INTRODUCTION

A. ORIGINS AND SCOPE OF *Manual X*

Manual X was originally conceived as an aid to population experts, mainly from developing countries, to carry out the best possible evaluation and exploitation of local data sources. It was also meant to serve as the framework on which the detailed country reports prepared by the Committee on Population and Demography and its panels should be based. Use of preliminary sections of the *Manual* to carry out the Committee's work programme demonstrated that it is a useful research tool, disseminating the most recent techniques of demographic estimation. The collaboration of the Population Division of the Department of International Economic and Social Affairs of the United Nations Secretariat with the Committee on Population and Demography assures an even wider audience for the *Manual*, making it yet more important that users understand something of its origins in order to be forewarned of the limitations of the presentation.

Manual X is particularly valuable because it is the most complete compilation to date of techniques suited for the analysis of incomplete or defective demographic data. Even in this respect, however, the *Manual* is not perfect. During the past four years, the field of demographic estimation has expanded considerably, and the task of maintaining the *Manual* completely up-to-date was beyond the capabilities of its authors. Furthermore, as new methods were proposed and tested, it frequently became apparent that their robustness to error was not as high as expected. Therefore, in several instances, methods included in preliminary versions of the *Manual* had to be altered. It is possible that some of the methods included in the present version will require future modifications to improve their performance. Thus, in terms both of scope and of detail, this volume cannot be considered the last word on demographic estimation.

The second quality of *Manual X* is that it is "user-oriented". Emphasis has been given to the way in which estimation techniques are applied. Detailed examples are provided, and the overall tendency has been to include too much detail rather than too little. This strategy contrasts sharply with the practice normally followed in papers published by professional journals, often the only published sources of information about recently developed methods. Obviously, those already familiar with the methods described may consider such a level of detail unnecessary or overwhelming. However, experience has shown that the detail is useful for the less sophisticated user, for whom it is essential that no errors resulting from the faulty application of a method be confounded with those already present in the data used.

As a consequence of this user-oriented approach, a conscious effort has been made to simplify as much as possible the description of each method and of the rationale underlying it. Such a strategy is certainly in agreement with the "how to" character of the *Manual*, but it reduces its theoretical significance. Therefore, the user of *Manual X* will find information concerning the available techniques, the basic hypotheses underlying them, certain guiding principles about cases in which it is not appropriate to use given methods, detailed examples of how to apply the techniques proposed and some guidance as to the interpretation of the results they yield; but even the mastery of the entire *Manual* will provide the user with little insight into how new techniques or alternative approaches to available techniques can be developed. In this sense, the *Manual* is more a tool than a general theory of indirect estimation. The user who is interested in a more theoretical treatment should consult the papers available in the demographic literature, references to which are provided in the text.

A further limitation of the *Manual* is that it deals mainly with the estimation of fertility and mortality in developing countries. There are other demographic processes affecting the populations of these countries (migration, for example) which are not treated here. There is no theoretical reason that the techniques described should not be applied to accurate data from developed countries (and indeed many of them have been, with generally very satisfactory results) except that such applications are not necessary from the point of view of evaluating the data, though they are often of interest in evaluating the methodology itself. One area of demographic analysis for developed countries in which the techniques have proved useful is historical demography, where the data cannot always be assumed to be accurate. As long as the underlying assumptions are more or less met, the methods presented can be applied under any conditions, though their use may not be necessary.

Perhaps the most serious limitation of *Manual X* is that it does not provide sufficient guidance for the assessment of results, an aspect of analysis that is also somewhat perfunctorily considered in the literature. The reason for this omission is that the task of assessment is extremely complex, varying widely from application to application. Hence, it is almost impossible to lay down general rules to perform it. This limitation is exacerbated to some degree by the modular structure of the *Manual*. Indeed, since every section is designed to be as self-contained as possible, most examples are treated independently, leaving little room for the intermethod, intertemporal or interdata comparisons that are so illuminating in actual analysis. Because the country studies prepared for the Committee on Population and Demography of the National Research Council attempt to establish coherent pictures of the demographic evolution of a population, taking into account all the relevant data available, they provide more insights into the assessment of results from different methods, data sets and time periods, and into the way that sense can be made of a variety of imperfectly consistent estimates. Therefore, it is recommended that users of the *Manual* refer to the country reports listed at the end of this introduction as examples of how the results yielded by the methods described herein can be used effectively in combination rather than in isolation.

