

Studies in Population

The Methods and Materials of Demography

Henry S. Shryock
Jacob S. Siegel
and Associates

Condensed Edition by
Edward G. Stockwell

The Methods and Materials of Demography

Condensed Edition

This is a volume in
STUDIES IN POPULATION

A complete list of titles in this series is available from the Publisher upon request.

The Methods and Materials of Demography

**Henry S. Shryock
Jacob S. Siegel
and Associates**

Condensed Edition by
Edward G. Stockwell

Bowling Green University
Bowling Green, Ohio



ACADEMIC PRESS, INC.
Harcourt Brace Jovanovich, Publishers
San Diego New York Boston
London Sydney Tokyo Toronto

COPYRIGHT © 1976, BY ACADEMIC PRESS, INC.
ALL RIGHTS RESERVED.
NO PART OF THIS PUBLICATION MAY BE REPRODUCED OR
TRANSMITTED IN ANY FORM OR BY ANY MEANS, ELECTRONIC
OR MECHANICAL, INCLUDING PHOTOCOPY, RECORDING, OR ANY
INFORMATION STORAGE AND RETRIEVAL SYSTEM, WITHOUT
PERMISSION IN WRITING FROM THE PUBLISHER.

ACADEMIC PRESS, INC.
1250 Sixth Avenue
San Diego, California 92101

United Kingdom Edition published by
ACADEMIC PRESS LIMITED
24-28 Oval Road, London NW1 7DX

Library of Congress Cataloging in Publication Data

Shryock, Henry S
The methods and materials of demography.

(Studies in population)
Includes bibliographies.
1. Demography. I. Siegel, Jacob S.
II. Stockwell, Edward G. III. Title. IV. Se-
ries.
HB881.S526 1976 301.32'01'82 76-18312
ISBN 0-12-641150-6

Contents

PREFACE	vii	CHAPTER 16	
CHAPTER 1		Natality: Measures Based on Vital	
Introduction	1	Statistics	273
CHAPTER 2		CHAPTER 17	
Basic Sources of Statistics	11	Natality: Measures Based on Censuses	
CHAPTER 3		and Surveys	297
Collection and Processing of		CHAPTER 18	
Demographic Data	29	Reproductivity	313
CHAPTER 4		CHAPTER 19	
Population Size	49	Marriage and Divorce	333
CHAPTER 5		CHAPTER 20	
Population Distribution		International Migration	349
Geographic Areas	63	CHAPTER 21	
CHAPTER 6		Internal Migration and Short-Distance	
Population Distribution—Classification		Mobility	373
of Residence	83	CHAPTER 22	
CHAPTER 7		Population Estimates	407
Sex Composition	105	CHAPTER 23	
CHAPTER 8		Population Projections	439
Age Composition	113	Chapter 24	
CHAPTER 9		Some Methods of Estimation for	
Racial and Ethnic Composition	145	Statistically Underdeveloped Areas	483
CHAPTER 10		APPENDIX A	
Marital Characteristics and Family		Reference Tables for Constructing	
Groups	161	and Abridged Life Table by the	
CHAPTER 11		Reed-Merrell Method	511
Educational Characteristics	177	APPENDIX B	
CHAPTER 12		Selected “West” Model Life Tables and	
Economic Characteristics	191	Stable Population Tables, and Related	
CHAPTER 13		Reference Tables	521
Population Change	211	APPENDIX C	
CHAPTER 14		Selected General Methods	531
Mortality	221	AUTHOR INDEX	561
CHAPTER 15		SUBJECT INDEX	565
The Life Table	249		

This page intentionally left blank

Preface

The present work is a condensed version of the two-volume work *The Methods and Materials of Demography* first published by the U.S. Bureau of the Census in 1971. The original work has enjoyed widespread use both in the United States and abroad.

The idea for a condensed edition of *The Methods and Materials of Demography* was suggested by Professor Halliman H. Winsborough, the editor for Academic Press's Series on Studies in Demography. This condensed edition is intended to serve several purposes. Among these are to reach a wider audience, particularly in academic circles, and to serve the needs of many users and prospective users of the previous work more effectively. It is believed that a one-volume edition of about 555 pages can much more easily serve as a text in a one-semester course in the techniques of demographic analysis and could be covered more thoroughly in a two-semester course, than the original two-volume work of 900 pages.

The authors are pleased that Professor Edward G. Stockwell was willing to undertake the time-consuming and demanding task of abridging the two-volume original. He worked from the latest edition then in print, namely the second printing of 1973; but he had available the record of the numerous but relatively minor corrections that were to be made in the third printing, which was completed in September 1975. In our opinion he has done a consistent and judicious job of selecting parts to be omitted and of providing the necessary editorial modifications.

Like the original two-volume work, this work attempts to present a systematic and comprehensive exposition, with illustrations, of the methods used by technicians and research workers in dealing with demographic data. The book is concerned with how data on population are gathered, classified, and treated to produce tabulations and various summarizing measures that reveal the significant aspects of the composition and dynamics of populations. It sets forth the sources, limitations, underlying definitions, and bases of classification, as well as the techniques and methods that have been developed for summarizing and analyzing the data.

This book is intended to serve both as a classroom text for courses on demographic methods, aimed at instructing students in how to use population data for analytic studies, and as a reference for professional workers who have occasion to use population data; but it is designed primarily as a classroom text. For courses that are focused on the subject matter of population, this condensed edition might well be used for supplementary reading. As a reference work we hope that the book will be helpful to statistical analysts and planners in private industry, city and regional planners, technicians in State and regional development agencies, and also statisticians and analysts in national agencies concerned with population matters.

The book could be employed as a text in a single semester, particularly with the omission of a few topics. On the other hand, with a more complete coverage of the book, and with supplementary readings from other appropriate publications, a full year could be devoted to the book. In addition to copious citations of relevant publications in the footnotes, there are lists of suggested readings at the end of each chapter that will identify for the interested student the more important works on demographic methods.

In the hope of making the book as widely useful as possible, we have presented the methodological material in the simplest mathematical form. For the most part, only high school algebra and

elementary statistics are needed for a proper understanding of the book. However, a knowledge of elementary mathematical analysis is required for understanding some parts, particularly sections of Chapters 18 and 24, and Appendix C.

Considerable attention is paid to the data and measurement problems of the less developed countries, the kinds and quality of data that are available for these areas, and the special methods that have been worked out for handling incomplete and defective data. Although their demographic publications tend to be less adequate and less available, numerous illustrations for the less developed countries are given. The materials of the United States are also covered in relatively great detail, especially with respect to sources, definitions, and historical developments. Although national demographic materials differ considerably with respect to availability, definitions, classifications, etc., certain demographic principles and methods are essentially "culture free," and measures worked out for the United States could serve as well as for any other country.

The original work was intended to be comprehensive in its treatment of demographic methods and materials; yet, inevitably, choices had to be made as to what to include and what to exclude. These choices reflected the judgment of the authors as to what is most important and useful to demographers. This condensed version has tried to preserve the broad scope of the original work. As may be seen from the table of contents, the present book includes material on both formal demography and social and economic demography (i.e., the study of many social and economic characteristics of individuals as well as of such social groups as families and households). Several important but peripheral or applied fields are barely touched on, however. These include worker commuting, morbidity, the evaluation of family planning programs, optimum population, population quality, and various types of applied projections (e.g., needs for housing).

The derivation of most demographic measures described is illustrated by step-by-step examples using actual or, occasionally, hypothetical statistics. The official statistics are usually accepted as published for use in these calculations, and the handling of a problem is sometimes simplified, in order to focus attention on the method being illustrated. In actual practice, the official statistics for many countries would need to be evaluated and perhaps adjusted, and related procedures might be required. Accordingly, the results from the illustrative examples given here should not be regarded as necessarily valid for substantive purposes.

The original work was written in the years 1967 to 1970. Thus, it takes account of recommendations made by the United Nations and other international agencies for the 1970 round of population censuses and the plans for the enumeration, processing, tabulation, and publication of the 1970 Census of the United States. In this condensed version, the material on the 1970 Census of the United States has been updated so that the treatment of the census more nearly describes what was actually done.

The condensed edition retains the essential organization of the earlier work. In particular, all the main topics covered in that work are retained, although necessarily many details had to be omitted in the condensation. The condensation did not involve a mere mechanical reduction to achieve the brevity intended, however. Where sections were deleted, the new text was woven together again to achieve logical continuity.

The earlier work consisted of 25 chapters, grouped under five broad headings, and four appendixes. The major structural changes in the abridged edition are (1) the deletion of Chapter 22 on "Selected General Methods" and the consequent renumbering of subsequent chapters; (2) the incorporation of what was Appendix A in the larger two-volume edition into the new Chapter 23 on "Population Projections," and the renumbering of Appendixes B and D as A and B; and (3) an expansion of Appendix C to include some of the material originally covered in the now deleted Chapter 22. As compared with the unabridged edition, the present one contains less emphasis on historical development, the recommendations of international organizations, the practices of national statistics agencies with respect to demographic data, and the more technical and mathematical aspects of demographic analysis.

Either Shryock or Siegel as principal authors wrote, or reviewed and edited, each chapter in the

original work. Associate authors contributed drafts of many of the chapters or parts of chapters, either working under the immediate direction of one of the principal authors or independently. These associate authors were Maria Davidson, Paul C. Glick, Elizabeth A. Larmon, Wilson H. Grabill, and Charles R. Kindermann of the Bureau of the Census; Robert D. Grove and Robert A. Israel of the National Center for Health Statistics; Charles B. Nam of Florida State University; Abram J. Jaffe of Columbia University; Francisco Bayo of the Social Security Administration; and Paul Demeny of the University of Hawaii. Elizabeth Larmon was editorial coordinator of the unabridged edition and was generally in charge of the technical production aspects of that work. Adriana Weininger performed these tasks in the final stages. We are grateful to the Bureau of the Census and particularly to the Population Division and the International Statistical Programs Division for its generous support of the work on this book. Acknowledgement is made of the role of the U.S. Agency for International Development in providing major financial support for the original publication.

The bulk of the retyping that was necessary for the condensed edition was done by members of the secretarial staff in the Department of Sociology at Bowling Green State University, with particular acknowledgement being due to Laretta Lahman and Lynn Schmid. Thanks are also due to Helen S. Curtis, Anne Donnelly, and Mary Hartman of the Center for Population Research, Kennedy Institute, Georgetown University, for helping with the voluminous correspondence in connection with the present edition, and to Georgetown University for other courtesies.

The first-named author would like to acknowledge his appreciation to his wife, Annie Frances King Shryock, for her patience during his work on the various drafts and editions, and for her help with some of the editorial work. The second-named author would also like to express his appreciation to Rose V. Siegel and Lorise V. Siegel for their forbearance and sympathetic support during the long period when the original work was being prepared.

Henry S. Shryock
Jacob S. Siegel

This page intentionally left blank

CHAPTER 1

Introduction

THE FIELD OF DEMOGRAPHY

Demography is the science of population. The word was coined by a Belgian, Achille Guillard, who published his *Eléments de statistique humaine, ou démographie comparée* (Elements of human statistics or comparative demography) in 1855.¹ He defined it as the natural and social history of the human species or the mathematical knowledge of populations, of their general changes, and of their physical, civil, intellectual, and moral condition.

Like most other sciences, demography may be defined narrowly or broadly. The narrowest sense is that of "formal demography." Formal demography is concerned with the size, distribution, structure, and change of populations. **Size** is simply the number of units (persons) in the population. **Distribution** refers to the arrangement of the population in space at a given time, that is, geographically or among various types of residential areas. **Structure**, in its narrowest sense, is the distribution of the population among its sex and age groupings. **Change** is the growth or decline of the total population or of one of its structural units. The components of change in total population are births, deaths, and migrations. In analyzing change in structure, however, we have to include the transition from one group to another. In the case of age, this is expressed as a simple function of time.

A broader sense includes additional characteristics of the units. These include ethnic characteristics, social characteristics, and economic characteristics. Ethnic characteristics like race, legal nationality, and mother tongue shade into social characteristics. Other examples of social characteristics are marital and family status, place of birth, literacy, and educational attainment. Economic characteristics include economic activity, employment status, occupation, industry, and income, among others. Other characteristics that might be encompassed are genetic inheritance, intelligence, and health; but the usual sources of demographic data, such as censuses, seldom deal with these directly. Furthermore, demography may look beyond the basic personal units to such customary social groupings as families and married couples.

