receives a great deal of publicity with attendant patriotic appeal.

The uses of censuses and surveys are sometimes interrelated. The use of the sample survey for testing new questions has already been mentioned. New procedures may also be tested. Census statistics may serve as benchmarks for analyzing and evaluating survey data and vice versa. The census can be used as a sampling frame for selecting the population to be included in a survey or may be a means of selecting a population group, such as persons in specified occupations.

CENSUS RECOMMENDATIONS

Methods of data collection vary among countries according to their cultural and technical advancement, the amount of data-collecting experience, and the resources available.

Both the methods used and the practices recommended by international agencies are covered in a number of sources. The Statistical Office of the United Nations has produced a considerable body of literature on the various aspects of the collection and processing of demographic statistics from censuses and surveys.

Definitions of Concepts

One requirement of a well-planned and executed census or survey is the development of a set of concepts and classes to be covered and adherence to these definitions throughout all stages of the collection and processing operations. These concepts provide the basis for the development of question wording, instructions for the enumerators, and specifications for editing, coding, and tabulating the data. Only when concepts are carefully defined in operational terms and consistently applied can there be a firm basis for later analysis of the results. Definitions of all of the recommended topics for national censuses and household surveys are presented in the manuals of the United Nations and are recognized by many countries as international standard definitions for the various population characteristics (United Nations, 1998).

Organization of National Statistical Offices

The statistical programs of a country may be largely centered in one national statistical office, which conducts the census and the major sample surveys, or they may be scattered among a number of government agencies, each with specific interests and responsibilities. Considerable differences exist among countries in the organization and permanence of the national census office, which may be an autonomous agency or part of the central statistical office. The United Nations groups countries into three categories according to types of central organizations: (1) those with a permanent census office and subsidiary offices in the provinces, (2) those with a permanent central office but no continuing organization of regional offices, so that they depend on provincial services or officials or field organizations of other national agencies, and (3) those that have no permanent census office but create an organization for the taking of each census and dissolve it when the census operations are complete.

There are many advantages to maintaining a permanent census office. Much of the work, including analysis of the data from the past census and plans and preparations for the next census, can best be accomplished by being spread throughout the intercensal period. The basic staff retained for this purpose forms a nucleus of experienced personnel to assume administrative, technical, and supervisory responsibilities when the organization is expanded for taking the census. The maintenance of this staff helps assure the

timeliness and maintenance of maps and technical documents necessary to conduct the census, as well as the security of historical census records.

Administration and Planning

The collection of demographic data by a census must have a legal basis, whereas a national sample survey may or may not have a legal foundation. The need for a legal basis is to establish administrative authority for the census. The administrative agency or organization is granted the authority to conduct a census and to use funds for this purpose within a specified time frame. The law must also provide for the conscription of the public to answer the census questions, and to do so truthfully. However, the legal basis that establishes the national program of census taking must also ensure the confidentiality of responses and ethical treatment of census respondents.

Any national census or major survey involves a vast amount of preparatory work, some aspects of which may begin years before the enumeration or survey date. Preliminary activities include geographic work, such as preparing maps and lists of places; determining the data needs of the national and local governments, business, labor, and the public; choosing the questions to be asked and the tabulations to be made; deciding on the method of enumeration; designing the questionnaire; testing the forms and procedures; planning the data-processing procedures; and acquiring the equipment to be used. Proper publicity for the census is important to the success of the enumeration, especially in countries where a census is being taken for the first time and the citizens may not understand its purpose. The public should also be assured of the confidentiality of the census returns—that is, that personal information will not be used for other than statistical purposes and will not be revealed in identifiable form by census officials.

Development of procedures for evaluation of the census should be part of the early planning to assure that they are included at the appropriate stages of the fieldwork and data processing and to assure that funds will be set aside for them. The funding of the census itself is one of many administrative responsibilities involved in the taking of a national census. Legislation must be passed to provide a legal basis, funds must be appropriated and a budget prepared, a time schedule of census operations must be set up, and a huge staff of census workers must be recruited and trained.

Quality Control

It is important from the outset of data collection to establish quality control measures for each step. Many of the processes for conducting and evaluating a census are similar to those of a large sample survey. Having quality control

measures at each step of the process is important in order to recognize and identify problems as they occur, enabling proper intervention measures. In countries with only recent experience in conducting a census, a quality control program is necessary to measure how census operations are proceeding. Even in countries with long-established censuses and large surveys, fluctuating numbers and the quality of workers, differences in data across multiple geographic layers, multiple types of data inputs and outputs over time, and technological advances require a solid quality control program to be in place.

Geography

In a national census, the geographic work has a twofold purpose: (1) to assure a complete and unduplicated count of the population of the country as a whole and of the many subdivisions for which data are to be published; (2) to delineate the enumeration areas to be assigned to individual enumerators.

To successfully carry out censuses and surveys, a formal ongoing cartographic program should be established. An ongoing operation not only affords a greater degree of comparability over time, but also saves the resources necessary to create such a program every time it is needed.

The boundaries that must be observed in a census include administrative, political, and statistical subdivisions (such as states or provinces and smaller political units). In countries that have a well-established census program, the geographic work is continuous and involves updating maps for changes in boundaries (e.g., annexations), redefining statistical areas, and so forth. When maps are not available from a previous census, they may be developed from existing maps obtained from various sources such as military organizations, school systems, ministries of health or interior, or highway departments, or they may be prepared from aerial photographs. The materials from these various sources may be compiled to produce working maps for the enumeration.

Once the maps have been prepared, the enumeration areas are delineated. There are two requirements for the establishment of enumeration areas. First, the enumeration area must not cross the boundaries of any tabulation area. Second, in the case of a direct-interview type of census, the population of the enumeration area as well as its physical dimensions must be such that one canvasser can complete the enumeration of the area in the time allotted. In some countries, the preparation of adequate maps is not feasible because of a lack of qualified personnel or because of the cost of producing the maps. In these cases, a complete listing of all inhabited places may be made by field workers as a substitute for maps.

The geographic work is sometimes supplemented with a precanvass of the enumeration areas shortly before

enumeration. A prec canvass serves to prepare the way for the enumeration by filling in any missing information on the map, providing publicity for the census, arranging with village chiefs or town officials for the enumerator's visit, determining the time necessary for covering the area, and planning the enumerator's itinerary. Geographic work is equally important as a preparatory phase of sample surveys. The selection of the sample usually depends on the delineation of certain geographical areas to serve as primary sampling units, then subdivisions of those areas, and finally delineation of small area segments of suitable size for the interviewer to cover in the allotted time period.

One of the most difficult tasks in conducting a census or survey is to identify and delineate small areas. Not only do small areas pose problems for data collectors but for data publication as well. The refinement of a geographic base is usually closely related to available resources. Each finer level of geographic detail usually entails an exponentially greater cost in conducting a census or survey. With limited resources, the best method is to establish a hierarchical coding of all geographic, political, and statistical subdivisions. The smallest of these may be limited by a minimum population, oftentimes established as 1000 or 2500. In a technically more advanced setting, if more resources are available, it is possible to coordinate cartographic operations with specific geographic identifiers. In such geocoding, each census or survey record may be identified on a coordinate or grid system, such as latitude and longitude. More information on geographic information systems and geocoding are available in Appendix D.

Once a geographic base is established, records of living quarters and housing-unit listings should be established and preferably associated with unique geographic, political, or statistical codes. This is particularly helpful in establishing enumeration districts, regardless of the type of areas for which the data are tabulated. Address lists, group quarters, government housing, shelters, and the like may be found in population registers and the records of tax authorities and other administrative agencies.

Census Instruments

Census questionnaires may be classified into three general types: first, the single individual questionnaire, which contains information for only one person; second, the single household questionnaire, which contains information for all the members of the household or housing unit; and third, the multihousehold questionnaire, which contains information for as many persons as can be entered on the form, including members of several households. Each of these has certain advantages and disadvantages.

The single individual questionnaire is more flexible for compiling information if the processing is to be done without the help of mechanical equipment. The single

household questionnaire has the advantage of being easy to manage in an enumeration and is especially convenient for obtaining a count of the number of households and for determining the relationship of each person to the household. If part of the census questions is to be confined to a sample of households, a single household schedule is required. The multihousehold questionnaire is more economical from the standpoint of printing costs and is convenient for processing on conventional or electronic tabulating equipment, but it may be awkward to handle because of its size.

Another type of questionnaire is that described earlier for group enumeration of nomadic people, when only the number of persons for broad age-sex groups is recorded. Although these summarized data do not provide census data in the strictest sense of the term, the group enumeration procedure has been used to enumerate classes of the population for whom conventional enumeration methods are not practical.