B. Definition of indirect techniques in demographic estimation

Demographic estimation consists of the attempt to measure values of basic demographic parameters, such as the birth rate, the death rate or the level of total fertility, under less than perfect conditions. These basic parameters indicate the way in which a population will evolve, in terms of size and age structure, over time.

The term "indirect" used to qualify some of the techniques used in demographic estimation has its origin in the fact that such techniques produce estimates of a certain parameter on the basis of information that is only indirectly related to its value. The classic example is the use of the proportion of children dead among those ever borne by women aged 20-24 years to estimate the probability of dying before age 2. The observed proportion of children dead is clearly related to mortality conditions, but it is not a pure mortality measure because it is affected by other, non-mortality parameters. In order to transform this proportion into the desired life-table function, the other parameters must be allowed for, generally by using procedures founded on demographic models. Therefore, not only is the information used "indirect", but the procedure followed, although considerably simplified in practice, is by no means straightforward theoretically. The extent of indirectness varies greatly, however, among procedures, in terms both of the reliance on models and of the number of unwanted factors that have to be allowed for. The term "indirect" is therefore used to describe any estimation method that depends upon models or uses consistency checks, or indeed uses conventional data in an unconventional way.

C. Need for indirect estimation

Traditionally, demographic estimation has been based on data collected by censuses and by a vital registration system. A continuous registration system usually has the task of recording vital events (births, deaths, marriages, divorces etc.) as they occur. When this system is coupled with periodic counts of the population (censuses), the calculation of demographic parameters becomes possible. Assuming that both the registration of vital events and the census counts were perfect, demographic parameters could be calculated directly from the data reported and there would be no need for indirect estimation. Unfortunately, however, in many countries today, either the data-collection systems described above do not exist or their performance is so poor that the estimates obtained directly from the data they produce are severely flawed.

Some of the possible deficiencies of a vital registration system, where it exists at all, may be outlined. The main deficiency is its failure to record all vital events as they occur. For example, it is well known that births may go unregistered for several years. Only when the child is ready to join the public education system, or some other type of organization for which a birth certificate is required, is the birth reported. In countries where a sizeable proportion of the population has never attended school or been in any way connected with official organizations demanding identity certificates, many births may never be registered. The same may be true of the births of children who die very young; in such cases, the parents may consider the registration of either the birth or the death to be futile. Adult deaths are likely either to be reported near the time of their occurrence or not at all. In most countries, a death must be registered before a burial permit is issued. Hence, more people may find it necessary to report deaths soon after they occur, especially in urban areas where burial grounds are often restricted to certain areas whose administrators are closely supervised by government officials. However, the necessity for a burial permit may be practically non-existent in rural areas, where ties with the government administration are weaker. Therefore, it is not surprising to find that, in spite of the legal necessity of registering deaths, many of them are never recorded. The vital registration system may also be deficient in recording characteristics of events, such as age at death, age of mother at a birth, or mother's parity after a birth.

Censuses, the second component of the traditional demographic estimation input, are also far from yielding perfect data. They suffer mainly from two types of error: the failure to enumerate all the members of the relevant population (though, occasionally, some censuses have also produced population overcounts) and poor age-reporting on the part of the population canvassed. Differential coverage of the population by age and sex is also very often present (young children, especially, tend to be undercounted to a greater extent than the adult population), and its effects cannot always be separated from those of age misstatement.

D. Evolution of indirect estimation and the contents of *Manual X*

Faced with the impossibility of obtaining reasonable measures of demographic parameters directly from the traditional data sources, demographers have developed a set of techniques that allow their indirect estimation. The development of these techniques has taken two main courses: either the search for robust methods to analyse data that have been collected by the traditional systems (such as a method designed to estimate a death rate from vital-registration data of uncertain accuracy); or the search for questions that can be answered with reasonable accuracy and that provide enough information about a certain demographic phenomenon to permit the indirect estimation of its level (for example, the use of information on the incidence of orphanhood to estimate adult mortality).