The widest sense of demography extends to applications of its data and findings in a number of fields including the study of problems that are related to demographic processes. These include the pressure of populations upon resources, depopulation, family limitation, eugenics, the assimilation of immigrants, urban problems, legislative apportionment, manpower, and the

maldistribution of income. This book covers very little of these fringe areas of demography.

Hauser and Duncan regard the field of demography as consisting of a narrow scope—demographic analysis—and a wider scope—population studies. "Demographic analysis is confined to the study of components of population variation and change. Population studies are concerned not only with population variables but also with relationships between population changes and other variables—social, economic, political, biological, genetic, geographical, and the like. The field of population studies is at least as broad as interest in the 'determinants and consequences of population trends.'"² In the same work, Lorimer also discusses demography as a discipline.³

DEMOGRAPHIC DATA AND THEIR USES

The types of demographic data have already been suggested. Counts of persons are obtained from censuses and sample surveys and from the files of continuous population registers. Counts of events are obtained from registered vital events (births, deaths, marriages, divorces, etc.), and also from continuous population registers. Sometimes censuses and surveys inquire about events, e.g., the number of children women have borne in the preceding 12 months. Any of these counts may be shown in the form of multiple classifications, e.g., population by sex and age for urban areas or deaths by age and cause. Various demographic measures, such as percentages, ratios, and averages may be derived from them. Furthermore, numbers of registered vital events can be related to the corresponding population to produce vital rates, for example, the number of deaths per 1,000 of the population or the number of births to married women 20 to 24 years old. Such vital rates can also be derived from continuous population registers without the use of other sources of demographic data.

The resulting demographic statistics can then be used to describe the distribution of the population in space, its density and degree of concentration, the fluctuations in its rate of growth, its movements from one area to another, and the force of natality, nuptiality, and mortality within it. These demographic statistics have many and increasingly varied applications. The fields of application include public health; local planning for land use, school and hospital construction, public

¹ Adolphe Landry, *Traité de démographie* (Treatise on demography). Paris, Payot, 1945, p. 7.

² Philip M. Hauser and Otis Dudley Duncan, "Overview and Conclusions" in: *The Study of Population*, Hauser and Duncan, eds., Chicago, University of Chicago Press, 1959, pp. 2-3.

³ Frank Lorimer, "The Development of Demography" in Hauser and Duncan, op. cit., pp. 157-166.

utilities, etc.; marketing; manpower analysis; family planning programs; land settlement; immigration and emigration policy; and many others. An analysis of current demographic levels and past trends is the necessary first step in the construction of population forecasts that in turn form the underpinning of national plans for economic development and other programs, including explicit population policies in some cases.

COVERAGE OF TOPICS

As may be seen from the table of contents, the scope of this book corresponds pretty much to the first and second concentric circles of the field of demography that were described in the opening section, that is, it covers formal demography and many additional social and economic characteristics of the population as well as such social groups as families and households. Several important but peripheral or applied fields are barely touched on. These include worker commuting, morbidity, pregnancy wastage through miscarriages and induced abortions, the evaluation of the effectiveness of family limitation programs, and population quality.

ORGANIZATION OF THIS BOOK

Arrangement of Chapters

The book consists of 24 chapters covering the sources of demographic data, the major topics of population size, distribution, and composition, the basic components of demographic change, and population estimates and projections.

A major problem that had to be faced in the writing was the fact that the subject and methods of a given chapter cannot be adequately described without drawing on those of a later chapter. This problem would arise no matter what order of topics or chapters was followed. In the analysis of age composition in chapter 8, for example, it is essential to make use of survival rates, which are derived by methods explained in chapter 15, "The Life Table." To understand the life table, on the other hand, we have to know how age is defined in censuses and in vital statistics. More particularly, we need to be familiar with the sources of error in the underlying age statistics in order to have some appreciation of the accuracy of a given life table.

A related problem is that a given method may apply to a number of subject fields within demography. Standardization is used with almost all kinds of rates and averages—birth, death, and marriage rates; migration rates; enrollment rates; worker participation rates and unemployment rates; median years of school completed; mean income; etc.

We have tried to cope with these problems in several ways. Frequent use has been made of forward and backward references. Nonetheless, a certain amount of duplication in the exposition was inevitable.

Treatment of Subjects Within Chapters

In addition to the conventional treatment of standard methodological issues and measures, each chapter generally contains a discussion of their uses and limitations, the quality of the statistics, and some of the factors important in their analysis.

Uses and Limitations.—The uses to which the given statistics and measures derived from them are or can be put are sometimes illustrated by actual examples from the literature. These uses often represent contributions to the investigation of basic social and economic problems or of challenging scientific questions.

The limitations of the statistics have to do with their scope and pertinency. Because of the definitions used, the restricted number and form of the questions, the scope of the census, survey, or register, and the extent of cross-classifications in the tabulations, the statistics can at best give only approximate answers to the broad questions that they were designed to answer and may be even less adequate for the purposes of some imaginative analysts. Demographers often find new and previously unforeseen uses to which they put the statistics available to them. When they do so, they need to recognize that the statistics may not apply to just the universe that is of interest to them or may measure obliquely rather than directly the phenomenon they are investigating. In short, they should be aware of these limitations and the implications of them for their analysis and conclusions.

The demographer should also be aware of the assumptions that he has to make when he extends the interpretation of standard, readily available measures or models in certain ways. For example, a life table based on the mortality experience of a given year does not describe the mortality experience of any actual group of persons as they pass through life. Neither does a gross reproduction rate based on the fertility experience of a given year describe the fertility experience of any actual group of females who started life together. With due caution, however, these measures may be used in many important descriptive and analytical applications.

We may cite also the limitations of population projections when they are actually used to predict the future. The fact that, in some situations, this type of use is unavoidable does not relieve the forecaster of considering the realism and appropriateness of the assumptions underlying the projections. Often, too, he may realize that if he had additional data at his disposal, he could express his assumptions in terms of more meaningful metrics.

Quality of the Statistics.—By the "quality" of the statistics is meant the degree to which they measure what was intended to be measured when the questions and procedures were designed. Hence, the term quality as used here is measured *within* the limitations of the statistics and does not include those limitations. "Accuracy" could be used as a synonym for quality.

Factors Important in Analysis.—Only a limited understanding of demographic phenomena is obtained from statistics (percentage distributions, rates, etc.) presented for a topic in isolation. Such so-called "inventory" statistics have mainly a descriptive value. Inventory statistics for a large number of small geographic areas can be very useful in ecological analysis; but here an important additional factor, namely, geographic location has already been introduced. In some countries, racial or ethnic group may be associated with striking variations in the topic (mortality, literacy, family income, etc.) being investigated.

The "factors" treated in these subsections can be regarded in several different ways. In ordinary statistical terms, they are the "independent variables" used to study variations in the "dependent variable" (or the topic under consideration). Hence, they may be introduced into the analysis in various ways as "controls" or as variables to be held "constant." The most important control factors in much demographic analysis are age and sex. Any of these factors may be introduced by means of cross-classification (and then by computation of factor-specific rates), by standardization, or by correlation.

To a lesser extent, the topic in question may itself be treated

here as an independent variable. In that case, we study its effects on other demographic variables (or qualitative attributes).

More broadly, these subsections occasionally mention a few comprehensive studies in which the topic is an important element. These studies place the topic in the context of substantive research and give the reader a broader view of how the data and methods he has just examined in detail have been applied in some of the classical demographic works.

DESIDERATA IN DEMOGRAPHIC DATA AND STATISTICS

There are certain properties that demographers would like their data to possess. Taken literally these desiderata are of the nature of ideals in that even in the most advanced countries they are never fully attained. Nonetheless, they are goals that should be kept in view by agencies that produce demographic data. Moreover, whenever compromises are made with these goals because of practical restrictions, the nature of these deviations should be imparted to the public in the resulting statistical publications. Furthermore, when unintended errors of coverage, classification, etc., occur in the data gathering and processing, some resources should be devoted to evaluating them, on a quantitative basis if possible.

We are concerned with all types of demographic data—those from censuses, sample surveys, vital records, continuous population registers, and immigration and emigration controls. In general, the data should completely cover the units of the specified universe or a representative sample of them. The data should be comparable over time and space. They should be published promptly and in adequate detail. Some considerations apply to only censuses and others to vital statistics, and so on. Let us start with more specific consideration of censuses.

Censuses

A recent U.N. publication lists four essential features of a population census:

1. Each individual is enumerated separately and the characteristics of each person are recorded separately.
2. The census covers a precisely defined territory and includes every person present or residing within its scope.
3. The population is enumerated with respect to a well-defined point of time, and data are in terms of a well-defined reference period.
4. The censuses are taken at regular intervals.⁴

The Census as Part of an Integrated Program of Demographic Statistics

In planning a national population census, attention should be paid to its interrelationships with sample surveys, vital statistics, population registers, etc. The census may serve as a frame for sampling surveys to be taken shortly after the field work of the census is completed; such surveys may probe in greater depth topics included in the census, or they may investigate additional topics. In the case of periodic intercensal surveys, the census may serve as a benchmark in various ways. For example, it may serve as an occasion for updating the population "controls" to which the percentage distributions obtained in the survey are applied. As part of the census evalua-

tion program, matching studies can be carried out between a sample of cases in the census and a survey taken at a nearby date. Such matching studies can also be carried out between the census and vital records, universal population registers, etc. Statistical totals can also be compared to measure the net differences. In the case of comparisons between the census, on the one hand, and a survey, population register, etc., on the other, both demographic programs will benefit from these matching studies and statistical comparisons unless one source is vastly superior to the other.

To achieve an integrated program of demographic statistics, it is essential that consistent concepts and definitions be employed throughout. When the respective programs are administered by different agencies, however, consistency may be difficult to achieve in some particulars since different program needs may seem to call for different treatments.

By the same token, international comparability is even more difficult to accomplish. Moreover, there may be such profound cultural, social, and economic differences among countries that, even if the phrasing and definition of questions are identical (and if identity can be achieved through translation), the resulting data from certain questions may not be comparable. Let us cite a few examples. If most of the people in a Far Eastern country are accustomed to figure their ages by the Oriental mode, simply asking for the age in completed years will not necessarily yield the Western-style age. More elaborate artifices may be required, and even these may yield only an approximation to the Western-style age distribution. If plural marriage or consensual unions are widespread, the conventional Western categories of marital status may give an unreal description of the population's distribution by marital status. Finally, if most of the working population are not part of a market economy and if there is considerable underemployment, the conventional labor force questions will yield meaningless statistics on unemployment. As Linder puts it, "A standard international definition for such characteristics can achieve comparable statistics only to the extent that the social and economic features of the country are similar. In two dissimilar countries, it may be that greater comparability of meaning for a census item can be obtained if the census question is asked in distinctly different ways."⁵

As source material, demography requires both regular censuses and the registration of vital events. In theory, a continuous, universal population register might suffice; but, in practice, even the best registers need to be checked periodically by a census. Moreover, a register cannot ordinarily accommodate as many items of social and economic characteristics of the person as a census can nor can new items be introduced so readily. Where census and vital records are not both adequate, as is true in many statistically underdeveloped countries, the demographer will work with what is available, using, for example, the techniques described in chapter 24.

Sample Surveys, Registration Systems, and Other Sources

Current sample surveys like censuses represent periodic stock-takings, whereas registration systems (universal population registers, vital records, etc.) and immigration and emigration control systems involve the continuous recording of data

⁴ United Nations, *Principles and Recommendations for the 1970 Population Censuses*, Statistical Papers, Series M, No. 44, 1967, pp. 3-4.

⁵ Forrest E. Linder, "World Demographic Data," in Hauser and Duncan, eds., op. cit., pp. 323-324.

even though the published statistics are extracted only at periodic intervals. Thus, some of the desiderata in census data are not applicable to registration data.