Census Content

The census subjects to be included are a balance between needs for the data and resources for carrying out the census program. National and local needs are of primary importance, but some consideration may also be given to achieving international comparability in the subjects chosen. As a rule, the list of subjects included in the previous census or censuses provides the starting point from which further planning of subjects proceeds. In general, it is desirable that most questions be retained from census to census in essentially the same form to provide a time series that can serve for analysis of the country's progress and needs. Some changes in subjects are necessary, however, to meet the changing needs of the country. Advice is usually sought from various national and local government agencies. Advisory groups including experts covering a wide range of interests may be organized and invited to participate in the formulation of the questionnaire content.

Census subjects may be classified as to whether they are mandated, required, or programmatic, as does the U.S. Census Bureau. Mandated subjects are those whose need for decennial census data is specifically cited in legislation. Required subjects are those that are specifically required by law and for which the census is the only source that has historically been used. Programmatic subjects are used for program planning, implementation, and evaluation and to provide legal evidence (U.S. Census Bureau, 1995).

Given this context, the United Nation's list of recommended items for censuses is valuable as an indicator of the basic items that have proved useful in many countries and as a guide to international comparability in subjects covered (United Nations, 1998, pp. 59-60). Its list of topics to be

included on the census questionnaire is as follows, with basic items shown in bold type:

1. Geographic and internal migration characteristics

Place of usual residence	Duration of residence
Place where found at time of census	Total population (Derived)
Place of birth	Locality (Derived)
Place of residence at a specified time in the past	Urban and rural (Derived)
Place of previous residence	

2. Household and family characteristics

Relationship to head or other reference person	Household and family composition (Derived)
Member of household	Household and family status (Derived)

3. Demographic and family characteristics

Sex	Religion
Age	Language
Marital status	National and/or ethnic groups
Citizenship	

4. Fertility and mortality

Children ever born	Maternal or paternal orphanhood
Children living	Age, date, or duration of first marriage
Date of birth of last child born alive	Age of mother at birth of first child born alive
Deaths in the past 12 months	

5. Educational characteristics

Literacy	Educational attainment
School attendance	Field of education and educational qualification

6. Economic characteristics

Activity status	Status in employment
Time worked	Income
Occupation	Institutional sector of employment
Industry	Place of work

7. International migration characteristics

Country of birth	Year or period of arrival
Citizenship	

8. Disability characteristics

Disability	Causes of disability
Impairment or handicap	

Regional interests are another consideration in the planning of census content. Organizations such as the Economic Commission for Europe, the Economic Commission for Asia and the Far East, the Economic Commission for Africa, ECLA, and the Inter-American Statistical Institute often conduct conferences with the United Nations to consider census content and methods and to make recommendations for the forthcoming census period. Neighboring countries sometimes cooperate in census planning through regional conferences or advisory groups for census subject matter

and practices. Public reaction to a subject also may influence the choice of census topics, since some questions may be too difficult or complicated for the respondent or the public may object to the substance of the question.

Survey Content

The contents of a survey are obviously significantly more guided by the objective and type of the survey than the standardization and continuity sought by a census. Although some sample surveys are multisubject surveys, it is more common for the survey to be restricted to one field, such as demographic characteristics or events, health, family income and expenditures, or labor force characteristics. One way in which sample surveys achieve multisubject scope is to vary the content from time to time. The UN *Handbook of Household Surveys* presents a list of recommended items for demographic surveys (United Nations, 1983). Content may also be determined by the type of survey being conducted, whether one-time (cross-sectional) or a series (longitudinal).

While the content of a census may be mandated, required, or programmatic, or combinations thereof, the requirements of specific survey questions are rarely well established and legal mandates for the content rarely exist. Therefore, consideration must not only be given to the value of each question in fulfilling the goal of the survey, but also the practicability of obtaining useful answers. Yates (1981, 58) wrote,

If the information is to be furnished in response to questions, the points of consideration are whether the respondents are sufficiently informed to be capable of giving accurate answers; whether, if the provision of accurate answers involves them in a good deal of work, such as consulting previous records, they will be prepared to undertake this work; whether they have motives for concealing the truth, and if so whether they will merely refuse to answer, or will give incorrect replies.

Tabulation Program

Closely related to the choice of subjects to be included in a census or survey is the planning of the tabulation program. Potential cross-tabulations in a census are boundless. Therefore, the selection of material is dictated partly by the uses of the results. The capacity of the financial and human resources and equipment for processing the data and the available facilities for publishing the results (e.g., page space available) place some restrictions on the material to be tabulated. The tabulation plans, as well as the choice of subjects on the questionnaire, should undergo review by the public, governmental, and commercial potential users of the statistics. Recommended tabulations for each of the subjects covered in national censuses and in various types of surveys are listed in the UN manuals previously listed.

Part of the planning of the tabulation program involves determining the number of different levels of geographic detail to be presented. Data are usually presented for the primary administrative divisions of the country and their principal subdivisions and for cities in various size categories as well as for the country as a whole. For the smallest geographic areas, such as small villages, the results as a rule are limited to a report of the total number of inhabitants or perhaps the male and female populations only. At the next higher level, which may be secondary administrative divisions, the tabulations may provide only "inventory statistics." These statistics are simply a count of persons in the categories of age, marital status, economic activity, and so forth, with little cross-classification with other characteristics. For the primary administrative divisions and major cities, most subjects are cross-tabulated by age and sex, and often there are also cross-classifications with other social and economic characteristics, such as educational attainment by economic activity or employment status by occupation. Also, more detailed categories may be shown on such subjects as country of birth, mother tongue, or occupation. The greatest degree of detail, sometimes termed "analytical" tabulations as opposed to "inventory" statistics, is that in which cross-tabulations involve detailed categories of each of the three or four characteristics involved.

Conducting the Census or Survey

Recruitment and Training

One of the largest tasks in conducting a survey, and especially a census, is the recruitment and training of staff. Anderson (1988, p. 201) states of the 1950 U.S. Census,

It was extraordinarily difficult to recruit in a number of months a reliable, competent staff of census enumerators and to guarantee uniform application of census procedures in the field. The 1950 evaluation studies indicated that on simple census questions, such as age and sex, the enumerators performed well. But in recording the answers to such complex questions as occupation and industry, two different interviewers recorded the answers differently in a sufficient number of cases to render the data suspect.

While retaining staff with the skills necessary for preparatory work (such as coding and data entry) is relatively easy, it is having a sufficient number of skilled workers conducting the enumeration that must be especially prepared for.

Pretesting

Pretesting of census content and methods has been found to be very useful in providing a basis for decisions that must be made during the advance planning of the census. This is especially so in countries without a long history of census

taking. Such pretests vary in scope. They may be limited to testing a few new subject items, alternate wording of a question, different types of questionnaires, or different enumeration procedures. Most census testing includes at least one full-scale pretest containing all questions to be asked on the census itself and sometimes covering part or all of the processing phases as well. The suitability of topics that have not been tried before may be determined from a small-scale survey in two or three localities. With enough other questions on the questionnaire to achieve something close to a normal census situation, a reasonable assessment of the question may be made. A test involving only the employees of the census office and their families may sometimes suffice for this purpose. Countries having an annual sample survey sometimes use this survey as a vehicle for testing prospective census questions.

Enumeration

The crucial phase of a census or survey comes when the questionnaires are taken into the field and the task of obtaining the required information begins. The kinds of problems encountered and the procedures used for collecting the data are similar for censuses and surveys. In a census the procedures for enumeration are affected by the type of population count to be obtained. The census may be designed to count persons where they are found on census day (a *de facto* count) or according to their usual residence (a *de jure* count).

In a *de facto* census, the method is to list all persons present in the household or other living quarters at midnight of the census day or all who passed the night there. In this type of enumeration, there is a problem of counting persons who happen to be traveling on census day or who work at night and consequently would not be found in any of the places where people usually live. It may be necessary to count persons on trains and boats or to ask households to include such members on the census form as well as those persons actually present. In some countries all persons are requested to stay in their homes on the census day or until a signal announces the completion of the enumeration.

In a *de jure* census, all persons who usually live in the household are listed on the form whether they are present or not. Visitors who have a usual residence elsewhere are excluded from the listing but are counted at their usual residence. Provisions must be made in a *de jure* census for persons away from home if those persons think it is likely that no one at their usual residence will report them. The usual practice is to enumerate such persons on a special form, which is forwarded to the census office of their home address. The form is checked against the returns for that area and is added to the count there if the person is not already listed. This is a complicated and expensive procedure, and

there still remains a chance that some persons will be missed and some counted twice.

There are two major types of enumeration, the direct-interview or canvasser method and the self-enumeration or householder method. In the direct-interview method, a census agent visits the household, lists the members living there, and asks the required questions for each person, usually by interviewing one member of the household. The advantage of this method is that the enumerator is a trained person who is familiar with the questions and their interpretation and he or she may assume a high degree of responsibility for the content of the census. Also, this method reduces the difficulty of obtaining information in an area where there is a low level of literacy. For these reasons it is considered possible to include more complex forms of questions in the direct-interview type of enumeration.