Because of their reliance upon special types of questions, indirect methods of estimation of the second type have come to be closely associated with sample surveys. Indeed, the sample survey provides perhaps the best means for collecting the data required for these techniques.[1]

During the first attempts to obtain estimates of demographic parameters from data gathered by sample surveys, the questions tried were the straightforward queries suggested by the traditional data sources. If a birth rate was to be estimated, the number of births occurring during a given year had to be known; hence, the question "How many births have occurred in this household during the past year?" was to be asked. It was soon realized that one could do better by asking each woman of childbearing age in the household: "How many children have you had during the past year?"; and recording her age. With that information, age-specific fertility rates could also be estimated. If an estimate of the death rate was needed, a question on deaths in the household during the year preceding the survey would be used. If a death had occurred, recording the age and sex of the deceased would permit the estimation of age-specific and sex-specific central mortality rates from which a life table could be derived. Thus, a complete demographic profile was at hand, just by asking two or three simple questions.

Unfortunately, the results obtained from these surveys did not match expectations. The death rates obtained were almost invariably too low and the birth rates usually seemed implausible. The failure of the question on deaths was explained by arguing that people do not like to talk of such events; they wish to forget them. Also, it was argued that the occurrence of a death within a household might bring about the household's disintegration, so that the households existing at the time of the survey would be more likely not to have experienced deaths of any of their adult members during the preceding year than would those which were in existence at some time during the course of the year.

A further reason postulated for the poor performance of these questions was the existence of a misconception on the part of the respondent of the length of the time period being alluded to. The vague expression "during the preceding year" might not be interpreted by those being interviewed as exactly the 12 months preceding the interview, but instead as a longer or shorter period, according to cultural factors.

Further trials with other ways of phrasing these questions have shown that their performance may be improved to some extent by making them more specific. For example, the period being considered should be specified as exactly as possible, preferably by quoting dates that have special relevance for the population being interviewed. Thus, when a survey is taking place during the second half of November 1978, for example, it may be better to ask "How many children have you had since 1 January of this year?" than "How many children have you had since 22 November 1977?" The first version of the question does not cover a complete year, but it would perform better than the second in a calendar-conscious society because 22 November is not likely to be a significant date for the respondents. The systematic errors introduced by restricting the information collected to periods that are shorter than a year in length are easier to correct than are those arising from incorrect responses.

Yet, although the data produced by these questions (or by the most recent and best version available to explore recent births—namely, "What was the date of your most recent birth?"—posed to all women of reproductive ages) are better, they are still far from perfect. Omission and recall errors are always present to varying degrees so that special analytical tools are necessary for the assessment and correction of the basic data. Some 20 years ago, Brass[2] proposed a method that made possible the adjustment of the reported data on births in the past year. This method is based on the simple idea of comparing the declared lifetime, or cumulated, fertility of young women (children ever born) with cumulated period fertility obtained from a question on births in the past year. Brass argued that although the number of births reported as occurring in the year before the survey might not be correct, the proportionate error might be more or less constant with the mother's age, so that the age pattern of recent fertility could be accepted even if the level could not. Similarly, one would expect younger women to report their lifetime fertility more accurately, since their births would have occurred fairly recently, so that their level of lifetime fertility could be considered reliable. The correction proposed by Brass was to adopt the two most reliable parts of the information available, taking the age pattern of fertility from the observed number of births in the past year and scaling it to match the level of fertility indicated by the lifetime fertility reported by younger women.

That idea gave rise to a series of methods exploiting the data on children ever born, obtained by asking women of reproductive ages the question: "How many children, who were born alive, have you ever had?" These methods are presented in chapter II. As far as possible, all of them take into account the usual deficiencies of data on children ever born, such as the tendency of older women to declare fewer children than the number they have really borne. This tendency has

---

1 For an in-depth analysis of the various data-collection techniques now available, see Panel on Data Collection, *Collecting Data for the Estimation of Fertility and Mortality*, National Research Council Committee on Population and Demography Report No. 6 (Washington, D.C., National Academy Press, 1981).