Sample Surveys.—Unlike censuses, many surveys are taken under private auspices. Although a periodic sample survey constitutes a continuing “micro-census” of a country or other area, a one-time survey can provide useful data for an *ad hoc* purpose. Surveys taken as census pretests or as experiments in survey methodology need not necessarily result in published tables based on the substantive demographic data. In other respects (individual enumeration, universality, and simultaneity), the scope of the survey and the methods of recording the data should be subject to the same principles as are censuses.

Registration Systems.—In the interests of national uniformity, it is best for a register to be operated by an agency of the national government. In some countries with a federal form of government or at least with relatively autonomous state governments, however, good vital statistics have been produced under an arrangement whereby the states handle the registration process and the national government exercises a coordinating function, setting standards and publishing national reports, for example.

The equivalent of the desideratum “individual enumeration” in the census or survey is the individual vital record. As Hauser and Duncan put it, “To provide adequate data for demography, a vital registration system must include procedures to assure the filing of a uniform record for every vital event—for example, live birth, death, stillbirth, and marriage; to provide for complete and usable answers to the inquiries on the record form; and to enable the information in the record form to be processed for statistical purposes—that is, edited, coded, tabulated, and presented, preferably through some central office which provides vital statistics for the nation and its subdivisions on a comparable basis. Principles and procedures for achieving these objectives have been evolved over the years.”⁶ (See also United Nations *Handbook of Vital Statistics Methods*, Studies in Methods, Series F, No. 7, 1953.)

A registration system should likewise have “universality within a defined territory.” The principle of “simultaneity” does not apply, of course; but it suggests another criterion that is pertinent to registration, namely, a specified maximum interval of time between the occurrence of the event and its recording. For most demographic purposes, moreover, the data should be tabulated as of the date of occurrence, not the date of recording. On the other hand, tabulation by the place of occurrence is less useful than tabulation by the place of residence of the person concerned.

Again, “defined periodicity” would not apply to the continuous recording of data, but it does suggest to us that compilations of the records for statistical purposes should be made periodically. (The date—day, month, year—of both occurrence and registration should be on the record form.) The usual interval for publication purposes is the year; but some series should be published quarterly, or even monthly. Finally, the compilers of statistics from registers, like the agencies that collect and tabulate census data, are obligated to publish, evaluate, and, to some extent, analyze them.

SOME BASIC DEMOGRAPHIC METHODS

As was noted above, it is not possible adequately to discuss demographic methods following a simple linear path. At many points, we have to refer to what lies ahead, or to the left or right, as well as to what lies behind. Partly for this reason, an early overview of how demographers organize and analyze their data seems necessary.

A good summary account of the essence of demographic methods is given by Hauser and Duncan.⁷ After first pointing out that demography shares many of its methods with science in general and especially with statistics, they list certain groups of techniques that tend to be peculiar to demography. These techniques include: (a) techniques of data collection; (b) techniques of data evaluation and adjustment and of statistical estimation; and (c) techniques of analysis including demographic projections or forecasts. We will not comment here on (a) but rather will focus on (c) with some attention to (b), since the respective techniques are interdependent to some extent.

“For dealing with population ‘statics’ the demographer depends largely on general statistical descriptive techniques with some special rates and graphic devices which have become widely used. In this latter category are such things as the sex ratio, the dependency ratio, the index of displacement, and the population pyramid . . . To deal with population dynamics, demography has developed a rather comprehensive and elaborate set of ‘rates’ designed to measure vital events or components of population change, such as natality, mortality, morbidity, marriage, divorce, and migration. . . . On the whole, demographic rates are designed to measure change and are calculated as approximations to a posteriori probability statements.”⁸

The Balancing Equation

The most basic method of demography is the decomposition of population change into its components, or, conversely, the synthesis of the components to estimate the total population change. Schematically, we may express this process in terms of the fundamental equation

$$P_t - P_0 = B - D + I - O, \quad (1)$$

where P_t is the population at the end of the period, P_0 that at the beginning of the period, B is births, D is deaths, I is in-migration, and O is out-migration.

This simple equation, which is called the “balancing equation” (or the “inflow-outflow relationship” or the “component equation”) has many forms and many uses. To be exactly true (i.e., represent a necessary relationship), it must apply to a fixed territory and there must be no measurement errors. In fact, the equation may be used to estimate the net error in this system of demographic statistics. If we find that the right-hand side differs from the left-hand side by an amount e , then we can write,

$$P_t - P_0 = B - D + I - O + e \quad (2)$$

Here e can be called the “residual error” or the “error of closure.” On the basis of additional knowledge about the

⁶ Hauser and Duncan, “The Data and Methods,” in Hauser and Duncan, eds., *op. cit.*, p. 62.

⁷ *Op. cit.*, pp. 70–73.

⁸ *Ibid.*, p. 70.

accuracy of the various terms, one may be able to decide whether e can be attributed as a measurement error almost wholly to a particular term in the equation. For example, if there is evidence that the right-hand terms are all based on very accurate registration data and the population figures come from successive censuses, then e would represent the relative accuracy of coverage of the two censuses. If e is positive, P_t is more nearly complete than P_o ; if e is negative, then the reverse would be true.

Let us consider some other illustrations of the uses of (1) or its variations. Suppose that a country has adequate vital statistics and statistics on immigration and emigration. Then t years after the last census but before the next census, it is desired to make a postcensal estimate of the current national population. We have

$$P_t = P_o + B - D + I - O \quad (3)$$

In this form, the equation may be thought of as the "basic estimating equation," which uses a straightforward book-keeping procedure.

If we are interested in projecting the population to a future date, we can use equation (3) in principle by making assumptions about the future births, deaths, and migration. Especially in the case of births and deaths, however, these assumptions are ordinarily made in the form of fertility and mortality rates, not in the form of the absolute numbers of births and deaths. The nature of these rates is another fundamental part of demographic methodology and will be discussed presently.

For another application, suppose that we have two successive population counts for a subnational area (province, county, commune, etc.). We also have vital statistics on births and deaths but no statistics on internal migration (or on the extent to which external migration affects the individual subnational areas). Then we may write

$$M = I - O = (P_t - P_o) - (B - D) \quad (4)$$

where M is the net migration to or from the area. In other words, to estimate the intercensal net migration for the subnational area, we subtract the natural increase, $B - D$, from the total population change, $P_t - P_o$.

First, however, we should mention another kind of elaboration of the balancing equation, namely, its use for a population subgroup, such as the male population, the female population of childbearing age, the native population, or university graduates. For some subgroups, the males in the native population, for example, we simply have to obtain the corresponding components, i.e., statistics on births, deaths, and migration for that subgroup. This restriction may be expressed by using the superscript i , to denote the subgroup, thus:

$$P_t^i - P_o^i = B^i - D^i + I^i - O^i \quad (5)$$

In the case of an age group, however, the very specification of the group, i.e., its age, changes over the period. For example, if t is 10 years, then we should compare age x at time 0 with age $x + 10$ at time t . The identification of age is more complicated in the case of the components. For example, we shall need to have death statistics for ages x , $x + 1$, $x + 2$, . . . , depending on the time elapsed since the first census, and then we shall need to redistribute them into other age groupings. For persons aged x at that census, for example, we shall need to

have part of the deaths at age x , part of those at age $x + 1$, $x + 2$, etc., depending on the time elapsed since the census.

In this situation although age changes, there is something about the population subgroup that remains the same. A group having such a common property is what demographers call a "cohort." Here the cohort is all the people who were born in a given year or their survivors. There are other types of demographic cohorts, such as marriage cohorts, or all the marriages that occurred in a given year. A great deal of demographic analysis is carried out in terms of cohorts.

Not all types of population subgroups are entered or left only by means of birth, death, migration, or growing older. The numbers in most social and economic subgroups are also affected by changes of status. For example, the number of citizens of a country is changed by naturalizations and losses of citizenship as well as by demographic factors. The number of married persons is increased by marriages and decreased by the number of "widowings" and divorces. In general, there are at least two gross components, a minimum of one positive and one negative one, that must be added to the right hand side of equation (3) to allow for these changes in status.

We analyze changes in population subgroups not only because we are interested in particular subgroups *per se* but also because this form of decomposition of population change gives us a better understanding of the nature of changes in the total population and of variations over time in these changes. The demographer's interest in the factors to be taken into account in this decomposition process is often limited only by the data that are available to him. The decomposition is more effectively carried out by multiple cross-classification rather than by simply distributing the population by one factor at a time. An example of a multiple cross-classification would be the distribution of the population by age, sex, and marital status, in which one subgroup or cell is the number of single females 15 to 19 years old. "Having accomplished what he regards as a suitable decomposition of population changes (or variations in rates of change from one period or place to another), the demographer may, of course, put the components back together again, e.g., in the form of a 'balance sheet,' a mathematical model, a statement of the relative importance of the several components, or the backward or forward projection of population changes, which may involve assumptions about the several components and the ways in which they are likely to change."⁹

The general topic dealing with deaths is called **mortality** and that with births, **natality** or **fertility**. These terms apply both to absolute numbers and to rates, and both to totals and to births and deaths specific with respect to various characteristics.

Rates and Ratios

As we have mentioned, demographic analysis also makes abundant use of rates. The term "rate" most appropriately applies to the number of demographic events in a given period of time divided by the population at risk during that period. Thus, we may speak of the number of deaths divided by the population as the death rate. The population at risk is usually only approximated. It may be the population at the middle of the period (which is roughly the average population during the period), the population at the beginning of

⁹ Hauser and Duncan, "Overview and Conclusions," in Hauser and Duncan, eds., *op. cit.*, p. 4.

the period, or a more complex definition. The period is usually a year and the rate is often expressed per 100 or per 1,000 of the population. For example, the **crude** birth rate, which is based on the midperiod population, is

$$b = \frac{B}{P} \times 1,000 \quad (6)$$

where B is the number of births in the year and P is the mid-year population. Similarly, the **crude** death rate is given by,

$$d = \frac{D}{P} \times 1,000 \quad (7)$$

The adjective “crude” is especially appropriate in the case of the crude birth rate because obviously men, children, and old people are not at risk of having a baby. (Although rarely used, there are “paternal fertility rates,” which relate the number of children sired to the number of adult males.) In the case of the crude marriage rate, persons who are in the married state throughout the year are likewise not at risk. Various refinements are introduced in the population base in order to obtain a more meaningful rate. These may not only omit that part of the population for which the risk is zero, but they may also take account of the fact that the risk is much greater for some population subgroups than for others. For example, all persons are at risk of dying but the risk is much greater at age 90 than at age 10. Accordingly, we have age-specific rates, for example, the number of deaths during the year to persons aged 20 to 24. The general formula is

$$n r_x = \frac{n E_x}{n P_x} \quad (8)$$

where E is the number of events, x is the initial age, and n is the number of years in the age group. Note that here the numerator is restricted so as to correspond to the age restriction in the population. Rates may be specific for any other characteristic into which the demographic data are subdivided. Common characteristics are sex, race, nativity, and marital status. Moreover, the rates may be specific for two or more characteristics simultaneously.

In addition to the distinction between crude and refined rates, demography follows actuarial science in distinguishing central rates from probabilities. In central rates, the denominator is the population at the midpoint of the period (or the average population during the period). Probabilities, on the other hand, are based on the population at the beginning of the period, which is then viewed as the population at risk of experiencing the event during the period. The distinction is somewhat blurred for short periods when the population is not closed, e.g., when it is subject to immigration or emigration. In an “open” population, the midperiod population may be viewed as the average population at risk during the period, or the number of person-years that the population is at risk.

Suppose we are concerned with deaths at a given age, x , in a given year. The central death rate would be given by the number of deaths occurring at age x in the given year divided by the population of that age at the middle of the year. The probability of death, on the other hand, would be given by the population at exact age x at the beginning of the period divided into the deaths occurring at age x during the year; the deaths are in the same cohort as the population. When the population is not closed, the inclusion of any deaths of immigrants at age x and the exclusion of any deaths of emigrants make the

resulting rate not strictly a probability because the deaths are no longer restricted to those occurring to the initial population.