In self-enumeration, the census forms are distributed, usually one to each household, and one or more members of the household complete the form for all persons in the household. With this method of enumeration, there is less need for highly trained enumerators. The census enumerator may distribute the forms and later collect them, or the mail may be used for either the distribution or collection of the forms or for both. If enumerators collect the forms, they can review them for completeness and correctness and request additional information when necessary. In a mail census, the telephone may be used to collect information found to be lacking on the forms mailed in, or the enumerator may visit the household to obtain the missing information. In some cases the enumerator may complete an entire questionnaire if the household is unable to do so.

Self-enumeration has the advantage of giving the respondents more time to obtain the information and to consult records if necessary. People can supply the information about themselves, rather than having the information supplied by a household member who may not have complete or correct information. The possibility of bias resulting from a single enumerator's erroneously interpreting the questions is minimized in this method of enumeration. It is also more feasible to achieve simultaneity with self-enumeration because all respondents can be asked to complete the questionnaires as of the census day. Thus, in this respect, self-enumeration is the more suitable method if a *de facto* count is desired.

Self-enumeration is the more frequently used method in European countries, the United States, Australia, and New Zealand, whereas direct interview is the usual method in other countries. A combination of these two main types of enumeration is often used. The self-enumeration method may be considered appropriate for certain areas of the country and the interviewer method for others, or some of the information may be obtained by interview and the remainder by self-enumeration. In a census that uses the interviewer method as its basic procedure, self-enumeration

may be used for some individuals, such as roomers, when the head of the household cannot supply the information or when confidentiality is desired.

One of the goals of censuses and surveys is to minimize response burden. For years it has been possible to conduct surveys over the telephone, and more recently on the Internet. To make answering the census questionnaire easier and to ease respondent burden, many countries are exploring the possibility of allowing respondents to complete the basic demographic questions online over the World Wide Web, with Internet access to explanations about the questions asked in the census. Another innovation is telephone interviewing, whereby dedicated telephone lines are provided for the public to provide answers to the basic demographic questions, instead of their completing and mailing the census questionnaire.

Some special procedures for enumeration are required for certain groups of the population, such as nomads or people living in inaccessible areas (i.e., icy, mountainous, or forested areas). Levels of literacy may be low among certain social or geographically concentrated groups, who may have little understanding of the purpose of a census or interest in its objectives. A procedure sometimes followed is to request that all the members of such groups assemble in one place on a given day, since enumerating them at their usual place of residence might require from 4 to 5 months. For some of these, a method of group enumeration has been used. Rather than obtaining information for each individual or household, the enumerator obtains from the head of the group a count of the number of persons in various categories, such as marital status, sex, and age groups.

Enumeration of persons in hotels, *pensions*, missions, hospitals, and similar group quarters usually requires special procedures. Since some are transients, inquiry must be made to determine whether they have already been counted elsewhere. If a *de jure* count is being made, steps must be taken to assure that they are counted at their usual residence. Special individual census forms are usually used in group quarters, since the proprietor or other residents of the place could not provide the required information about each person. Another segment of the population that presents an enumeration problem is the homeless population, because people in this group have no fixed addresses and possibly occupy public spaces or temporary residences.

In some households the enumerator is unable to interview anyone even after repeated visits because no one is at home or, more rarely, because the occupants refuse to be enumerated. Since the primary purpose of a census is to obtain a count of the population, an effort is made to obtain information from neighbors about the number and sex of the household members. Neighbors may also be able to supply information about family relationships and marital status, which may, in turn, provide a basis for estimating age. Reliable information on other subjects usually cannot be

obtained except from the members themselves, and these questions are left blank, perhaps to be supplied during processing operations according to procedures that are discussed in "Processing Data."

In a sample survey, it is less practical to get information from neighbors because the emphasis is on characteristics rather than on a count of the population. The usual procedure is to base the results on the cases interviewed and adjust the basic weighting factors to allow for noninterview cases when the final estimates are derived from the sample returns. The effect of this procedure is to impute to the population not interviewed the same characteristics reported by the interviewed population. Since this assumption may not be very accurate, the presence of numerous noninterview households may bias the sample.

When a conventional enumeration has been completed in the field, questionnaires are assembled into bundles, usually corresponding to the area covered by one enumerator. The number of documents, the geographic identification of the area, and other appropriate information are recorded on a control form, which accompanies the set of documents throughout the various stages of processing. The tremendous volume of records involved in a census or large survey makes the receipt and control of material a very important function. The identification of the geographic area provides a basis for filing the documents and a means of locating a particular set of documents at any stage of the processing.

Processing Data

Regardless of the care expended on the preparation of a census and the enumeration of the population, the quality and the usefulness of the data will be compromised if they are not properly processed. The processing of the data includes all the steps, whether carried out by hand or by machine, that are required to produce from the information on the original document the final published reports on the number and characteristics of the population. The extent to which these operations are accomplished by mechanical or electronic equipment or by hand varies among countries and among surveys and censuses within countries.

Recent innovations in data processing have advanced processing capabilities immensely. However, few censuses are processed entirely electronically. Usually, some of the data, such as preliminary counts of the population for geographic areas, are obtained from a hand count. Even data that are produced primarily by machine must undergo some manual processing to correct for omissions or inconsistencies on the questionnaire and to convert certain types of entries into appropriate input for the electronic equipment. Electronic output may undergo a certain amount of hand processing before it is ready for reproduction in a published report. Such factors as the cost and availability of

equipment, the availability of manpower, and the goals in terms of tabulations to be made, reports to be published, and time schedules to be met determine the degree to which electronic processing is used. The data-processing operations to be performed in a census or survey usually consist of the following basic steps: editing, coding, data capture, and tabulation.

Editing

There are two principal points at which data errors may arise. The first occurs when a respondent provides erroneous or conflicting information, or an enumerator misrecords given information. The other occurs when data are coded and entered for computer processing. In both instances, concise rules should be established to determine how these errors should be edited. Census or survey procedures often include some editing of the questionnaires in the field offices to correct inconsistencies and eliminate omissions. Errors in the information can then more easily be corrected by checking with the respondent, and systematic errors made by the enumerator can more easily be rectified. Whether the editing is done in the field office or is part of central office processing, elimination of omissions and inconsistencies is a necessary step preliminary to coding.

A "not reported" category is permitted in some classifications of the population, but it is desirable to minimize the number of such cases. Where information is lacking, a reasonable entry can often be supplied by examining other information on the questionnaire. For example, a reasonable assumption of the relationship of a person to the head of the household or the householder can be made by checking names, ages, and marital status; or an entry of "married" may be assigned for marital status of a person whose relationship entry is "wife." Other edits may be made by comparing data entries with noncensus information, such as administrative records. For example, in 1980 the Census Bureau asked, "How many living quarters are in the building in which you live?" During editing, clerks were required to compare answers with the census mailout count for addresses with 10 or fewer units. If the clerk found that more units were reported in a building than questionnaires mailed, an enumerator was sent to investigate (Choldin, 1994, p. 57).

In manual editing, the clerks are given detailed specifications for assigning characteristics. Nonresponse cases may be assigned to a modal category (e.g., persons with place of birth not reported may be classified as native), or they may be distributed according to a known distribution of the population based on an earlier census. Since much of the editing for blanks and inconsistencies is accomplished by applying uniform rules, the use of electronic equipment for performing this operation is now commonplace. Electronic processing is designed to reject or to correct a record with missing

or inconsistent data and assign a reasonable response on the basis of other information.

Problems with data entry and coding can lead to voluminous errors in raw data files, making testing and quality control procedures throughout the census especially important. Errors of this type are typically systematic and can lead to much more pervasive problems than erroneous individual records. Strict editing and error-testing rules should be established by data experts and operationalized by programmers to ensure a minimum of problems.

Coding

Coding is the conversion of entries on the questionnaire into symbols that can be used as input to the tabulating equipment. Many of the responses on a census or survey require no coding or may be "precoded" by having the code for each written entry printed on the schedule. For those that do, there are three different types of coding techniques possible. For questions that have a small number of possible answers, such as sex or marital status, and questions that are answered in terms of a numerical entry, the appropriate code may be entered directly. If there are multiple answers, then computer-assisted coding may be used. In this process, codes are stored in a database and are automatically accessed and inserted at the prompting of the operator. The third alternative is automatic coding, which may be used if the coding scheme is extraordinarily complex—such as when the codes for an answer need to be recorded in more than one place.