2 W. Brass, "Uses of census or survey data for the estimation of vital rates", (E/CN.14/CAS.4/V57), paper prepared for the African Seminar on Vital Statistics, Addis Ababa, 14-19 December 1964.

been explained by arguing that older women may either omit those children who have died or report only those who still live with them. Omission may be minimized by asking more specific questions, such as: "How many children do you have who live with you?", "How many children do you have who live elsewhere?", and "How many children have you had who have died?" The sum of the responses to these questions gives the desired number of children ever born.

A by-product of the data on children ever born gathered by the set of three questions just presented is the number of children surviving among those ever born. The ratio of these two numbers, that is, the proportion of the children who are still alive, has long been recognized as an indicator of child mortality, and Brass[3] was the first to suggest a method to transform these ratios (computed for women in each five-year age group) into life-table measures of mortality. Some of the variants of his original method are presented in chapter III. They constitute probably the most powerful techniques of indirect demographic estimation, especially in view of the difficulty in obtaining valid estimates of child mortality from registration data where, as pointed out earlier, deaths of young children are particularly susceptible to omission.

In the area of adult mortality estimation, two approaches were developed almost simultaneously. The first was based on the possibility of gathering data that would indirectly provide information on adult mortality. The demographic sample survey provides an ideal means of collecting these data, because it allows the types of questions that are expected to provide reasonably reliable results to be tested without great expense. So far, two questions have proved useful for the estimation of adult mortality. One proposed by Brass and Hill[4] to estimate female mortality investigates the maternal orphanhood status of each member of the population ("Is your mother alive?"); the other is used to estimate male or female mortality by investigating the survival status of the first spouse of women or men, respectively, who have been married at least once ("Is your first husband still alive?"; "Is your first wife still alive?"). Responses to these questions have been found to exhibit some characteristic biases. Young children whose true mothers have died are very often living with adoptive mothers; and the respondent to a question on survival of mother is not likely to make the distinction when answering, either because he or she simply does not know about the adoption or because he or she feels that the matter should not be reported. (Indeed, with young children, the "respondent" is not really the respondent at all, since most information is provided by an adult household member.) Therefore, the true incidence of orphanhood at younger ages is often underreported and the adult mortality of young adult females cannot be estimated with confidence. Data on widowhood, on the other hand, is affected by remarriage. To avoid this problem, only the survival of the first spouse should be investigated, and the question should make this distinction as clear as possible. Furthermore, in countries where a substantial proportion of the unions are of a consensual type (that is, not legalized) one must make it clear to the respondent that "spouse" and "consensual union partner" are equivalent. When phrased properly, this question can yield data of good quality. All existing methods of adult mortality estimation based on orphanhood and widowhood data are discussed in chapter IV.

In the second approach to adult mortality estimation, Brass[5] tried to make use of already existing data, obtained either by a registration system or by a question on deaths in the past year asked in a survey, on the number of deaths classified by age and sex occurring during a given period (usually a year) to a certain population of known age distribution. A simple model associating the reported age distribution of deaths with the age distribution of the population was developed, based on the assumption that deaths are equally underreported at each age and that the population is stable, and was used as a method allowing the estimation of the extent of relative underregistration. This method and a similar technique proposed by Preston and Coale[6] are presented in chapter V.

Most of the techniques described above are applied to data that are generally gathered by sample surveys, though it is primarily cost factors which prevent them from being gathered by censuses, which are, after all, only a special type of survey. Yet, typically, censuses have not included many of the questions necessary for indirect estimation, in part because of the necessity of maintaining the census questionnaires short in order not to prejudice the overall quality of the data. Hence, it is fairly common for censuses to establish only the age and sex distribution of the population being counted and, perhaps, to gather some information on household structure, as far as demographic information is concerned. Only three questions are required to fulfil this goal: one on age; one on sex; and one exploring the relationship of each reported household member to the head of the household. Almost every census carried out during the twentieth century has included questions on age and sex, so that age distributions of the population by sex are available for many countries for different points in time. The development of mathematical models relating different aspects of population growth to the observed age distribution[7] has given rise to several

[3] *Ibid.*

[4] W. Brass and K. H. Hill, "Estimating adult mortality from orphanhood", *International Population Conference, Liège, 1973* (Liège, International Union for the Scientific Study of Population, 1973), vol. 3, pp. 111-123.