The term rate is also loosely used to refer to the ratio between a population subgroup and the total population where the definition of the subgroup reflects a prior event. For example, the number of persons reported in the census as migrants to an area during t preceding years divided by the total population of the area is often called the in-migration rate. The underlying justification here is that the number of enumerated migrants roughly approximates the number of persons who migrated into the area during the specified t years—although it always falls short of that number. The population counted in the census is likewise only an approximation to the population at risk of migration during a specified period. A rate like an illiteracy rate is even more similar to a simple proportion or percentage. It is simply the percent of persons in a given population subgroup who are classified as illiterate. The acquirement of literacy may have occurred many years earlier.

Still other types of ratios are called “rates” in demography. For example, the infant (or infantile, in British usage) mortality rate,

$$\text{IMR} = \frac{\text{Deaths to children under 1 year of age during the year}}{\text{Births during the year}} \times 1,000 \quad (9)$$

Refinement of the rate takes account of the facts that some of the deaths in the numerator occurred to births of the preceding year and that some of the births in the denominator are at risk in the following year.

Survival rates are very frequently used in demographic analysis. In a closed population (i.e., a population with no external migration), a survival rate is the ratio of the number of persons in a cohort at one date to the number at an earlier date. Survival is from a given age to a subsequent age. There are 1-year, 5-year, 10-year, etc., survival rates. Survival rates are typically calculated from age distributions in two successive censuses or from a life table. Survival rates, too, may be specific for various characteristics or population subgroups, e.g., we may have age-specific survival rates for females, married men, or the country's aboriginal population.

By multiplying a population subgroup at one census by its appropriate survival rate, we obtain the **expected population** t years later. The expected population can be used in a number of ways—to make estimates or projections of the population, to estimate net migration during the period, to help in estimating the contribution of certain factors to the total change during the period, or to estimate the relative completeness of two successive census counts. By dividing the survival rate into the population, one obtains an estimate of the size of the cohort t years earlier, for example at the preceding census date.

In general, rates are of interest in their own right in describing the dynamics of population. They may also be applied, as just illustrated, by multiplication or division to demographic aggregates to obtain various types of estimates. For example, a death rate for a prior period or one from a life table may be multiplied by an appropriate population to estimate the number of deaths in a more current period. Conversely, on the assumption that the death rate has remained constant, a rate for an earlier period (say the census year) may be divided into the number of registered deaths to obtain a postcensal estimate of the population.

We have shown that the distinction between “rates” and “ratios” in demography, as elsewhere, is somewhat fuzzy.

Restriction of the term "rate" to probabilities or to fractions where the numerator is part of the denominator is not observed in actual practice. There is a tendency for what are called "ratios" (e.g., the sex ratio or the dependency ratio) to be used for descriptive purposes in population studies and for rates to be used in the analysis of change, i.e., in population dynamics: but this is only a tendency not a universal rule. Some writers even use the term "survival ratio."

Life Tables

Several references have already been made to the life table. A whole chapter, chapter 15, is devoted to this subject—the definition and interpretation of life table functions, the construction of life tables, and some of the applications of life tables in demography. Other applications are scattered through the book.

The primary purpose of the life table for the actuary is to measure the expectation of life at each age, e_x . The table is built up from age-specific death rates, conventionally those observed in a single year or period of years for all cohorts that were alive and hence subject to the risk of dying in that period.

An example of a life table is given in table 15-1. In addition to the expectation of life at age x , it may be observed that the life table displays a number of other functions. These are:

- ${}_nq_x$ proportion of persons alive at beginning of age interval dying during interval
- l_x number living at beginning of age interval of 100,000 born alive
- ${}_nd_x$ number dying during age interval
- ${}_nL_x$ number living in the age interval
- T_x number living in this and all subsequent age intervals

The subscript n is the number of years of age in the age interval. For a complete life table, it is 1; for an abridged life table, it is usually 5. A more precise definition of e_x is the average number of years of life remaining at the beginning of the age interval. ${}_nL_x$ may also be defined as the number of man-years lived in the age interval, and ${}_nT_x$ as the number of man-years lived in this age interval and all subsequent age intervals.

A life table ordinarily starts with 100,000 births, a benchmark which is called the radix. ${}_nL_x$ may also be viewed as describing the age distribution of a population called the **stationary population** because it is constant over time. This population is generated by the constant 100,000 annual births and the death rates of the life table. The survival rates previously referred to may be derived from the ${}_nL_x$ values. Thus, the probability of surviving k years for persons in the age group x to $x+n$ is given by

$$\frac{{}_nL_{x+k}}{{}_nL_x} \quad (10)$$

The survival rate is a particularly useful measure in demography.

Unabridged life tables are usually constructed by actuaries, who are concerned with the careful graduation or smoothing of the basic data or the calculated values. There are, of course, many actuarial applications of life tables in the field of insurance. Demographers often construct their own life tables from deficient data, for special population subgroups, or by shortcut methods.

Both actuaries and demographers are interested in what are called **multiple-decrement tables**. In these tables, the popu-

lation is subject to attrition not only from the force of mortality but also from that of some other factor or factors as well, for example, marriage, widowhood, or entry into the labor force. Thus, we might start with 100,000 single persons and apply to their "survivors" at successive ages not only mortality rates (${}_nq_x$ values) for single persons but also a proportion that represents the probability of marrying during the interval.

Demographic Models

Increasing use is being made of model-building in demography. The life table is one such model. More complex models link together component models of fertility, mortality, and migration. Most models, however, represent closed populations. In addition to the life table, one of the best-known models is the "stable population."

Stable Population.—Demographers have long been interested in the question of what population structure would result if fixed age-specific schedules of mortality and fertility rates remained in effect indefinitely. More specifically, what would be the age-sex composition of such a population and what rate of growth would it have? The solution was found by Lotka and was presented in a classic paper.¹⁰ He proved that the resulting population would eventually have a fixed age composition and a fixed rate of growth and provided the formulas for these. This population is called the **stable population** because its age composition is stable. Unlike the life table population, which is stationary as well as stable, it may increase or decrease in absolute numbers.

There are other parameters of (measures associated with) the stable population that are widely used in demographic analysis. These include the **true (or intrinsic) birth and death rates**, the **mean length of generation**, and the **gross and net reproduction rates**. The reproduction rates are particularly important and deserve a brief explanation at this point.

The gross reproduction rate represents the average number of daughters that would be borne by a cohort of females all of whom lived to the end of the childbearing period if the cohort bore children according to a given set of age-specific fertility rates. The net reproduction rate removes the assumption of no mortality before age 50, say, and represents the average number of daughters borne by a cohort of females starting life together, if there were no changes in the age-specific schedules of fertility and mortality.

The concept of stable population has proven itself a powerful tool in demographic analysis. For example, when an underdeveloped country has very inadequate demographic statistics but can be assumed to have had roughly constant fertility for many years, its age structure, life table, gross reproduction rate, etc., can be estimated fairly closely under some circumstances on the basis of a very few simple statistics using stable population theory.

On the other hand, when demographers tried to interpret or even apply such measures as the net reproduction rate in situations where fertility was changing very sharply, certain inconsistencies and limitations in these measures became apparent. Efforts were then made to refine the concepts and measures. Probably the chief modification was the shift from period to generation (cohort) measures, not only for making

¹⁰ Louis I. Dublin and Alfred J. Lotka, "On the True Rate of Natural Increase," *Journal of the American Statistical Association*, 20(150):305-339, September 1925.

projections of reproduction but also for gauging its force more realistically. Another refinement was to take account of the parity of the women and order of birth of the child when computing age-specific fertility rates. **Parity** is the number of children previously borne. Zero-parity (childless) women are at risk of having a first birth, one-parity women of having a second birth, and so on. Measures using the **interval** between marriage and a birth of a given order and between births of successive order have also been developed.

Cohort Analysis

In an earlier section, we defined and illustrated a cohort. Other examples are the persons immigrating in a given year and the persons completing a given year of school in a given year. A great deal of demographic analysis is carried out by arranging the data in cohort form, particularly when changes over time are being studied.

Cohort statistics are distinguished from **period statistics**, the latter applying to a combination of cohorts in a given year or other period. A life table or a conventional reproduction rate describes a period such as a year or group of years. There are corresponding cohort forms, which are called "generation life tables" and "generation reproduction rates," respectively.

The statistical ingredients are the same for cohort as for period analysis, the difference lies in how they are put together. The death rate of a given age group in a given year can be viewed as applying either to a period of time or to a cohort. When we compare the number of persons 40 years old at two different censuses, we are comparing two different cohorts at the same stage of life, whereas when we are comparing the number of persons 40 years old with the number 50 years old 10 years later we are analyzing change within a single cohort. Both comparisons are of value to the demographer.

Cohort analysis may be carried out either forward in time or in the reverse direction. For example, **reverse survival rates** are applied to estimate the size of a cohort at earlier dates.

Controlled and Uncontrolled Factors

Much of demography is concerned with the comparisons of populations and of demographic processes in space and time. In making such comparisons, one quickly encounters a common problem in scientific analysis—that of "uncontrolled" factors. For example, if the force of mortality is being compared for two populations, one must ask whether the difference in the crude death rates can be explained by differences in their age composition—since mortality is so highly correlated with age. The demographer applies a number of methods that are in general use in social statistics in order to cope with this problem. Some of these methods, however, have been adapted in specific ways in demography; indeed, some of them have had their greatest technical elaboration in demographic applications.

One general method, namely that of **cross-classification**, has already been mentioned. Reverting to the previous illustration, we may compare the mortality of two populations using specific rates, age for age. If there are no errors in reporting age in the census or in the death records, the finer the age detail, the better the control on that factor. We may also control or "hold constant" more than one factor at a time by multiple cross-classification. Thus, we can have death rates specific for age, sex, and race.

This method of decomposition enables us to make mortality comparisons without the disturbing influence of the specific factors used in the cross-classifications. We have lost, on the

other hand, one advantage of the crude rates, namely an overall summary comparison. By the technique of **standardization**, we can derive from these factor-specific rates a summary measure that, to some extent, holds constant the influence of the factor. Thus, we can compare age-standardized mortality rates for two populations or for the same population at two different dates. What is required here is the age distribution (or distribution by the factor to be controlled) for a **standard population**. The standard population may be one of the two populations concerned, their averaged distributions, or a third population, for example, in the case of comparison of two or more provinces, the total population of their country. The population of England and Wales in 1901 was long used as such a standard, and its distribution by age and sex was referred to as the **standard million**. The stationary population of a life table may also be used as the standard for the corresponding observed population.

What we have just sketched is called the **direct** method of standardization. There is also an indirect method. Suppose we know, for the two areas to be compared, their crude rates and their age distributions and, for a standard population, its factor-specific rates (age-specific, etc.). We can then compute another kind of standardized rate. There are still more types of standardization, and this general technique is closely related to **Westergaard's Method of Expected Cases** used in conjunction with multiple classifications.¹¹

The concept of an expected population is a very general one in demography and is very widely used. Suppose we have a population distributed by age and sex and, for each age-sex group, the percent having some other characteristic such as being married, enrolled in school, or in the labor force. Such percentages are sometimes called **participation rates**. Then we assume that the particular participation rate remains constant. By applying such rates to the observed age-sex structure at the later date, we can calculate the expected number of married persons, students, workers, etc., on the assumption of constant participation rates. Comparison of these expected numbers with those actually observed tells us what part of the total change in the number of married persons, etc., was attributable to the change in the participation rates as opposed to change in the age-sex structure of the populations.

Summary demographic factors that are used to control the influence of age include the expectation of life from the life table and gross and net reproduction rates. There are other general statistical methods that hold constant a given factor or that measure the relative effects of two or more factors upon a dependent variable. These include partial correlation, analysis of variance, analysis of covariance, and factor analysis.

In this section we have given a brief overview of the typical approaches in demographic analysis and some of the methods used. This discussion was by no means intended to explain those methods to the extent that the reader would feel equipped to apply them. Detailed definitions and explanations are given in the body of the book. As he reads the following chapters, the reader will sometimes encounter certain terms before he encounters their fullest explanation. We hope then that he will have received a general notion of their meaning and uses and will be able to look ahead, if necessary, by means of the forward references that are given, the table of contents, or the index.

¹¹ Robert M. Woodbury, "Westergaard's Method of Expected Deaths as Applied to the Study of Infant Mortality," *Journal of the American Statistical Association*, 18(139):366-376, September 1922.