Data Capture

In most data-processing systems, there must be some means of transferring the data from the original document to the tabulating equipment. After going through editing and coding, the data on the questionnaire may be transferred to a format that is electronically recognizable. There is a lengthy history of improvements in this field. In the 1880s, the U.S. Census Bureau sponsored the development of punched-card tabulation equipment. By 1946, the Census Bureau had contracted with the Eckert-Mauchley Computer Corporation to design a machine for processing the 1950 census, and the result of this collaboration was the UNIVAC. Special equipment developed for the 1960 census of the United States "reads" microfilmed copies of the questionnaires and transfers the data directly to computer tape. This equipment, known as FOSDIC (film optical sensing device for input to computers), reads the schedule by means of a moving beam of light, decides which codes have been marked, and records them on magnetic tape. By the 1980s, optical mark reading (OMR) was being widely used. Akin to a "scan-tron," OMR dramatically improved the speed and

accuracy with which data were captured. However, OMR limited the format on which survey and census responses could be printed.

Today, there are three techniques commonly used to capture data. The first is simple keyboard entry by clerks. At an average rate of between 5000 and 10,000 keystrokes per hour (depending on equipment and the skill of the clerk), manual entry is reserved for only the smallest data-capture tasks. The second is optical character recognition (OCR). OCR devices are programmed to look for characters in certain places on a census or survey response and convert them to an accurate, electronically recognizable value. The third is electronic optical scanning, which can be especially useful for recording handwritten answers and especially voluminous data. Recent developments in OCR and scanning have led to substantial improvements in accuracy through better character recognition, higher rates of input, and the acceptability of a wider range of paper and other media for input.

It was noted earlier that during the planning stage of a census or survey, decisions are made about the tabulations to be produced, and outlines are prepared showing how the data are to be classified and what cross-tabulations are to be made. The outlines may be quite specific, showing in detail the content of each proposed table.

On the basis of these outlines, specifications for computer programs are written for the various operations of sorting, adding, subtracting, counting, comparing, and other arithmetic procedures to be performed by the tabulating equipment. The input is usually punched cards or computer tape, and the output is the printed results in tabular arrangement. In the most advanced systems of tabulation, the final results include not only the absolute numbers in each of the prescribed categories but derived numbers such as percentage distributions, medians, means, and ratios as well.

One of the most obvious indicators of the quality of the data from a census or survey is the nonresponse rate. Even when a nonresponse category is not published and characteristics are allocated for those persons for whom information is lacking, a count of the nonresponse cases should be obtained during processing. One advantage of performing the edit in the computer is that not only the number of nonresponses on a given subject but also the known characteristics of the nonrespondents may be recorded. This provides a basis for analyzing nonresponses and judging the effects of the allocation procedures.

The nonresponse rate for a given item has more meaning if it is based on the population to which the question applies or to which analysis of that subject is limited. The base for nonresponse rates on date of first marriage, for example, would exclude the single population, and nonresponse rates for country of birth would be limited to the foreign born. A problem arises in the establishment of a population base

if the qualifying characteristic also contains a substantial number of nonresponses.

Planning the tabulations includes making some basic decisions about the treatment of nonresponses. Nonresponses may be represented in a separate category as "not reported" or they may be distributed among the specific categories according to some rule, ideally on the basis of other available characteristics of the person. Practices vary on the extent to which responses are allocated, but the elimination of "unknowns" before publication is a growing practice, partly because the greater capabilities of modern tabulating equipment have improved the possibilities of assigning a reasonable entry without prohibitive cost and partly because convenience to the user of the data favors the elimination of nonresponses.

Data Review

It has been mentioned that maintaining quality control and testing for errors while conducting a census or survey are imperative. Several steps may be taken to improve the accuracy and validity of results.

Supervisors should review samples of each enumerator's work for completeness and acceptability and accompany the enumerator on some of his or her visits. Progress-reporting of the enumeration enables census officials to know when an individual enumerator or the enumerators in a given area are falling seriously behind schedule and thus jeopardizing the completion of the census within the allotted time. Hand tallies of the population counted in each small area are compared with advance estimates, and the enumeration is reviewed if the results vary too widely from the expected number.

Reinterviewing is a common technique used for quality control of the data-collection process in sample surveys. A sample of households visited by the original interviewer is reinterviewed by the supervisor, and the results of the check-interview are compared with the original responses. Such checking determines whether the recorded interview actually took place and reveals any shortcomings of the interviewer.

Verification

Verification of the operation is an important element of each stage in the processing. Verification is not done for the purpose of removing all errors, as this is virtually impossible and does not justify the expense of time and resources. The purpose rather is (1) to detect systematic errors throughout the operation that can be remedied by changes in the instructions or by additional training of personnel, (2) to detect unsatisfactory performance on the part of an individual worker, and (3) to determine whether the general error

rate of the operation is within tolerance. Therefore, it is seldom necessary to have 100% verification. A procedure often followed is to verify an individual's work until the worker is found to be qualified in terms of a maximum allowable error rate, and thereafter to verify only a sample of the individual's work. If during the operation, a worker is found to have dropped below the acceptable level of accuracy, his or her work units may be subjected to a complete review and correction process.

Verification may be "dependent," in which the verifier reviews the work of the original clerk and determines whether it is correct, or "independent," in which two persons do the same work independently and then a comparison is made of the results. Tests have shown that in dependent verification, a large proportion of the errors are missed. Independent verification, in which the verifier is not influenced by what was done by the original worker, has been found to be more successful in discovering errors.

The statistical tables produced by the tabulating equipment are usually subjected to editorial and statistical review before being prepared for publication. On the basis of advance estimates and data from previous surveys or other independent sources, judgments are made regarding the reasonableness of the numbers. Figures that are radically different from the expected magnitudes may indicate an error in the specifications for tabulation. Review at this stage may show the need for expansion of the editing procedure. For example, early tabulations of educational statistics occasionally showing impossible combinations of age and educational attainment may lead to an addition to the editing specifications to eliminate spurious cases of this nature. Tables are reviewed for internal consistency. It is not necessary that corresponding figures in different tables agree perfectly to the last digit, since minor differences are common in tables produced by different passes through the tabulating equipment. Arbitrary corrections for all small differences are not feasible, and such changes would add little to the accuracy of the data. If the tables printed out by the tabulating equipment are to be used for publication, the spelling, punctuation, spacing, and indentation are also carefully reviewed so that corrections can be made before the tables are reproduced.

Evaluation

The evaluation of census results is frequently cited as a requirement of a good census. An initial distinction must be made between the *products* of an evaluation program and the *uses* of these products. The products of an evaluation are measures of census error and identification of the sources of error.

Census errors may occur at any of the various stages of enumeration and processing and may be either *coverage*

errors, that is, the omission or double-counting of persons, or *content* errors, that is, errors in the characteristics of the persons counted, resulting from incorrect reporting or recording or from failure to report. Methods for measuring the extent of error include reenumeration of a sample of the population covered in the census; comparison of census results with aggregate data from independent sources, usually administrative records; matching of census documents with other documents for the same person; and demographic analysis, which includes the comparison of statistics from successive censuses, analysis of the consistency of census statistics with estimates of population based on birth, death, and immigration statistics, and the analysis of census data for internal consistency and demographic reasonableness.

Uses of the results of census evaluation include guiding improvements in future censuses, assisting census users in interpreting results, and adjusting census results. Evaluation can identify certain geographic areas or persons with characteristics that made it problematic to enumerate them. The results of special enumeration efforts in relation to their costs may also be examined. Evaluation may also illustrate the usefulness and limitations of the census data, especially to novice users. It can alert the user to errors in the data and the magnitude of those errors. Moreover, the introduction of evaluation may inform users of additional sources of demographic data. Finally, evaluation may be used to adjust census results. Adjustment may be decided upon if evaluation indicates serious methodological, content, or coverage errors in the census (U.S. Census Bureau, 1985).

While there are a large number of methods for evaluating censuses, two predominant techniques have emerged. The first is the use of post-enumeration surveys, which employ case-by-case matching of the census and the survey to evaluate coverage and content error. The second is demographic analysis, which applies demographic techniques to data from administrative records to develop population estimates for comparison with the census.

Post-Enumeration Surveys

Post-enumeration surveys (PES) may be conducted in order to test census coverage and content error. While a PES may provide valuable insight into coverage and content error, caution must be used when designing and conducting a PES, as it is a statistically complex task. A simplified explanation of the method used by the U.S. Census Bureau in 1990 follows.

The Census Bureau's coverage measurement program in 1990 involving a post-enumeration survey was one in a series from 1950 to 2000. It was modeled after capture-recapture techniques used to estimate the size of animal populations. In essence, by sampling the population shortly after the census is taken and matching the two sets of data, estimates of census omissions may be derived. In the PES,

the traditional census enumeration corresponds to the original capture sample, and the PES to the recapture sample. However, equating the proportion of the PES sample not found in the census with the proportion of the census that was missed implicitly assumes that the chances of being counted in the capture sample and of being counted in the recapture sample are independent. It is known that the probability of being counted differs by age, sex, geographic area, and race, among other factors. For this reason, the results of the PES cannot be simply applied to the entire population, but instead must be stratified by small areas and various demographic and socioeconomic characteristics. In this way different coverage ratios are derived according to these factors.¹

Demographic Analysis

In addition to the information afforded by a PES, simple demographic techniques can be used to evaluate a census for accuracy and reasonableness. Visually identifying results that are statistically improbable can be considered demographic analysis. However, much more refined demographic techniques are available not only for detecting error, but for identifying its source as well. The goal of demographic analysis is to provide population estimates that are independent of the census being evaluated, using data from other sources, including principally administrative records on demographic variables such as births, deaths, and migration, and demographic techniques such as sex ratio and survival analysis (Kerr, 1998, p. 1).