[5] William Brass, *Methods for Estimating Fertility and Mortality from Limited and Defective Data* (Chapel Hill, North Carolina, Carolina Population Center, Laboratories for Population Statistics, 1975).

[6] Samuel H. Preston, Ansley J. Coale, James Trussell and Maxine Weinstein, "Estimating the completeness of reporting of adult deaths in populations that are approximately stable", *Population Index*, vol. 46, No. 2 (Summer 1980), pp. 179-202.

[7] Ansley J. Coale and Paul Demeny, *Regional Model Life Tables and Stable Populations* (Princeton, New Jersey, Princeton University Press, 1966); and Ansley J. Coale, *The Growth and Structure of Human Populations: A Mathematical Investigation* (Princeton, New Jersey, Princeton University Press, 1972).

techniques of estimation based on the age structure of the population at a given time.[8] These techniques often make use of sets of mortality models or of models of age distributions of theoretically stable populations. Such sets are known as "model life tables" and "model stable populations", and some of the sets available are described in chapter I. The estimation techniques that use them are presented in chapter VII.

Traditionally, the information yielded by a census question on relationship to the head of the household was not used for demographic estimation purposes. That situation changed when Cho[9] suggested a method that uses the stated relationship to head of the household to identify, where possible, the mother of each enumerated child. Once the child's mother is identified, the age of the child and that of the mother are available; and it is possible to compute age-specific fertility rates for each of the 10 or 15 years preceding the census whenever estimates of child and female adult mortality are available. In countries where young children are not differentially undercounted, where most young children live with their mothers and where age-reporting is reasonably accurate, this method yields good results. It is especially suited for the study of the age structure of fertility—provided, of course, that age-misreporting is not severe. Since this method requires for its application estimates of demographic parameters that are usually obtained by methods described in earlier chapters, it is presented in chapter VIII.

Lastly, if a country possesses a relatively good death registration system, it may be possible to compare its performance with indicators of mortality provided by successive censuses, since the difference between counts of the same cohort at two different points in time is, in the absence of net migration, only the result of mortality. A method that, by comparing death registration and census counts, permits the estimation of death under-registration and relative census completeness of enumeration[10] is presented in chapter IX.

E. USE OF *Manual X*

Manual X is intended to describe some of the more promising techniques currently available to make indirect estimates of demographic parameters. Each of the techniques presented is based on a more or less simplified model of reality, where by "model" is meant a set of mathematical relations between relevant demographic variables. Some understanding of the model underlying each technique is essential for its adequate use, and to understand a model is to have a clear idea of the hypotheses or assumptions on which it is based. In describing each technique in this *Manual*, the assumptions on which its underlying model is based are always emphasized.

The models underlying the techniques presented often take into account some of the flaws that are likely to be present in actual data. For example, the technique that allows the estimation of total fertility by comparing cumulated current fertility with reported average parity specifically assumes that recent fertility has the right shape but not necessarily the right level. In other words, it takes into account the deficiencies noticed in data referring to recent fertility and, in consequence, it is designed to be robust to the typical errors found in actual data. As far as possible, other techniques are also based on models that make some allowance for the existence of typical data errors and are therefore well suited to permit the estimation of demographic parameters from data sets that are less than ideal.

However, even though these techniques are quite powerful, they do not solve all the problems that arise in demographic estimation. For example, their performance is not totally immune to the effects of age-misreporting, nor is it entirely adequate when differential errors by age are present. Therefore, the analyst using these techniques is well-advised to assess fully the quality of the data being analysed before accepting or rejecting the parameter values yielded by particular indirect methods. If a range of indirect techniques can be applied in a particular case, it is unlikely that their results will be entirely consistent. The choice of a final estimate among the possible candidates will be based on knowledge of the ways the methods generally perform, on knowledge of the particular data sets being used, and on knowledge of the country itself and its customs. Such judgements are based largely on experience, and no attempt has been made in this *Manual* to provide general rules for them beyond indicating how particular errors affect the results of particular methods.