SUGGESTED READINGS

- Barclay, George W. *Techniques of Population Analysis*. New York. John Wiley & Sons. 1958. Pp. 1-15.
- Benjamin, Bernard. *Demographic Analysis*. New York, Frederick A. Praeger, Inc., 1968.
- Bogue, Donald J. *Principles of Demography*. New York. John Wiley & Sons. 1969. Pp. 1-31.
- Cox, Peter R. *Demography*. Cambridge, England. Cambridge at the University Press, 3rd ed., 1959. Pp. 1-13.
- Hauser, Philip M., and Duncan, Otis Dudley, eds. *The Study of Population*. Chicago, University of Chicago Press, 1959. Pp. 1-117; 157-166; 321-360.
- Pressat Roland. *Demographic Analysis: Methods, Results, Applications*. Chicago, Aldine-Atherton, 1972. (Translated by Judah Matras.)
- Spiegelman, Mortimer. *Introduction to Demography*. Cambridge, Mass., Harvard University Press, 1968, 2nd ed. Pp. 1-7.
- United Nations. *Handbook of Population Census Methods. Studies in Methods, Series F, No. 5, rev. 1, Vol. I. General Aspects of a Population Census*. 1958. Pp. 3-8.
- . *Handbook of Vital Statistics Methods. Studies in Methods, Series F, No. 7*. April 1955. Pp. 3-18.
- . *Principles and Recommendations for the 1970 Population Censuses*. Statistical Papers, Series M, No. 44. 1967. Pp. 3-9.

This page intentionally left blank

CHAPTER 2

Basic Sources of Statistics

TYPES OF SOURCES

The primary sources of demographic statistics may be regarded as the published reports, unpublished worksheets, tally tapes, etc., that are produced by official or private agencies. The reports may contain texts that tell how the report is organized, how the statistics were obtained, and how accurate the statistics are deemed to be. The texts may also contain descriptive or analytical material based on the statistics. Moreover, graphic devices are often incorporated.

The same statistics may be selectively reproduced or rearranged in secondary sources such as compendia, statistical abstracts, and yearbooks. Additional derived figures may be introduced at this stage. Other secondary sources that present some of these statistics are journals, textbooks, and research reports. Occasionally, a textbook or research report may include demographic statistics based on the unpublished tabulations of an official agency.

Many important kinds of demographic statistics are produced by combining census and vital statistics. Examples are vital rates, life tables, and population estimates and projections. Population estimates may also draw on immigration and emigration statistics and even on such local statistics as those on school enrollment, residential building permits, and registered voters.

Primary Demographic Data and Statistics

A country may have a central statistical office, or there may be separate agencies that take the census and compile the vital statistics. Even when both kinds of statistics emanate from the same agency, they are usually published in separate reports, reflecting the fact that censuses are customarily taken decennially or quinquennially and vital statistics are compiled annually or monthly. Both kinds of statistics may be included in the same compendium, of course, along with other kinds of official statistics.

In some countries, the province or state may have important responsibilities in conducting the census or operating a registration system. Even here, however, the national government customarily issues a summary report or series of reports. At one extreme, the central office simply combines the statistics that were tabulated in the provincial offices. On the other hand, it may receive the original records, copies of them, or abstracts from them and make its own tabulations. In either situation, both national and provincial offices may publish their own reports. The statistics from the different governmental sources may differ with respect to their arrangement, detail, and choice of derived figures. Moreover, what purport to be comparable statistics may differ because of variations in

classification or editing rules and because of processing errors.

The *raison d'être* of a census or sample survey is the statistics that it produces. The registration of vital events or a population register, on the other hand, may be at least as much directed toward the legal and administrative uses of its records. In fact, the compilation and publication of statistics from a population register may be rather minimal, partly because these activities tend to disturb the day-to-day operation of the register. Even though the equivalent of census statistics could be compiled from a population register, the countries concerned still find it necessary to conduct censuses through the usual method of delivering forms to all households simultaneously. This partial duplication of data-gathering is justified as a means of making sure that the register is working properly and of including additional items (characteristics) beyond those recorded in the register. There are often restrictions imposed on the public's access to the individual census or registration records in order to protect the interests of the persons concerned.

Statistics Produced from Combinations of Census and Registration Data

Some examples of the measures based on combinations of population figures from a census with vital statistics were given above. Rates or ratios that have the number of some kind of vital event as the numerator and a population as the denominator are the most obvious type. The denominator may be a subpopulation, such as the number of men 65 to 69 years old—divided into the number of deaths occurring at that age—or the number of women 15 to 44 years old—divided into the total number of births. Moreover, the population may come from a sample survey or a population estimate, which in turn was based partly on past births and deaths.

Products of more complex combinations include current population estimates, life tables, net reproduction rates, estimates of net intercensal migration by age, and estimates of relative completeness of enumeration in successive censuses. The computation of population projections by the component method also starts with a population by age and sex, mortality rates by age, and natality rates by age of mother. Here there may be a series of successive computations in which population and vital statistics are introduced at one or more stages.

All of these illustrative measures can be produced by the combination of statistics. A different approach is to relate the individual records. This approach is that of the matching study. By matching birth certificates, infant death certificates, and babies born in the corresponding period of time and counted in the census, one can estimate both the proportion

of births that were not registered and the proportion of infants that were not counted in the census. Other statistics of demographic value can be obtained by combining the information from the two sources for matched cases so as to be able to obtain a greater number of characteristics for use in the computation of specific vital rates. For example, if educational attainment is recorded on the census schedule but is not called for on the death certificate, a matching study can yield mortality statistics for persons with various levels of educational attainment. When the same characteristic, such as age, is called for on both documents, the matching studies yield measures of the consistency of reporting.

In a country with a population register, matching studies with the census can also be carried out. Again, the resulting statistics could be either of the evaluative type or could represent cross-classifications of the population in terms of a greater number of characteristics than is possible from either source alone.

Secondary Sources

Secondary sources may be either official or unofficial and include a wide variety of textbooks, yearbooks, periodical journals, research reports, gazetteers, and atlases. In this section we mention only a few of the major sources of international population statistics. The United Nations is today the chief producer of secondary demographic statistics for the countries of the world. Some of its relevant publications include the following:

Demographic Yearbook (since 1948) presenting basic population figures from censuses or estimates, and basic vital statistics yearly, and in every issue, featuring a special topic that is presented in more detail (natality statistics, mortality statistics, population distribution, population censuses, ethnic and economic characteristics of population, marriage and divorce statistics, population trends, etc.). This *Yearbook* now includes a cumulative subject-matter index in English and French.

Statistical Yearbook (since 1948) containing fewer demographic series than the foregoing, but also including four tables of manpower statistics.

Monthly Bulletin of Statistics (since 1947) carrying four tables each on demographic topics and on manpower, for the countries of the world.

Statistical Papers, Series A, Population and Vital Statistics Reports (quarterly since 1949).

Statistical Papers, Series K, No. 3, Compendium of Social Statistics: 1967, containing tables on population and vital statistics, health, education, labor force, etc.

Epidemiological and Vital Statistics Report, published monthly since June 1947 by the World Health Organization, a specialized agency of the United Nations.

The recently established Social and Economic Statistics Administration publishes several series of reports on international population statistics. In 1969 the U.S. Bureau of the Census inaugurated Series P-96, *Demographic Reports for Foreign Countries*, under an agreement with the U.S. Agency for International Development, in 1973 ISP-30, *Country Demographic Profiles*, and in 1974 ISWP (International Statistical Programs - World Population). The Foreign Demographic Analysis Division (formerly in the Bureau of the Census but now in the Bureau of Economic Analysis, another part of S.E.S.A.) publishes Series P-90, *International Population Statistics Reports* and Series P-91, *International Population Reports*.

Several recent publications by university scholars have pre-

sented derived demographic measures for large numbers of countries, all based on a common computer program. Noteworthy among these is the book by Keyfitz and Flieger, which presents 10 tables covering 63 countries.¹ The first table presents population, births, and deaths, all by age and sex, and from these are derived life tables, population projections, intrinsic rates for the stable population, and a wide variety of other demographic measures, some of them of an evaluative nature.

Population Index, which is published by the Population Association of America and the Office of Population Research at Princeton University, has appeared since 1937. It has always contained a tabular section, which gives more attention to international than domestic demographic statistics.

United States.—The U.S. Bureau of the Census publishes a number of secondary compilations of statistics; although typically demographic statistics usually account for only a minority of the whole, nonetheless they usually have priority of place. These sources include the following:

Statistical Abstract of the United States, published since 1878.

Supplements to the *Statistical Abstract*, such as *Pocket Data Book. USA*, biennially since 1966–67

Directory of Federal Statistics for Local Areas, 1966, a finding guide, which does not itself present statistics
County and City Data Book, 1944, 1947, 1949, 1952, 1956, 1962, 1967, 1972 (The 1944 edition was called *Cities Supplement*)

Congressional District Data Book (Districts of the 87th, 88th, 93rd, and 94th Congresses) and *Supplements, Redistricted States*.

Historical Statistics of the United States, Colonial Times to 1957

Historical Statistics of the United States, Colonial Times to 1957; Continuation to 1962 and Revisions

From 1850 through 1930, all but one census had a one-volume *Compendium*, or *Abstract*, or both. Several census studies that were not limited to the statistics of a single census are useful for historical research and contain some interesting material that is not available in the decennial reports. These are:

A Century of Population Growth, 1909

Negro Population, 1790–1915, 1918

Negroes in the United States, 1920–1932, 1935

Health, Education and Welfare Indicators was published monthly April 1958 to February 1967 by the Department of Health, Education, and Welfare. It contained vital statistics, statistics on educational attainment and enrollment, etc. The annual supplement to this publication is called *Health, Education, and Welfare Trends*.

A number of other official publications are less appropriately regarded as secondary sources than as sources of important collections of statistics (population estimates and projections, rates, etc.) that were developed expressly for these publications. In this category, for example, are:

U.S. Bureau of the Census, *Vital Statistics Rates in the United States: 1900–1940*, by Forrest E. Linder and Robert D. Grove, 1943.

United States, National Center for Health Statistics,

¹ Nathan Keyfitz and Wilhelm Flieger, *World Population: An Analysis of Vital Data*, Chicago, University of Chicago Press, 1968.

Vital Statistics Rates in the United States: 1940-1960, by Robert D. Grove and Alice M. Hetzel, 1968.

Various reports of the National Resources Board (later the National Resources Committee and the National Resources Planning Board) during the late 1930's and the 1940's.

A bibliography on "Sources for National Demographic Statistics" is contained in an appendix to one of the 1950 Census Monographs.²

CENSUSES AND SURVEYS

The distinction between a census and a survey is far from clear-cut. At one extreme, a complete national canvass of the population would always be recognized as a census. At the other extreme, a canvass of selected households in a village to describe their living conditions would probably be regarded as a social survey. But neither the mere use of sampling nor the size of the geographic area provides a universally recognized criterion. Most national censuses do aim at a complete count or listing of the inhabitants; however, in many modern censuses, sampling is used at one or more stages—in the collection of some of the questions, in the units to be tabulated, etc. When the U.S. Bureau of the Census, at the request and expense of the local government, takes a canvass of the population of a village with 100 inhabitants, it has no hesitation in calling the operation "a special census." The Current Population Survey in the United States and other national sample surveys with equally low rates of sampling are sometimes referred to as "microcensuses." That term might more appropriately apply to very large samplings that provide statistics in considerable geographic detail, such as the 10-percent sample census of Great Britain in 1966. Here, however, we will arbitrarily designate complete official canvasses of an area as censuses but canvasses or mailings made on only a sample basis, as surveys. This convention tends to be the usual practice. Finally, the main objective of a population census is the determination of the number of inhabitants. This is rarely true of a sample survey.

The definition used by the United Nations is as follows: "A census of population may be defined as the total process of collecting, compiling and publishing demographic, economic and social data pertaining, at a specified time or times, to all persons in a country or delimited territory."³

The typical scope of a census or demographic survey is the size, distribution, and characteristics of the population. In countries without adequate registration of vital events, however, a population census or survey may include questions about births or deaths of household members in the period (usually the year) preceding the census. Moreover, even when vital statistics of good quality exist, the census or survey may include questions on fertility (children ever born, children still living, date of birth of each child, etc.) because the distribution of women by number of children ever born and by interval between successive births cannot be discovered from birth certificates. Recently, in the United States, the National Center for Health Statistics has mailed questionnaires to mothers identified from a sample of birth statistics for a given year in order to obtain additional particulars concerning the pregnancy and the characteristics of the parents.