Demographic analysis can be used in two contexts. The first is to evaluate the quality of the results themselves, and the other is to provide measures of error for possible adjustment of the census. Countries may use different types and even different combinations of methods of demographic analysis to evaluate census results. The results of this analysis may be used not only to estimate the overcoverage or undercoverage, but also to provide a basis for adjustment to the official census population statistics. In cases where demographic analysis shows results similar to those of the census, confidence in the census may be increased.

Different formal procedures of coverage evaluation may be used, and in fact some may be more appropriate in certain countries, based on their record-keeping systems. In Canada, for example, a combination of a reverse record check (RRC) and an overcoverage study are used for evaluating the census. The RRC is a comprehensive record-linkage system, which entails taking a sample from various administrative

¹ Further information on post-enumeration surveys may be found in William Bell, "Using Information from Demographic Analysis in Post-Enumeration Survey Estimation." *Statistical Research Report Series No. RR92/04*, Washington, DC: U.S. Census Bureau, Statistical Research Division, 1992.

records of people who should have been enumerated and surveying for those who were missed. The overcoverage study involves reenumerating a sample of enumerated households to test whether the members should have been enumerated and where they should have been enumerated (Kerr, 1998, pp. 3–4). In Australia, the National Demographic Data Bank, established in 1926 to measure births, deaths, and international migration, is used to develop estimates, which are used in conjunction with a PES to evaluate that country's census (Kerr, 1998, p. 20).

In the United States, the Census Bureau applies demographic analysis, distinguished as being a *macrolevel* approach to measuring coverage, and a Post-Enumeration Survey distinguished as being a *microlevel* approach. In the analytic method, estimates of the population below age 65 are derived from the basic demographic accounting equation, while Medicare data are used to estimate the population aged 65 and over. Some population groups, such as illegal entrants, have no associated administrative records and therefore must be estimated. While demographic analysis was not formally used to provide corrected populations in the 1990 U.S. census, it was used to measure net coverage error and "evaluate" the results of the PES (Robinson, 1996, p. 59).

The evaluation techniques of PES, RRC, overcoverage surveys, demographic analysis, and others are not without their shortcomings. The PES and RRC techniques are hindered by difficulty in measuring nonsampling error. Overcoverage is always difficult to measure, as in the case of *de jure* censuses, and the respondents often do not know that they have been recorded twice. The quality of demographic estimates declines in older age categories as the length of the times series for births used in estimation grows, and difficulty in measuring certain components (such as international migration) may compound error. Additionally, geographic detail is often lost, affording analysis only for large census regions or a nation as a whole.

The benefits of demographic analysis, however, are that it may be applied at a very low cost and that most of the administrative records necessary for demographic analysis oftentimes exist already and only need to be compiled and summarized for an evaluation. Demographic analysis is also easy to complete on a timely basis and works independently of the census, thus affording a quick and valid evaluation of census results. Finally, demographic analysis provides a benchmark of decennial census quality, affording the only consistent historical time series of measures of census net undercount for age, sex, and race groups (Robinson, 1996, pp. 60–61).

Dissemination

Once data are tabulated and reviewed, they are disseminated to users. Private, governmental, and other non-

commercial groups rely on timely and convenient access to census data. Historically, census data have primarily been provided as a series of printed tables and more recently as data tapes and CD-ROMs. Recent advances in Internet technology now afford data users the opportunity to gather data online and to design data sets and tabulations not previously possible.²

Publication of Results

The output of the tabulation equipment may be used as the final statistical tables suitable for reproduction in the published reports, or it may be an interim tabular arrangement of the data from which the final tables will be produced. In the latter situation, typing of the final tables is either done directly from the machine printouts or requires preliminary hand posting of the data on worksheets to arrange them as required for the publication tables. These additional steps, of course, require verification, proofreading, and machine-checking.

Electronic Dissemination

The continuous improvement of computers and high-speed printers has made the automatic production of final tables both feasible and economical. The elimination of one or more manual operations in the production process reduces the burden of quality control, improves the timeliness of publication, and reduces manpower requirements. The use of high-speed printer output demands very precise advance planning of the content of each table, the wording of captions and stubs, and the spacing of lines and columns. The technical skill involved and the lead time required for such planning have led some countries to use a compromise procedure in which the machine printout is used for the body of the table but the stubs and captions are provided by means of preprinted overlays. The programming of the computer printout in these instances is designed to display the data in the desired arrangement and to include rudimentary captions, which identify the numbers.

As discussed in Chapter 2, the trend in the dissemination of survey and census data has been heavily toward electronic dissemination on CD-ROM and other high-capacity media, and it is now turning toward the Internet. There are many potential methods for data dissemination on the Internet, ranging from free public access of easily downloadable

² A valuable source of international census enumeration, data tabulation, and dissemination is *Diffusion: International Forum for Census Dissemination*, 1985, Statistics Canada. Published approximately every year, editorship rotates among participating countries. The journal provides international perspectives on testing forms, designs, topics, and questions. The journal also provides evaluation of data tabulation and dissemination methods.

data files and products, to interactive online software for the creation of customized data sets by the user to commercial "for a fee" data available by subscription only. Data security on the Internet is an important consideration, not only for users, but for data suppliers as well. Commercial data vendors often contend with security issues, such as unauthorized users' accessing their files without permission. In addition to the emplacement of sophisticated security systems, techniques have been devised whereby encoded/encrypted data are placed on the Internet, and authorized users are privately given special software with which to access it.

Storage

In addition to these improvements in data dissemination, consideration must be given to the voluminous data in existence on other media. As already mentioned, many data have been stored on computer tape. Four alternate technological applications are used to replace traditional hard-copy records. These include microforms, computer-assisted microforms systems, optical disk systems, and computer-based systems (Suliman, 1996). It should be noted that these applications are used for a wide variety of data-storage purposes in addition to censuses and surveys, including civil registers, vital statistics, and population registers.

Microforms were one of the earliest replacements of hard-copy records and developed into both roll microfilm and flat microfiche. This application provides very long-term preservation of written information and often enhances written items on older records. An improvement of the microform system has been the computer-assisted microform system (CAM). If records already exist in a manual microform system, they can be indexed electronically, allowing very fast searches and record retrieval. If records do not already exist in a microform system, they may be filmed and placed directly into a CAM system. Shortcomings of both microform systems are the inability to evaluate the data statistically and to make any subsequent changes once the data have been filmed. The third application is known as an optical disk system. In this application, large volumes of records may be scanned electronically and stored on an optical disk. An electronic index may be created at the time of scanning, again allowing for very fast data searches and record retrieval. The optical disk system has the same limitations as microform, however, in that tabulations and calculations may not be made within the application, and revisions or corrections must be rescanned. The final system is the computer-based system. This has been described as the system in which data are entered directly via keystrokes or optical scanning systems that are compatible with software that enables conversion to an electronic format (Suliman, 1996).

Use of Sampling in Censuses

Although censuses as a rule involve a complete count of the number of inhabitants according to certain basic demographic characteristics, sampling is often used as an integral part of the enumeration to obtain additional information. As noted by the United Nations:

The rapidly growing needs in a number of countries for extensive and reliable demographic data have made sampling methods a very desirable adjunct of any complete census. Sampling is increasingly being used for broadening the scope of the census by asking a number of questions of only a sample of the population. Modern experience in the use of sampling techniques has confirmed that it is not necessary to gather all demographic information on a complete basis; the sampling approach makes it feasible to obtain required data of acceptable accuracy when factors of time and cost might make it impracticable, or other considerations make it unnecessary, to obtain the data on a complete count basis. (United Nations, 1998, p. 25)

Many data items may have to be collected on a complete-count basis because of legal requirements or because of the need for a high degree of precision in the data on basic topics so as to establish benchmarks for subsequent studies. However, the need in most countries for more extensive demographic data has driven the collection of other items on a sample basis. This practice not only expands the potential coverage of subjects, but also saves time and money throughout the enumeration and processing stages as well.

Even when data collection is on a 100% basis, a representative sample of the schedules may be selected for advance processing to permit early publication of basic information for the country as a whole and for large areas. Many of the final tabulations in a census may be limited to a sample of the population; thus the cost of tabulation is reduced considerably, especially when detailed cross-classifications are involved. In addition to its use in enumeration and processing, sampling is important in the testing of census questionnaires and methods prior to enumeration, in the application of quality-control procedures during enumeration and processing, and in the evaluation of the census by means of a PES and field checks (United Nations, 1998, p. 47).