The strategy followed in producing *Manual X* has been to describe in the simplest possible way the assumptions and, if necessary, the actual models on which each technique is based; and to provide the reader with a detailed description of the actual application of each method. The organization of each section describing a technique consists usually of four subsections. The introductory portion describes the basis of the method and its rationale; the second lists the data required for its application; the third describes, in general terms, how the method is applied; and the fourth is an example of its application to an actual data set. Both the subsection describing the computational procedure and that corresponding to the detailed example are divided into different and clearly labelled steps.

As far as possible, the section describing a particular technique is intended to be independent from other sections of the *Manual*, so that the user may refer to it without having to work through the entire volume. Occasionally, when a section is not entirely self-contained or self-explanatory, the reader is referred to

---

[8] *Manual IV: Methods of Estimating Basic Demographic Measures from Incomplete Data* (United Nations publication, Sales No. E.67.XIII.2).

[9] Lee-Jay Cho, "The own-children approach to fertility estimation: an elaboration", *International Population Conference, Liège, 1973* (Liège, International Union for the Scientific Study of Population, 1973), vol. 2, pp. 263-280.

[10] S. Preston and K. Hill, "Estimating the completeness of death registration", *Population Studies*, vol. 34, No. 2 (July 1980), pp. 349-366.

other parts of the *Manual* where the necessary concepts or procedures are explained. However, the user should always read the introductory comments of a chapter before using a method from that chapter, since the general characteristics of the data are described in that section.

In general, all the techniques presented are fairly easy to apply. The procedures described are especially suited for the analyst who possesses a hand calculator and some patience, and who is willing to carry out sometimes lengthy computations carefully. The background required to understand the assumptions and the procedures described in chapters II-IX consists of a familiarity with basic demographic measures and a command of algebra. However, it is realized that to grasp fully the mathematical derivation of the demographic models presented in chapter I, a good command of statistics and of calculus is necessary. This chapter, although presenting some background material that is very useful in understanding the other chapters of the *Manual*, does not conform to their level of simplicity; and although it is recommended that every user read it, full understanding of its contents is not necessary for the successful application of the methods described.

It is hoped that this *Manual* will fulfil two functions: as a didactic tool especially suited for those wishing to learn how to perform indirect demographic estimation; and as a handbook for the experienced demographer who needs a quick guide to the application of a variety of demographic estimation techniques. It is strongly recommended that both types of users read this introduction entirely before approaching the techniques themselves, as it contains the information necessary to understand the conventions and notations used throughout the *Manual*. Furthermore, when only one method in a given chapter is of interest to the user, it is recommended that the introductory section of the chapter be read before focusing on the method itself. This section usually states the assumptions on which the method is based and the errors typically found in the type of data being used. It should not, therefore, be overlooked.

The techniques presented have been grouped into chapters according to the type of data they require and not according to the type of estimates they yield. Therefore, several chapters need to be considered when one is interested in parameter types rather than in data types. For example, age-specific fertility rates may be estimated from data on children ever born or from census or survey data where the identification of own-children is possible; and while all the techniques using data on children ever born to estimate fertility are presented in chapter II, the own-children method of fertility estimation is presented in chapter VIII, together with other reverse-survival techniques. The same observation could be made about adult mortality estimation, which can be performed either when a distribution of deaths by age and sex exists or when data on orphanhood and widowhood or even when successive age distributions are available. It is hoped that table 1 and the annotated index tables of each chapter may make the selection of the techniques applicable to different data sets straightforward.