² Conrad and Irene B. Taeuber, *The Changing Population of the United States*, Philadelphia, John Wiley & Sons, 1958, pp. 327-334.

³ United Nations, *Principles and Recommendations for National Population Censuses*, Statistical Papers, Series M, No. 27, 1958, p. 3.

International View

History of Census Taking.—Census-taking had its beginning in ancient times—in Egypt, Babylonia, China, Palestine, and Rome. Few of the results have survived, however. Moreover, the counts were undertaken to determine fiscal, labor, and military obligations and were usually limited to heads of households, males of military age, taxpayers, or adult citizens. Women and children were seldom counted.

The first of two enumerations mentioned in the Bible is assigned to the time of the Exodus, 1491 B.C. The second was taken at the order of King David in 1017 B.C. The Roman Census, taken quinquennially, lasted for about 800 years. Citizens and their property were inventoried for fiscal and military purposes. This enumeration was extended to the entire Roman Empire in 5 B.C.

The Domesday inquest ordered by William I of England in 1086 covered landholders and their holdings. The Middle Ages, however, were a period of retrogression in census-taking throughout Europe, North Africa, and the Near East.

As Kingsley Davis points out, it is hard to say when the first census in the modern sense was undertaken since censuses were long deficient in some important respects.⁴ Nouvelle France (later Quebec) and Acadia (later Nova Scotia) had 16 enumerations between 1665 and 1754. In Europe, Sweden's census of 1749 is sometimes regarded as the first, but those in some of the Italian principalities (Naples, Sicily, etc.) go back into the 17th century; again, the issue hinges upon one's criteria for a census. The clergy in the established Lutheran Church of Sweden had been compiling lists of parishioners for some years prior to the time when they were required to take annual (or later triennial) inventories. Whereas in Scandinavia this ecclesiastical function evolved into population registers and occasional censuses, the parish registers of baptisms, marriages, and burials in England evolved into a vital statistics system, as will be described later in this chapter. Carr-Saunders, Davis, Lorimer, and Wolfenden mention 18th century censuses in some of the Italian and German States (Sardinia, Parma, Tuscany, Prussia), in Iceland and Denmark, and in British North America.

There were 25 colonial enumerations within what is now the United States, beginning with a census of Virginia in 1624-1625. The first census of the United States (1790) was followed in 1801 by censuses of England and France. The American experience is described in more detail in later sections of this chapter.

Turning again to the Far East, we may cite Irene Taeuber's generalization that the great demographic tradition of that region is that of population registration.⁵ This practice began in ancient China with the major function being the control of the population at the local level. Occasionally, the records could be summarized to successively higher levels to yield population totals and vital statistics. The family may be viewed as the basic social unit in this system of record-keeping. In theory, a continuous population register should have resulted, but in practice statistical controls were usually relatively weak and the compilations were either never made or they tended to languish in inaccessible archives.

⁴ Kingsley Davis, "Census", *Encyclopaedia Britannica*, 1966, Vol. 5, pp. 167-170.

⁵ Irene B. Taeuber, "Demographic Research in the Pacific Area," in *The Study of Population*, Philip M. Hauser and Otis Dudley Duncan, eds., Chicago, University of Chicago Press, 1959, p. 261.

The Chinese registration system diffused gradually to nearby lands. Until the present century, the statistics from this source were intended to cover only part of the total population and contained gross inaccuracies. Japan's adaptation of the Chinese system resulted in the *koseki*, or household registers. These had been in existence for more than a thousand years when, in 1721, an edict was issued that the numbers registered should be reported. Such compilations were made at six-year intervals down to 1852 although certain relatively small classes of the population were omitted. Thus, this use of the population register parallels that in Scandinavia in the same centuries. The first census of Japan by means of a canvass of households was not attempted until 1920; it presumably resulted from the adoption of the western practice that was then more than a century old. Fairly frequent compilations of populations and households were made in Korea during the Yi dynasty; the earliest was in 1395.⁶

In summary, then, the evolution of the modern census was a gradual one. The tradition of household canvasses or population registration often had to continue for a long time before the combination of public confidence, administrative experience, and technology could produce counts that met modern standards of completeness, accuracy, and simultaneity. Beginning with objectives of determining military, tax, and labor obligations, censuses in the 19th century changed their scope to meet other administrative needs as well as the needs of business, labor, education, and academic research. New items included on the census questionnaire reflected new problems confronting state and society.

Of continuing interest are the periodic national sample surveys of households that have been established in a number of countries, somewhat on the model of the Current Population Survey in the United States. These may be conducted monthly, quarterly, or only annually; in some countries, they have been discontinued after one or two rounds because of financial or other problems. Usually the focus is employment and unemployment or consumer expenditures rather than demographic information, but demographic items may be added in particular rounds.

Censuses and Surveys in the United States

Both population censuses and national sample surveys developed relatively early in the United States of America, the former much earlier than the latter, of course. Moreover, the decennial census and the Current Population Survey have never missed a scheduled round since their respective inception. Both have also tended to grow in the range of topics covered, in sophistication of procedures, in accuracy of results, and in the volume of statistics made available to the public.

Decennial Censuses.—The census of population has been taken regularly every 10 years since 1790 and was one of the first to be started in modern times. At least as early as the 1940's, there have been demands for a quinquennial census of population, which is the frequency in a fair number of other countries; but, so far, no mid-decade census has been approved by the Congress.

Evolution of the population census schedule.—The area covered by the census included the advancing frontier within continental United States. Each outlying territory and posses-

sion has been included also, but the statistics for these areas are mostly to be found in separate reports.⁷ Beginning as a simple list of heads of households with a count of members in five demographic and social categories, the population census has developed into an inventory of many of the personal, social, and economic characteristics of the American people.

A comprehensive account of the content of the population schedule at each census through the 11th in 1890 is included in *The History and Growth of the United States Census* by Wright and Hunt.⁸ A list of items included in each census through 1960 is given in table 2-1.

The changing content of the population schedule has reflected the rise and wane of different public problems. Since the Constitution provided that representatives and direct taxes should be apportioned among the States "according to their respective numbers, which shall be determined by adding to the whole number of free persons . . . excluding Indians not taxed, three-fifths of all other persons," early attention was directed to free Negroes and slaves. Indians were not shown separately until 1860, and most were omitted until 1890. Increasing detail was obtained on age and color; but it was not until 1850 that single years of age and sex were reported for whites, blacks, and mulattoes. Interest in immigration was first reflected on the census schedule in 1820 in an item on "foreigners not naturalized"; but the peak of attention occurred in 1920 when there were questions on country of birth, country of birth of parents, citizenship, mother tongue, ability to speak English, year of immigration, and year of naturalization of the foreign born.

Internal migration did not become a subject of inquiry until 1850 when State of birth of the native population was asked for, and it was not until 1940 that questions were carried on residence at a fixed date in the past. The first item on economic activity was obtained in 1820 ("number of persons engaged in agriculture," "number of persons engaged in commerce," "number of persons engaged in manufactures"). The items on economic characteristics have increased in number and detail; they have included some on wealth and, more recently, income. Welfare interests in the defective, delinquent, and dependent were first recognized on the 1840 schedule. Such inquiries were expanded over the course of many decades and did not completely disappear from the main schedule until 1920. In 1970, again, an item on disability has been introduced. Education and the veteran both received their first recognition in 1840 also. Attempts were made to collect vital statistics through the census before a national registration system was begun. Interest in public health led to a special schedule on mortality as early as 1850; but questions on marriages and births were carried on the population schedule itself, beginning in 1860 and 1870, respectively. A few questions on real property owned and on housing were included, beginning in 1850; but, with the advent of the concurrent housing census in 1940, such items were dropped from the population schedule. The topic of the journey to work ("commuting") did not receive attention until 1960 when questions on place of work and means of transportation were included. New items added in the 1970 census included major activity

⁷ Alaska and Hawaii, previously the subjects of separate reports, were included in the national totals in the 1960 census, after they had become States.

⁸ Carroll D. Wright and William C. Hunt, *The History and Growth of the United States Census*, Washington, Government Printing Office, 1900.

⁶ Kap Suk Koh, *The Fertility of Korean Women*, Seoul, The Institute of Population Problems, 1966, p. 6. (A volume in the 1960 monograph series.)

Table 2-1. —Questions Included in Each Population Census in the United States: 1790 to 1970**CENSUS OF 1970**

Information obtained from all persons: Address; name; relationship to head of household; sex; race; age; month and year of birth; marital status; if American Indian, name of tribe.

Information obtained from 20-percent sample: Whether residence is on farm;¹ place of birth; educational attainment; for women, number of children ever born; employment status; hours worked in week preceding enumeration; year last worked; industry, occupation, and class of worker; state or country of residence 5 years ago; activity 5 years ago; weeks worked last year; earnings last year from wages and salary, from self-employment; other income last year.

Information obtained from 15-percent sample: Country of birth of parents; county, and city or town of residence 5 years ago (and whether in city limits or outside); length of residence at present address; language spoken in childhood home; school or college attendance, and whether public, parochial, or other private school; veteran status; place of work--street address, which city or town (and whether in city limits or outside), county, State, ZIP code; means of transportation to work.

Information obtained from 5-percent sample: Whether of Spanish descent; citizenship; year of immigration; whether married more than once and date of first marriage; whether first marriage ended because of death of spouse; vocational training; (for persons of working age) presence and duration of disability; industry, occupation, and class of worker 5 years ago.

Supplemental schedule for Americans overseas.

CENSUS OF 1960

Information obtained from all persons: Address; name; relationship to head of household; sex; race; month and year of birth; marital status.

Information obtained from 25-percent sample: Whether residence is on farm;¹ place of birth; if foreign born, language spoken in home before coming to U.S.; country of birth of parents; length of residence at present address; State, county, and city or town of residence 5 years ago; educational attainment; school or college attendance, and whether public or private school; whether married more than once and date of first marriage; for women ever married, number of children ever born; employment status; hours worked in week preceding enumeration; year last worked; occupation, industry, and class of worker; place of work--which city or town (and whether in city limits or outside), county, State; means of transportation to work; weeks worked last year; earnings last year from wages and salary, from self-employment; other income last year; veteran status.

Supplemental schedule for Americans overseas.

CENSUS OF 1950

Information obtained from all persons: Address; whether house is on farm; name; relationship to head of household; race; sex; age; marital status; place of birth; if foreign born, whether naturalized; employment status; hours worked during week preceding enumeration; occupation, industry, and class of worker.

Information obtained from 20-percent sample: Whether living in same house a year ago; whether living on a farm a year ago; if not in same house, county and State of residence a year ago; country of birth of parents; educational attainment; school attendance; if looking for work, number of weeks; weeks worked last year; for each person and each family, earnings last year from wages and salary, from self-employment; other income last year; veteran status.

CENSUS OF 1950--Continued

Information obtained from 3-1/3 percent sample: For persons who worked last year but not in current labor force: occupation, industry, and class of worker on last job; if ever married, whether married more than once; duration of present marital status; for women ever married, number of children ever born.

Supplemental schedules for persons on Indian reservations; infants born in the first 3 months of 1950; Americans overseas.

CENSUS OF 1940

Information obtained from all persons: Address; home owned or rented; value or monthly rental; whether on a farm; name; relationship to head of household; sex; race; age; marital status; school or college attendance; educational attainment; place of birth; citizenship of foreign born; county, State, and town or village of residence 5 years ago and whether on a farm; employment status; if at work, whether in private or nonemergency government work, or in public emergency work (WPA, NYA, CCC, etc.); if in private or non-emergency government work, number of hours worked during week of March 24-30; if seeking work or on public emergency work, duration of unemployment; occupation, industry, and class of worker; number of weeks worked last year; wage or salary income last year and whether received other income of \$50 or more.