Sample Survey Methods

The role of sample survey methods in the collection of demographic data is well established. Some of the uses and advantages of sample surveys were discussed earlier in this chapter. While a complete discussion of probability, survey design, and sampling concepts is not presented here, it is important to consider three aspects of sampling. The first is the definition of the population. It is important for analysts to consider the population to be measured and characterized and to take precautions to ensure that the sample instrument affords generalizability to that population. The second is the sampling methods being used. The choice among conven-

ience, typical-case, quota, or other designs in nonprobability sampling and among systematic, stratified, cluster, or other designs in probability sampling can have widely varying effects on the results of a survey. The third is the precision being sought. While the variance of sample estimates is inversely proportional to sample size, the cost, efficiency, and proposed uses of the data must also be considered (Henry, 1990).

When deriving census values based on sample census data, the sampling ratio itself determines the basic weights to be applied to each record (e.g., a sample of one in five leads to a weight of five). The figures produced by the application of these weights, however, are often subjected to other adjustments to obtain the final estimates. The adjustments may be made to account for the population not covered because of failure to obtain an interview. Also, independent population "controls" often are available to which the sample results are adjusted. In a census, the data obtained on a sample basis may be adjusted to the 100% population counts for the "marginal" totals by means of a ratio-estimation procedure. In this case the ratios of complete-count figures for specified demographic categories (e.g., age, sex, race) to the sample figures for the same categories are computed and used for adjusting the more detailed tabulations based on the sample. Similarly, the results of sample surveys may be adjusted to independent population controls, which are postcensal estimates derived by applying the basic population estimating equation to population figures from the previous census.

Other Demographic Record Systems

The administration of population registers differs somewhat from country to country, but basically it calls for registration at birth and entering specified subsequent events (marriage, change of residence, death, etc.) upon the individual or household record. A copy of this record, or an extract thereof, may be required to follow the person when she or he moves from one local jurisdiction to another. There are always local registers, and there may also be a central national register. The discussion of population registers in Chapter 2 gave an indication of their general nature and cited a number of publications concerning them.

Some aspects of the collection and processing of immigration data, particularly the registration system associated with border control, are discussed in Chapter 18. Here we consider, next, vital statistics registration systems in detail.

VITAL STATISTICS

Dual Functions of a Vital Statistics System

Vital statistics systems are designed primarily to accomplish the registration of vital events. Vital statistics, are the

statistics derived from compiling vital events. Registration of births, deaths, marriages, and divorces was originally intended to meet public and private needs for permanent legal records of these events, and these needs continue to be very important. However, equally important are the demands for useful statistics that have come from the fields of public health, life insurance, medical research, and population analysis.

Viewed as one of several general methods of collecting demographic statistics, registration has certain advantages and disadvantages. If events are registered near the time of occurrence, the completeness of reporting and the accuracy of the information are potentially greater than if reporting depends on a later contacted by an official and recall of the facts by the respondent. Also, continuous availability of the data file tends to be assured by the dual uses of the information—for legal and for statistical and public purposes.

There are also certain limitations of the registration method. The fact that the vital record is a legal document limits the amount and kind of nonlegal information that can be included in it. The method is also affected by the number and variety of persons involved in registering the events. For example, birth registration in some countries requires actions by thousands or millions of individual citizens and hundreds of local officials. Thousands of physicians, nurses, or hospital employees may be involved, and all of these people have other duties that they consider more urgent. It seems inevitable that for the most part these many and diverse persons will have less training and expertise in data collection than the enumerators who interview respondents in censuses or other population surveys. The latter are usually given intensive training in which the importance, purposes, and exact specifications of the information sought are thoroughly explained.

Satisfactory conduct of registration, in terms of both the legal and the statistical requirements, is closely related to the completeness and promptness with which events are registered and the accuracy of the information in the registration records. Certain functions such as indexing and filing of certificates, issuance of copies, and amendment of records are important for their legal uses but do not significantly affect the statistics. However, if the legal functions are poorly performed, the statistical program will suffer because public pressures will demand that first priority be given to serving people's needs for copies of their personal records.

International Standards and National Practices

The Handbook of Vital Statistics Systems and Methods, Volume I: Legal, Organizational and Technical Aspects (United Nations, 1991) and *Handbook of Vital Statistics Systems and Methods, Volume II: Review of National Prac-*

tice (United Nations, 1985), published by the United Nations Statistical Office, are the principal sources of the material presented in this section on international recommendations for the collection and processing of vital statistics.

Definitions of Vital Events

As in all systems of data collection, clear, precise definitions of the phenomena measured are prerequisites for accurate vital statistics. Use of standard definitions of vital events is essential for comparability of statistics for different countries.

Live Birth

Most countries follow the definition of a live birth recommended by the World Health Assembly in May 1950, and by the United Nations Statistical Commission in 1953, which is as follows:

Live birth is the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of pregnancy, which after such separation, breathes or shows any other evidence of life, such as beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles, whether or not the umbilical cord has been cut or the placenta is attached; each product of such birth is considered live-born. (United Nations, 1991, p. 17)

Under this definition a birth should be registered as a live birth regardless of its "viability" or death soon after birth or death before the required registration date. Although variations in the statistical treatment of "nonviable" live births (defined by low birthweight or short period of gestation) do not significantly affect the statistics of live births, they can have a substantial effect on fetal death and infant death statistics.

Death

Until very recently, there has been less difficulty with respect to the definition of death than with definitions of live birth and fetal death. For statistical purposes, the United Nations has recommended the following definition of death:

Death is the permanent disappearance of all evidence of life at any time after live birth has taken place (postnatal cessation of vital functions without capability of resuscitation). This definition therefore excludes foetal deaths. (United Nations, 1991, p. 17)

Fetal Death

The definition of fetal death recommended by the World Health Organization (WHO) and the United Nations Statistical Commission is as follows:

Foetal death is death prior to the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of pregnancy; the death is indicated by the fact that after such separation the foetus does not breathe or show any other evidence

of life, such as beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles. (United Nations, 1991, p. 17)

Marriage

The Statistical Commission of the United Nations has recommended the following definition of marriage for statistical purposes:

Marriage is the act, ceremony or process by which the legal relationship of husband and wife is constituted. The legality of the union may be established by civil, religious, or other means as recognized by the laws of each country. (United Nations, 1991, p. 17)

Divorce

The United Nations Statistical Commission's recommended definition of divorce is as follows:

Divorce is the final legal dissolution of a marriage, that is, the separation of husband and wife by a judicial decree which confers on the parties the right to civil and/or religious remarriage, according to the laws of each country. (United Nations, 1991, p. 17)

This definition excludes petitions, provisional divorces, and legal separations since they do not imply final dissolution of marriage and the right to remarry. In some countries, legal annulment is a statistically significant method of marriage termination. It is desirable in such countries to include annulments with divorces in determining the statistics of marriage dissolution. The *Handbook* defines annulment as "the invalidation or voiding of a marriage by a competent authority, according to the laws of each country, which confers on the parties the status of never having been married to each other (United Nations, 1991, p. 17).

Collection of Vital Statistics

Vital statistics systems differ in the amount of authority given to the collecting agency, the degree of national centralization of its organization, and the type of agency carrying out the program. The basic features of a vital statistics collection system are discussed in the following sections.

Civil Registration Method

This method of collecting vital statistics data is defined as the "continuous, permanent, compulsory recording of the occurrence and characteristics of vital events . . . in accordance with the legal requirements of each country" (United Nations, 1991, p. 16). The registration of all vital events must be done as they occur and must be maintained in order to be retrieved as required. This must be done by a permanent governmental agency with administrative stability. The underpinning, however, is that vital registration is legally required and there are penalties for failure to comply with the law. "The compulsion or legal obligation to register a vital event is the basic premise of the entire civil

registration system. When registration is voluntary rather than compulsory, there can be no assurance of complete or accurate vital records or statistics" (United Nations, 1973, p. 159). Without specific penalties, the fact that it is compulsory is meaningless.

Governmental Organization

The registration systems may be classified as organized under centralized or decentralized control. Most nations have established a centralized national authority over registration. In some countries, it is the civil registration office, in others, the department of public health, and in others, the central statistical agency. Again, in some countries the same national agency is responsible for both registration and vital statistics, but in others two or occasionally three separate agencies control these two functions. Advantages of a central registration office include direct and effective control over the entire system, including a standard legal framework, uniform procedures, and consistent interpretation and enforcement of norms and regulations.

In a decentralized system, civil registration is administered by major civil divisions, for example, the state, province, or department. Many countries with federated political systems have decentralized registration systems.