Information from 5-percent sample: Place of birth of parents; language spoken in home in earliest childhood; veteran status, which war or period of service; whether wife or widow of veteran; whether a child under 18 of a veteran and, if so, whether father is living; whether has Social Security number, and if so, whether deductions were made from all or part of wages or salary; occupation, industry, and class of worker; of women ever married--whether married more than once, age at first marriage, and number of children ever born.

Supplemental schedule for infants born during the four months preceding the census.

(All inquiries in censuses from 1790 through 1930 were asked of all applicable persons.)

CENSUS OF 1930

Address; name; relationship to head of family; sex; race; age; marital status; age at first marriage; home owned or rented; value or monthly rental; radio set; whether family lives on a farm; school attendance; literacy; place of birth of person and parents; if foreign born, language spoken in home before coming to U.S.; year of immigration; naturalization; ability to speak English; occupation, industry, and class of worker; whether at work previous day (or last regular working day); veteran status; for Indians, whether of full or mixed blood, and tribal affiliation.

Supplemental schedules for gainful workers not at work on the day preceding the enumeration; blind and deaf-mutes.

CENSUS OF 1920

Address; name; relationship to head of family; sex; race; age; marital status; year of immigration to the U.S.; whether naturalized and year of naturalization; school attendance; literacy; place of birth of person and parents; mother tongue of foreign born; ability to speak English; occupation, industry, and class of worker.

Supplemental schedules for the blind and for the deaf.

¹ The question on farm residence was carried on the housing schedule but was available for cross-classification with population characteristics.

Table 2-1 —Questions Included in Each Population Census in the United States: 1790 to 1970—Con.

CENSUS OF 1910

Address; name; relationship to head of family; sex; race; age; marital status; number of years of present marriage; for women, number of children born and number now living; place of birth and mother tongue of person and parents; if foreign born, year of immigration, whether naturalized or alien, and whether able to speak English, or if not, language spoken; occupation, industry, and class of worker; if an employee, whether out of work on census day, and number of weeks out of work during preceding year; literacy; school attendance; home owned or rented; if owned, whether mortgaged; whether farm or house; whether a survivor of Union or Confederate Army or Navy; whether blind or deaf and dumb.

Supplemental schedules for the Indian population; blind; deaf; feeble-minded in institutions; insane in hospitals; paupers in almshouses; prisoners and juvenile delinquents in institutions.

CENSUS OF 1900

Address; name; relationship to head of family; sex; race; age; month and year of birth; marital status; number of years married; for women, number of children born and number now living; place of birth of person and parents; if foreign born, year of immigration to the U.S., number of years in the U.S., and whether naturalized; occupation; months not employed; months attended school during census year; literacy; ability to speak English.

Supplemental schedules for the blind and for the deaf.

CENSUS OF 1890

Address; name; relationship to head of family; race; sex; age; marital status; number of families in house; number of persons in house; number of persons in family; whether a soldier, sailor or marine during Civil War (Union or Confederate) or widow of such person; whether married during census year; for women, number of children born, and number now living; place of birth of person and parents; if foreign born, number of years in the U.S., whether naturalized or whether naturalization papers* had been taken out; profession, trade, or occupation; months unemployed during census year; months attended school during census year; literacy; whether able to speak English, and if not, language or dialect spoken; whether suffering from acute or chronic disease, with name of disease and length of time afflicted; whether defective in mind, sight, hearing, or speech, or whether crippled, maimed, or deformed, with name of defect; whether a prisoner, convict, homeless child, or pauper; home rented or owned by head or member of family; if owned by head or member, whether mortgaged; if head of family a farmer, whether farm rented or owned by him or member of his family; if owned, whether mortgaged; if mortgaged, post office address of owner.

Supplemental schedules for the Indian population; for persons who died during the year; insane; feeble-minded and idiots; deaf; blind; diseased and physically defective; inmates of benevolent institutions; prisoners; paupers and indigent persons; surviving soldiers, sailors, and marines, and widows of such; inmates of soldiers' homes.

CENSUS OF 1880

Address; name; relationship to head of family; sex; race; age; marital status; month of birth if born within the census year; married within the year; occupation; number of months unemployed during year; sickness or temporary disability; whether blind, deaf and dumb, idiotic, insane, maimed, crippled, bedridden, or otherwise disabled; school attendance; literacy; place of birth of person and parents.

CENSUS OF 1880--Continued

Supplemental schedules for the Indian population; for persons who died during the year; insane; idiots; deaf-mutes; blind; homeless children; prisoners; paupers and indigent persons.

CENSUS OF 1870

Name; age; sex; race; occupation; value of real estate; value of personal estate; place of birth; whether parents were foreign born; month of birth if born within the year; month of marriage if married within the year; school attendance; literacy; whether deaf and dumb, blind, insane, or idiotic; male citizens 21 and over, and number of such persons denied the right to vote for other than rebellion.

Supplemental schedules for persons who died during the year; paupers; prisoners.

CENSUS OF 1860

Name; age; sex; race; value of real estate; value of personal estate; occupation; place of birth; whether married within the year; school attendance; literacy; whether deaf and dumb, blind, insane, idiotic, pauper, or convict.

Supplemental schedules for slaves; public paupers and criminals; persons who died during the year.

CENSUS OF 1850

Name; age; sex; race; whether deaf and dumb, blind, insane, or idiotic; value of real estate; occupation; place of birth; whether married within the year; school attendance; literacy; whether a pauper or convict.

Supplemental schedules for slaves; public paupers and criminals; persons who died during the year.

CENSUS OF 1840

Name of head of family; age; sex; race; slaves; number of deaf and dumb; number of blind; number of insane and idiotic and whether in public or private charge; number of persons in each family employed in each of six classes of industry and one of occupation; literacy; pensioners for Revolutionary or military service.

CENSUS OF 1830

Name of head of family; age; sex; race; slaves; deaf and dumb; blind; foreigners not naturalized.

CENSUS OF 1820

Name of head of family; age; sex; race; foreigners not naturalized; slaves; industry (agriculture, commerce, and manufactures).

CENSUS OF 1810

Name of head of family; if white, age and sex; race; slaves.

CENSUS OF 1800

Name of head of family; if white, age and sex; race; slaves.

CENSUS OF 1790

Name of head of family; free white males of 16 years and up, free white males under 16; free white females; slaves; other persons.

and occupation five years ago, vocational training, and additional particulars designed to improve the classification of occupation.

Publications.—There is a cumulative list of census publications covering those from the beginning through 1945.⁹ Subsequent years are covered by annual catalogues, which are preceded by monthly and quarterly issues.¹⁰

The results of the 1970 Census of Population and Housing are available in the form of printed reports, microfiche copies of the printed reports, computer summary tapes, computer printouts, and microfilm. The population census statistics are published in two basic volumes:

Volume I, *Characteristics of the Population*, consists of separate reports for the United States, each of the 50 states, the District of Columbia, Puerto Rico, Guam, Virgin Islands of the United States, American Samoa, Canal Zone, and Trust Territory of the Pacific Islands. For each of the 58 areas, the data were first issued in four separate paperbound chapters designated as PC (1)-A, B, C, and D. For the outlying areas other than Puerto Rico, all the data on characteristics of the population are included in chapter B. A brief description of the content of each of these four chapters is as follows:

PC (1)-A, *Number of Inhabitants*. Final official population counts are presented for States, counties (by urban-rural residence), Standard Metropolitan Statistical Areas (SMSA's), urbanized areas, minor civil divisions, census country divisions, all incorporated places, and unincorporated places of 1,000 inhabitants or more.

PC (1)-B, *General Population Characteristics*. These reports contain statistics on age, sex, race, marital status, and relationship to head of household for States, counties (by urban-rural residence), Standard Metropolitan Statistical Areas, urbanized areas, minor civil divisions, census country divisions, and places of 1,000 inhabitants or more.

PC (1)-C, *General Social and Economic Characteristics*. These reports focus on the population subjects collected on a sample basis (see Table 2-1). Each subject is shown for some or all of the following areas: States (by urban, rural-nonfarm, and rural-farm residence), counties, SMSA's, urbanized areas, and places of 2,500 inhabitants or more.

PC(1)-D, *Detailed Characteristics*. These reports cover most of the population subjects collected on a sample basis, presenting the data in considerable detail or cross-classified by age, race, and other characteristics. Each subject is shown for some or all of the following areas: States (by urban, rural-nonfarm, and rural-farm residence), SMSA's, and large cities.

Volume II, *Special Reports*. There are 39 reports scheduled in this volume, each of which concentrates on a particular subject. Detailed information and cross-relationships are generally provided on a national and regional level; in a few reports, data for States or standard metropolitan statistical areas are also shown. Among the characteristics covered are national origin and race, fertility, families, marital status, migration, education, employment, unemployment, occupation, industry, and income. There is also a report on the geographic distribution and characteristics of the institutional population.

A number of reports have also been published that combine population items with those of the housing census. There are three

series of these reports based on the 1970 enumeration:

Series PHC(1) *Census Tract Reports*. One report for each SMSA, showing data for most of the population and housing subjects included in the 1970 Census. Some tables are based on the 100-percent tabulation, others on a sample tabulation.

Series PHC (2) *General Demographic Trends for Metropolitan Areas, 1960 to 1970*. This series consists of one report for each State and the District of Columbia, as well as a national summary report, presenting statistics for the State and for SMSA's and their central cities and constituent counties. Comparative 1960 and 1970 data are shown on population counts by age and race.

Series PHC (3) *Employment Profiles of Selected Low-Income Areas*. This series consists of 75 reports, each presenting statistics on the social and economic characteristics of the residents of a particular low-income area, as well as a United States summary. The data, derived from sample surveys conducted during late 1970 and early 1971, relate to low-income neighborhoods in 51 cities and seven rural areas. Each report provides statistics on employment, education, vocational training, availability for work, job history, and income.

Another type of publication is represented by the occasional census monograph dealing with detailed analysis of specific topics. For example, the following monographs were prepared as part of the 1960 publication program:

Herman P. Miller	<i>Income Distribution in the United States</i>
John K. Folger and Charles B. Nam	<i>Education of the American Population</i>
Dale E. Hathaway, J. Allan Beegle, and W. Keith Bryant	<i>People of Rural America</i>
Daniel O. Price	<i>Changing Characteristics of the Negro Population</i>
Irene B. Taeuber and Conrad Taeuber	<i>People of the United States in the 20th Century</i>

There are, in addition, a number of methodological publications that are either part of a decennial census program or that are included in one of the series of technical reports. A prominent example of the former is the *Procedural History* of the 1970 Census of Population and Housing.

Available unpublished data.—Unpublished data may be classified according to the stage of processing. Coded schedules are not available directly because of confidentiality considerations. Computer tapes containing individual records are ordinarily subject to the same restrictions, but an important breakthrough occurred in the 1960 census when, without violating confidentiality, 1/1,000 and 1/10,000 samples of the records were marketed at a subsidized price for use in training and special tabulations.¹¹ Thus, demographers in universities could produce additional cross-classifications of the data as part of their own research, and students were able to learn more about how census data are edited and tabulated. Special tabulations may be contracted for from the record tapes, or more cheaply, from the tally tapes (statistics in the form of tallies that were made in the process of preparing the computer printouts of tables). The provisions for use of unpublished data were greatly expanded in the 1970 Census. See Bureau of the Census, *Catalog 1974*, Part II, "Data Files and Special Tabulations."

Special Censuses Conducted by the Federal Government.—At the request and expense of local governments, many

⁹ U.S. Bureau of the Census, *Catalogue of the United States Census Publications: 1790-1945*, by Henry J. Dubester, 1950.

¹⁰ Bureau of the Census Catalog, [date] (published as *Census Publications—Catalog and Subject Guide* until 1952). This annual catalogue lists not only the official publications (with a subject index) but also the "outside" publications of staff members. Furthermore, it contains a section on "Data files and unpublished materials."

¹¹ U.S. Censuses of Population and Housing: 1960, *One-in-a-Thousand Sample, Description and Technical Documentation*, 1964.

complete enumerations have been taken at off-census dates by the Bureau of the Census.