The Statistical Office of the United Nations Secretariat undertook a Survey of Vital Statistics Methods during 1976–1979. Of the 103 countries reporting on the type of civil registration system, 88 were centralized and 15 decentralized (United Nations, 1985, p. 8).

Local registration areas are the basic units of a vital registration system. They must have clearly defined geographic boundaries and be small enough for the registrar to provide good registration services for the area and for persons reporting vital events to come to or communicate with the registration office without excessive difficulty. One of the most important responsibilities of the local registrar is to encourage the general population, physicians, midwives, and others to report occurrences of vital events promptly and to supply complete and accurate information about them.

Informants and Reporters

The person responsible by law for reporting the occurrence of a vital event may or may not also be the source of the facts associated with the event. In most countries, a family member is responsible for reporting the occurrence of a live birth, fetal death, or death, together with certain personal information, but the attendant physician or midwife is also responsible for reporting the event along with certain medical information. The officiant, civil or religious, at the marriage is required to report it in about one-half of the countries; in the other half, the participants, bride and groom, are responsible. Reporting of divorces is the responsibility of the court in slightly more than half of the

countries and of one or both of the parties to the divorce in the remaining countries (United Nations, 1985, pp. 20–22).

Place of Registration

The United Nations recommends and, with few exceptions, the countries of the world require registration of vital events in the local registration area where the event occurred. Statistics tabulated by the United Nations from the 1976–1979 survey of national practices show that the percentage of responding countries where vital events are registered by place of occurrence is 92 for births and deaths, 93 for fetal deaths, 90 for marriages, and only 55 for divorces (United Nations, 1985, pp. 29–30). Tabulations are frequently made by area of usual residence of the mother, decedent, and so forth; these are generally regarded as more useful for demographic purposes than tabulations by place of occurrence.

Time Allowed for Current Registration

The registration record usually calls for both the date of the event and the date of registration. National laws usually specify the maximum interval permitted between these two dates for each type of vital event. The 1976–1979 survey shows that the time allowed for registering deaths tends to be shorter than for births—94% within 30 days for deaths compared with 73% for births (United Nations, 1985, pp. 26–27). The United Nations recommends that final tabulations for any calendar period should be based on events that occurred during that period and not on those registered. Information from the 1976–1979 survey indicates that two-thirds to three-quarters of the countries tabulated the records by date of registration (United Nations, 1985, pp. 34–35).

Content of Statistical Records

The need for national vital statistics data is the primary determinant of what items should be collected on vital records. Another major consideration is international comparability. The United Nations has recommended lists of statistical items that should be included in the records of live births, fetal deaths, deaths, marriages, and divorces (United Nations, 1991, pp. 30–31). The World Health Organization recommended the form of the medical certificate of cause of death. Some of the recommended items are designated as priority items, that is, items all countries should include. Parallel listings of priority items for the various vital statistics records are shown in Table 3.1.

Compilation and Tabulation of Vital Statistics

The underlying purpose of a vital statistics system is to make available useful statistics for the planning, administration, and evaluation of public health programs and to provide basic statistics for demographic research. The documents undergo much the same processing that is required

TABLE 3.1 Priority Items Recommended for Inclusion in Statistical Reports of Live Birth, Fetal Death, Death, Marriage, and Divorce

Live birth	Fetal death	Death	Marriage	Divorce
Date of occurrence	Date of occurrence	Date of occurrence	Date of occurrence	Date of occurrence
Date of registration	Date of registration	Date of registration	Date of registration	Date of registration
Place of occurrence	Place of occurrence	Place of occurrence	Place of occurrence	Place of occurrence
Place of usual residence of mother		Place of usual residence	Place of usual residence ¹	Place of usual residence ²
Sex	Sex	Sex		
Legitimacy status	Legitimacy status	Marital status	Marital status ¹	
Date of marriage (legitimate births)	Date of marriage (legitimate births)			Date of marriage
Age of mother	Age of mother	Age	Age ¹	Age ²
Type of birth (single or multiple)	Type of birth (single or multiple)		Type of ceremony (civil, religious, etc.)	
Number of children born to this mother	Number of children born to this mother Number of previous fetal deaths to this mother			Number of dependent children of divorcee ²
Weight at birth	Gestational age			
Attendant at birth		Cause Certifier		

¹Of bride and groom.

²Of both divorcees.

Source: United Nations, 1991. "Handbook of Vital Statistics Systems and Methods," Volume I: "Legal, Organizational and Technical aspects." *Studies in Methods*, Series F, No. 35, pp. 30-31.

for census and survey data, and similar planning is required to produce the desired tabulations.

In a majority of countries, the central statistical office has been given responsibility for compilation of national vital statistics. In some countries, including the United States, this function has been located in the national public health agency. In other countries, responsibility has been divided between the health agencies and the statistical and registration agencies.

The United Nations has suggested four criteria for measuring the effectiveness of a national vital statistics program, (1) coverage of the statistics, (2) accuracy of the statistics, (3) tabulations of sufficient detail to reveal important relationships, and (4) timeliness of availability of the data (United Nations, 1991, p. 46). One of the basic premises of a vital statistics system is that every event should be reported for statistical purposes for all geographic areas and all population subgroups. The time reference for the data should be the date on which the event occurred. The geographic reference for the statistics may be either the place where the event occurred or the residence of the person to whom the event occurred. Final tabulations for subnational geographic areas should be by place of residence. This allows for computation of meaningful population-based rates. Tabulation by place of occurrence may also be useful for specific administrative purposes. Finally, the data and their analysis

need to be disseminated to be useful. Unless the data are available to the public, its willingness to support the system cannot be expected. A wide variety of dissemination media should be used, including printed publications, public use data tapes and disks, and the Internet.

It is also essential that statistics of births, deaths, and marriages be based on definitions and classifications that are identical to or consistent with those used in the population census. Computation of valid vital rates and use of these rates in population estimation depend on consistent treatment of vital statistics and population data. This objective is sometimes difficult to attain, however, especially when different agencies are responsible for the two statistical programs.

Other Methods of Obtaining Vital Statistics

Every nation has as a goal the coverage of all its states or other areas in its vital statistics system. This objective is often not achieved without a long period during which the registration system is being developed and its coverage gradually extended. Other data collection methods may supplement or be a substitute for the registration system. These may include surveys, censuses, and population registers.

Surveys

Vital statistics may be obtained from a household sample survey by questioning members of the household regarding vital events that occurred in that household in some specific past period. This method can be implemented in a relatively short time if the necessary technical skills can be mobilized to plan and conduct the survey; and it can be expected to provide some statistics rather speedily. Its success depends heavily on the willingness of persons in the sample to supply the information and on their ability to recall the vital events occurring during some past period of time, and the date, place of occurrence, and other facts about the events. Also, the considerable skills required for sample design, survey organization and operation, and questionnaire construction need to be available on a continuing basis.

Censuses

Information on vital events is sometimes obtained in the population census. Statistics on births, marriages, and deaths in the previous year are available from this source in some countries. This method is essentially a special survey, which includes the entire population rather than a sample. It is subject to the same limitations as surveys with respect to the recall of events.

Population Registers

In countries that maintain a population register, birth, death, marriage, and divorce registration may be an integral part of the register. The information obtained in the registration of vital events must not only serve the needs for statistics on these subjects but must also be consistent in definitions and classifications with the information to be kept in the population register on the entire population.

The United States Vital Statistics System

National-State Relationships

The United States system for collecting vital records is decentralized in that the legal authority over registration is located in each of the 50 states and the District of Columbia. New York City is an independent registration area that has its own laws and regulations and publishes its own reports, as do Guam, Puerto Rico, and the Virgin Islands of the United States. Many states are divided into local registration districts, for each of which a registrar is appointed. There are about 10,000 such registrars, appointed by the state governments or locally elected. Each state separately processes the statistics that it wishes for its own area and population. The processing of national vital statistics is centralized in the National Center for Health Statistics (NCHS), a federal agency located in the U.S. Public Health Service (US PHS). An extensive history of the U.S. vital registration and statistics system may be found in

History and Organization of the Vital Statistics System (Hetzel, 1997).

Uniformity of Reporting

Although registration of vital events is governed by state laws, a considerable degree of uniformity has been achieved in definitions, organization, procedures, and forms. Uniformity has been promoted primarily by the development of model laws and certificate forms that have been recommended for state use. The Model State Vital Statistics Law has been followed with variations in the laws enacted in the various states. It was first promulgated in 1907 and has been revised and reissued several times. The most recent version was promulgated in 1992 (US PHS, 1995). Standard certificates of the several vital events, issued by the responsible national agency, have been the principal means of achieving uniformity in the certificates of the individual states, which provide the information upon which national vital statistics are based. The last revision was promulgated in 1989 (US NCHS/Tolson *et al.*, 1991). The next revision is being implemented gradually beginning in 2003.