The local government almost invariably chooses to collect only the minimum of items of information—name, relationship to the head of the household, sex, age, and race. (The census is usually taken to obtain a certified count for some fiscal purpose.) Most of the special censuses are requested for cities; but counties, minor civil divisions, and annexations have also been covered and occasionally even an entire State. Results for individual areas are published in *Current Population Reports*, Series P-28 (formerly Series P-SC), *Special Censuses*, for areas with 50,000 inhabitants or more. Statistics are shown by age, sex, and color for all political areas of this size; if the area is tracted, these statistics are also published by tracts. Quarterly and annual summaries present abridged data for all the special censuses. Furthermore, all the special censuses taken in the preceding decade have been listed in the U.S. Summaries of the 1950, 1960, and 1970 Censuses.¹²

State and Local Censuses.—The trend in the number of censuses taken by States and localities has been quite unlike that in the number of special censuses taken by the federal government. In or around 1905, 15 States took their own census; in 1915, 15; in 1925, 9; in or around 1935, 6; in 1945, 4; and in 1955 and 1965, only 2. The last survivors are Kansas and Massachusetts.

Censuses conducted by cities and other local governments are not currently and never have been very plentiful. State agencies in three Pacific Coast States that have had official responsibility for preparing population estimates have implemented that function partly by conducting censuses, particularly of small incorporated places. These agencies are the Washington State Census Board, the Oregon Center for Population Research and Census, and the Financial and Population Research Section of the California State Department of Finance. They began conducting such special censuses, on the request of local governments, in 1943, 1955, and 1960, respectively.

Sample Surveys in the United States.—Beginning in the second third of the 20th century, sample surveys of many kinds have had a tremendous vogue in the United States. Demography has shared in this movement.

Surveys conducted by the Bureau of the Census.—The sample check on the 1937 unemployment census was a methodological forerunner of the many sample surveys that have been conducted by the Federal government, but demographic data in the narrow sense were not collected in it. The procedures inaugurated in the 1940 census of asking selected questions of a representative sample of persons and of tabulating for only a sample of persons certain data that could have been tabulated for all persons should be distinguished from surveys where only a sample of households is interviewed. The ancestry of the sample surveys of the Bureau of the Census traces rather to the Monthly Report on the Labor Force of the Works Progress Administration.

This survey was designed to provide an official national series on the labor force, employment, and unemployment. The total population and its distributions by age, color, and

sex are not obtained from this survey but are independent estimates based on other demographic data, and the survey results are inflated to these totals. After the liquidation of the Works Progress Administration in August 1942, this survey was transferred to the Census Bureau. In June 1947, it was renamed the Current Population Survey. This sample survey provides many demographic data, some from special tabulations of the contents of the regular monthly forms and others by means of the supplementary questions that are carried from time to time. The first of the special tabulations was for marital status in February 1944, and the first supplementary demographic question dealt with migration status as of March 1945.¹³

There are now regular annual series on marital status and family status, households and families, metropolitan and non-metropolitan residence, population mobility, farm population, school enrollment, educational attainment, Negro population, and consumer income.¹⁴ In addition, there are supplements about every two years on fertility and voting behavior. Furthermore, there are occasional supplements on other demographic topics. Many of these supplements are added at the request of another Federal agency. Here the contracting agency usually publishes the data, but sometimes data from the same supplement are also published by the Bureau of the Census. These special topics have included reasons for moving, religion, educational change in a generation, literacy, occupational mobility, financial resources and major expenditures of the elderly, outdoor recreational activity, and poverty.¹⁵

Prior to July 1, 1959, the Census Bureau's *Current Population Reports* also included Series P-50, *Labor Force*, and Series P-57, *Labor Force*, "Monthly Report on the Labor Force." P-50 comprised an annual report on the labor force and reports on special supplements relating to employment and unemployment as well as projections of the labor force. Although such data are still collected through the Current Population Survey, they are now published by the Bureau of Labor Statistics—the equivalent of the P-50 reports as articles in its *Monthly Labor Review* and the erstwhile P-57 statistics in *Employment and Earnings*.

The Bureau of the Census also operates other national sample surveys of households. For example, statistics based on its National Health Survey are published by the U.S. Public Health Service.

Surveys conducted by other federal agencies.—Since most other federal agencies do not have their own national field

¹³ U.S. Bureau of the Census, *Population—Special Reports*, Series P-S, No. 1, "Marital Status of the Civilian Population: February, 1944," June 20, 1944; idem, *Population—Special Reports*, Series P-S, No. 5, "Civilian Migration in the United States: December, 1941, to March, 1945," September 2, 1945.

¹⁴ For convenience, all the current publication series of the Population Division, U.S. Bureau of the Census, will be listed at this point, even though some of them are not based, or are not entirely based, on current surveys.

CURRENT POPULATION REPORTS

P-20 *Population Characteristics*

P-23 *Special Studies*

P-25 *Population Estimates and Projections*

P-26 *Federal-State Cooperative Program for Population Estimates*

P-27 (also Series Census-ERS), *Farm Population* (Jointly with Department of Agriculture)

P-28 *Special Censuses*

P-60 *Consumer Income*

¹² U.S. Census of Population: 1950, Vol. I, *Number of Inhabitants*, Washington, 1952, table 32, pp. 1-80 to 1-84; and U.S. Census of Population: 1960, Vol. I, *Characteristics of the Population, Part A, Number of Inhabitants*, 1961, table 40, pp. 1-127 to 1-138; U.S. Census of Population: 1970, Vol. I, *Characteristics of the Population, Part A: Number of Inhabitants*, 1972, table 45, pp. 1-223 to 1-245.

¹⁵ Gertrude Bancroft, "Special Uses of the Current Population Survey Mechanism," *Estadística*, 12(43):198-205, June 1954; Daniel B. Levine and Charles B. Nam, "The Current Population Survey: Methods, Content, and Sociological Uses," *American Sociological Review* 27(4):585-590, August 1962.

organizations for conducting household surveys, they tend to turn to the Bureau of the Census as the collecting agency when social or economic data are needed for their research or administrative programs. An exception that is of interest here is the Department of Agriculture. From a mail-questionnaire survey conducted for the Economic Research Service by the Department's Statistical Reporting Service, data are collected annually about the number of persons moving to and from a sample of farms, from which the Economic Research Service constructs estimates of the farm population by geographic divisions.¹⁶ These estimates are adjusted to the national estimates of farm population, which are based on the Census Bureau's Current Population Survey and are published jointly by the two agencies in Series P-27, *Farm Population*, as mentioned above. Furthermore, the Current Population Survey includes every December, for the Department of Agriculture, supplementary questions on hired farm workers.¹⁷ A feature of particular demographic interest from this source is the data on migratory agricultural workers, which are sometimes also the subject of special reports. Other topics concerning farm population or agricultural manpower are covered by occasional supplements to the Current Population Survey.

The Office of Education and the American Association of Collegiate Registrars and Admissions Officers have conducted several surveys on the migration of college students.¹⁸ The information was collected from the colleges and universities themselves on the basis of their registration records, and not from the individual students.

A rather different kind of survey using a probability sample has been conducted by the National Center for Health Statistics. It is called the National Natality Survey. Beginning with the births of 1963, a sample of about 3,700 families in which a birth had occurred was sent a questionnaire that contained a number of items concerning the mother's and father's social and economic characteristics. (This is information that was not called for on the standard birth certificate.)¹⁹

Surveys by State and local governments and by private agencies.—Sample surveys of population conducted by State or local governments were very rare until after 1960.

In recent years such surveys have proliferated, partly in connection with programs in the fields of manpower, unemployment, health, education, and welfare. Federal grants have been made in large numbers to State and city agencies, and especially to universities, for surveys and research. Few of the surveys are concerned directly with population but may include "background" questions on the demographic characteristics of the persons in the sample.

There are a great many private survey organizations in the United States, and some of them conduct national sample surveys in which demographic data are collected. In fact, much of what

we know about the determinants of differential fertility behavior in this country today has been derived from one or another sample surveys. The Survey Research Center at the University of Michigan includes among its many programs the Detroit Area Survey, which was established in 1951. These surveys frequently include questions on migration, family income, and other population topics.²⁰ In the past it has also conducted two surveys (in 1955 and 1960) underlying the study on the Growth of American Families.²¹ Likewise, the National Opinion Research Center at the University of Chicago and National Analysts have been data-collectors for important demographic studies. The latter, for example, carried out the field work for Princeton's fertility study, popularly identified as the Third Child Study, and the fieldwork for the 1973 round of the Family Growth Survey for the National Center for Health Statistics; and it has been the major data collecting agent for the National Fertility Studies of 1965, 1970, and 1975.²² The classic Indianapolis Study, which has engaged the research efforts of so many analysts of human fertility at so many different institutions over the course of a generation, was, however, carried out in the field by an *ad hoc* survey organization.²³ Demographic surveys conducted by universities in particular communities are legion, and their number grows at an accelerated pace. Moreover, some teams of demographers and other social scientists have conducted population surveys abroad or have given the essential supervision of technical assistance to such surveys; and, of course, many foreign surveys have been financed in whole or part by American foundations. Demographic survey data from a number of countries are being included in a network of archives, the Social Science Data Archives.²⁴

Public opinion surveys sometimes include attitudinal questions regarding population problems. For example, the Gallup Poll (American Institute of Public Opinion) has inquired about what was considered the ideal family size in 10 polls from 1936 to 1966.

REGISTRATION SYSTEMS

This major section of the chapter treats not only the statistics that are produced by registration of vital events and the recording of arrivals and departures at international boundaries but also universal population registers and registers of parts of the population (workers employed in jobs covered by social insurance plans, aliens, members of the armed forces, voters, etc.). In most cases one's name is inscribed in a register as the result of the occurrence of a certain event (birth, entering the country, attaining military age, entering gainful

¹⁶ For example: U.S. Department of Agriculture, Economic Research Service, *Farm Population—Estimates for 1968*, ERS 427, Sept. 1969.

¹⁷ The latest regular report from this source is: U.S. Department of Agriculture, Economic Research Service, *The Hired Farm Working Force of 1967*, by Robert C. McElroy, Agricultural Economic Report No. 148, September 1968.

¹⁸ The latest report is: U.S. Office of Education, *Residence and Migration of College Students, Fall 1963: State and Regional Data*, by Mabel C. Rice and Paul L. Mason, OE-54033, Circular No. 783, 1965; There has been a long series of such surveys under various auspices. See: H. Theodore Groat, "Internal Migration Patterns of a Population Subgroup: College Students, 1887-1958," *American Journal of Sociology*, 69(4):383-394, January 1964.

¹⁹ National Center for Health Statistics, *Vital and Health Statistics*, Series 22, No. 3, *Methods and Response Characteristics: National Natality Survey, United States, 1963*, by Gooloo S. Wunderlich, September 1966.

²⁰ University of Michigan, Survey Research Center, *The Detroit Area Study—A Bibliography of Material Based on Detroit Area Study Research: September 1, 1951—December 1, 1960*, Ann Arbor, Michigan, Detroit Area Survey, 1960; University of Michigan, Institute for Social Research, *List of Publications: 1961-1965*, compiled by William Goodrich Jones, Ann Arbor, Michigan, Survey Research Center, 1965.

²¹ Ronald Freedman, Pascal K. Whelpton, Arthur A. Campbell, *Family Planning, Sterility, and Population Growth*, New York, McGraw-Hill Book Co., 1959; and Pascal K. Whelpton, Arthur A. Campbell, and John E. Patterson, *Fertility and Family Planning in the United States*, Princeton, N.J., Princeton University Press, 1966.

²² Charles F. Westoff, et al., *Family Growth in Metropolitan America*, Princeton, N.J., Princeton University Press, 1961; idem, *The Third Child*, Princeton University Press, 1963. As of this writing, only the results of the 1965 National Fertility Survey had been published. See Norman B. Ryder and Charles F. Westoff, *Reproduction in the United States: 1965*, Princeton, N.J., Princeton University Press, 1971.

²³ P. K. Whelpton and Clyde V. Kiser, eds., *Social and Psychological Factors Affecting Fertility*, 5 vols., New York, Milbank Memorial Fund, 1946 to 1958.

²⁴ Ralph L. Bisco, "Social Science Data Archives: A Review of Developments," *American Political Science Review*, 60(1):93-109, March 1966.