The responsible national vital statistics agency (the Census Bureau, 1903–1946, NOVS, 1947–1959, and NCHS, 1960 to date) has actively assisted the state agencies in achieving complete, prompt, and accurate registration of vital events. Tests of registration completeness and intensive educational campaigns to promote registration have been joint federal-state efforts. The national office has developed and recommended to the states model handbooks designed to instruct physicians, hospitals, coroners and medical examiners, funeral directors, and marriage license clerks on current registration procedures and the meaning of the information requested in the certificates (e.g., US NCHS, 1987).

Functions Performed by State Offices

In the decentralized registration system of the United States, the primary responsibility for the collection of vital records rests with each state. This responsibility encompasses a number of functions that are carried out in each state's vital statistics office.

Planning Content of Forms

It is the responsibility of the state's vital statistics office to recommend the format and content of the vital records used in its jurisdiction. These recommendations are usually based to a large extent on the United States standard certificates but also often reflect special interests or needs not encompassed in the federal model forms. In spite of the efforts of the federal government to promote national uniformity, state and local uses of vital records, especially in the health field, produce differences in record content and format, which have an effect on the statistics. Some of the states have not included all of the standard

demographic or health items on their vital records. Currently, however, all states have birth and death certificates that conform very closely to the U.S. standard certificates in content.

Confidentiality of Records

It is the responsibility of each state or other registration area to determine the need for confidentiality and to maintain confidentiality of the vital records. In some areas, vital records are considered to be public documents; in other areas, the vital statistics laws and administrative regulations permit the release of information or certified copies of the record only to certain authorized persons.

Receipt and Processing of Records

One of the major functions of a state office is to serve as the repository for vital records of events occurring within the state, and thus to serve as a central source within each state for both the legal and statistical uses of the records. This function entails a number of related responsibilities, such as the handling of corrections, missing data, name changes, and adoptions and legitimations and issuing certified copies of records on file.

Electronic birth certificate (EBC) software has been developed for use in the capture of the information on the birth certificate at the reporting source (hospitals). This software has been designed to improve the timeliness and quality of birth registration. The information on the birth certificate is entered into the software by hospital personnel and transmitted to the appropriate registration authority within the state. Before transmission, it is checked for quality and completeness by an edit program designed and installed by the state. Currently all states are using EBC software and approximately 90% of births are currently registered through this process.

States are also in the process of developing Electronic death certificate (EDC) software. It is anticipated that within a few years most deaths will also be registered through an electronic process.

Tabulation and Publication of the Data

Just as each state prepares and processes its own vital statistics data, so does each state prepare an annual summary of its vital statistics. These summaries vary in analytic detail and comprehensiveness, but almost all states publish some kind of annual vital statistics report. Some of these reports merely present selected vital statistics data, whereas others contain, in addition to tabular material, an analysis and interpretation of the statistics.

Another activity of the state vital statistics offices is the transmittal of data to the National Center for Health Statistics (NCHS) for the purpose of assembling national statistics. The NCHS purchases the data in electronic form from each registration area through a contractual arrangement, which includes a guarantee of confidentiality

prohibiting the center from releasing any data other than statistical summaries without the written consent of the state's vital statistics office.

In order to issue provisional statistics in its *National Vital Statistics Report*, NCHS receives reports from the states on the total number of records (birth, death, infant death, marriage, and divorce) received during the month regardless of date of occurrence. Characteristics about these events are not published in these provisional reports.

Functions Performed by the National Center for Health Statistics

The NCHS performs a variety of functions designed to improve the national vital statistics system. It exercises leadership in the revision of the standard certificates and in evaluating the completeness of birth registration; represents the United States in international conferences on the standard classification of causes of death; conducts a training program on vital and health statistics; and helps the states in developing forms, procedures, draft legislation, definitions, and tabulations.

The NCHS serves as the focal point for the collection, analysis, and dissemination of national vital statistics for the United States. Because of the diversity of practices and procedures existing in the decentralized U.S. system, the production of national statistics involves more than the combination of statistics from each registration area to produce national vital statistics.

Detailed data on births, deaths, and fetal deaths are obtained in electronic form through contractual arrangements with the states. The data are subjected to a series of computer edits that eliminate inconsistencies in the data and impute missing data for certain items. This is generally done only when the number of items with missing data comprises a very small proportion of the total. Sex, race, and geographic classification are assigned if not reported on the birth or death certificates, and age and marital status of mother are assigned if not reported on the birth certificate.

The final computer tabulations of national vital statistics appear in various publications prepared by NCHS and mentioned in Chapter 2, "Basic Sources of Statistics." Unpublished material and resource data for special investigations are maintained by the NCHS and made available on the Internet (www.cdc.gov/nchs). In addition, unit record data on births, deaths, and linked birth-infant deaths are available on CD-ROMs.

References

- Anderson, M. 1988. *The American Census: A Social History*. New Haven, CT: Yale University Press.
- Choldin, H. 1994. *Looking for the Last Percent: The Controversy over Census Undercounts*. New Brunswick, NJ: Rutgers University Press.

- Henry, G. 1990. *Practical Sampling*. Newbury Park, CA: Sage.
- Hetzel, A. M., 1997. *History and Organization of the Vital Statistics System*. Hyattsville, MD: National Center for Health Statistics.
- Kerr, D. 1998. "A Review of Procedures for Estimating the Net Undercount of Censuses in Canada, the United States, Britain and Australia." *Demographic Documents*. Ottawa: Statistics Canada.
- Robinson, J. G. 1996. "What Is the Role of Demographic Analysis in the 2000 United States Census?" *Proceedings of Statistics Canada Symposium, 96: Nonsampling Errors, Nov. 1996*. pp. 57–63, Ottawa: Statistics Canada.
- Suliman, S. H. 1996. "Automation of Administrative Records and Statistics," <http://www.un.org/Depts/unsd/demotss/tenjun96/suliman.htm>, October 27, 1999.
- United Nations, 1973. "Principles and Recommendations for a Vital Statistics System." *Statistical Papers*, Series M, No. 19, Rev. 1. New York: United Nations.
- United Nations, 1983. "Handbook of Household Surveys." *Studies in Methods*, Series F, No. 10. New York: United Nations.
- United Nations, 1985. "Handbook of Vital Statistics Systems and Methods," Volume II: "Review of National Practices." *Studies in Methods*, Series F, No. 35. New York: United Nations.
- United Nations, 1991. "Handbook of Vital Statistics Systems and Methods," Volume I: "Legal, Organizational and Technical Aspects." *Studies in Methods*, Series F, No. 35. New York: United Nations.
- United Nations, 1998. "Principles and Recommendations for Population and Housing Censuses." *Statistical Papers*, Series M. No. 67 / Rev. 1. New York: United Nations.
- U.S. Census Bureau, 1985. "Evaluating Censuses of Population and Housing." *Special Training Document ISP-TR-5*. Washington, DC: U.S. Census Bureau.
- U.S. Census Bureau, 1995. "Solicitation of 2000 Census Content Needs from Non-federal Data Users: November 1994–March 1995." Special report of the Decennial Management Division. Washington, DC: U.S. Census Bureau.
- U.S. National Center for Health Statistics, 1987. *Hospitals' and Physicians' Handbook of Birth Registration and Fetal Death Reporting*. DHHS Pub. No. (PHS) 87–1107. Washington, DC: National Center for Health Statistics.
- U.S. National Center for Health Statistics, 1991. "The 1989 Revision of the U.S. Standard Certificates and Reports," by G. C. Tolson, J. M. Barnes, G. A. Gay, and J. L. Kowaleski. *Vital Health Stat 4(28)*. Hyattsville, MD: National Center for Health Statistics.
- U.S. Public Health Service, 1995. *Model State Vital Statistics Act and Regulations*. DHHS Pub. No. (PHS) 95–1115.
- Yates, Frank. 1981. *Sampling Methods for Censuses and Surveys*. New York: Oxford University Press.

Suggested Readings

- Anderson, M. 1988. *The American Census: A Social History*. New Haven, CT: Yale University Press.
- Edmonston, B., and C. Schultze (eds.). 1995. *Modernizing the U.S. Census*. Washington, DC: National Academy Press.
- Hetzel, A. M. 1997. *History and Organization of the Vital Statistics System*. Hyattsville, MD: National Center for Health Statistics.
- Hogan, H. 1993. "The 1990 Post-Enumeration Survey: Operations and Results." *Journal of the American Statistical Association* 88 (423), 1047–1060.
- Robinson, J. G., B. Ahmed, P. Das Gupta, and K. A. Woodrow. 1993. "Estimating the Population Coverage in the 1990 United States Census Based on Demographic Analysis." *Journal of the American Statistical Association* 88 (423), 1061–1071.
- United Nations, 1998. "Principles and Recommendations for Population and Housing Censuses." *Statistical Papers*, Series M. No. 67/Rev. 1. New York: United Nations.