To conclude, it should be pointed out that although all the examples presented are based on real data, the results obtained from the application of a certain method to an isolated set of information should not be considered definitive. Frequently, the application of a method requires several pieces of information obtained from different data sets or sources; and it is impossible within the scope of this *Manual* to assess in each case whether these external pieces of information are satisfactory. In other instances, the application of a certain method of estimation reveals inconsistencies in the data that make the results suspect. In such cases, further exploratory and explanatory work would be required to resolve these inconsistencies. Yet, since the purpose of *Manual X* is to illustrate the way in which the methods

TABLE 1. SCHEMATIC GUIDE TO *Manual X*

Chapter		Type of input data	Estimated parameters
II.	Estimation of fertility based on information about children ever born	Children ever born Births in the past year Number of women Total population	Total fertility Age-specific rates Birth rate
III.	Estimation of child mortality from information on children ever born and children surviving	Children surviving Children dead Children ever born Number of women	Probabilities of surviving from birth l(2); l(3); l(5); l(10); l(15); l(20)
IV.	Estimation of adult survivorship probabilities from information on orphanhood and widowhood	Respondents with mother alive Respondents with mother dead	Probabilities of surviving from age 25 to age 25 + x, for $x = 10, 15, ..., 60$ among females
		Respondents with father alive Respondents with father dead	Probabilities of surviving from age 32.5 (or 37.5) to age 35 + x (or 40 + x) for $x = 10, 15, ..., 55$ among males

TABLE I (*continued*)

Chapter	Type of input data	Estimated parameters
	Ever-married female population with first husband alive Ever-married female population with first husband dead	Probabilities of surviving from age 20 to age x, for $x = 25, 30, ..., 60$ among males
	Ever-married male population with first wife alive Ever-married male population with first wife dead	Probabilities of surviving from age 20 to age x, for $x = 25, 30, ..., 55$ among females
V. Estimation of adult mortality from information on the distribution of deaths by age	Deaths classified by age Population classified by age Growth rate	Adjustment factor for deaths Adjusted age-specific central mortality rates for ages over 10
VI. Derivation of a smooth life table from a set of survivorship probabilities	Estimates of child mortality Estimates of adult mortality from orphanhood, or widowhood, or death registration	Complete life table
VII. Fertility and mortality estimation using model stable age distributions	Growth rate Child mortality ($l(z)$) Age distribution	Birth rate Death rate Gross reproduction rate Life table
VIII. Estimation of fertility by reverse-survival methods	Population classified by age Estimates of child mortality Growth rate	Birth rate
	Enumerated children by single year of age and single year of mother's age Female population by single years Estimates of child mortality Estimates of female adult mortality	Age-specific fertility rates for the 10 or 15 years preceding enumeration Total fertility for the same time period
IX. Estimation of adult mortality using successive census age distributions	Population enumerated at two points in time five or 10 years apart, classified by five-year age group	Life table
	Population enumerated at two points in time classified by age Recorded intercensal deaths classified by age	Completeness of death registration Completeness of census enumeration Life table

are applied and not to obtain complete, coherent demographic profiles of the cases studied, a thorough validation of the results obtained has not generally been attempted. Hence, although this *Manual* is well suited to instruct the user on how to obtain the pieces of the puzzle, it does not go far enough in showing how to put them together to form a consistent picture of the demographic situation of the population concerned.

F. DEFINITIONS AND CONVENTIONS

This section discusses briefly the conventions used throughout this *Manual* concerning both notation and some general definitions. Of course, the specific definitions needed in particular chapters are given there; but in order to avoid repetition, frequently used symbols are defined in this section.

First to be discussed is notation. Because of the wide variety of concepts used, it has been impossible not to assign different meanings to the same symbol in different chapters. Even though this practice clearly violates strict mathematical conventions, its flexibility far outweighs the slight confusion, if any, that it may cause. In order to avoid such confusion completely, the reader must be aware of the cases in which different meanings are associated with the same symbol. Within any chapter, a given symbol always refers to the same type of data; however, its exact algebraic definition may vary according to indices associated with it. For example, the symbol N usually represents the number of persons in some age group, but $N(x)$ is the symbol used to represent the number of persons in the age group from x to $x+1$; ${}_5N_x$ is the number of persons whose exact ages range from x to $x+5$, $N(x+)$ represents the number of persons aged x and over, and $N(x-)$ is the number of persons below age x. Of course, the exact definition of each of these variations of the symbol N is stated within the chapter where it is used, but it would be helpful for the user to remember these frequently used variations of meaning.

Another example is the use of P to represent average parity. Usually, $P(i)$ represents the average parity of women whose ages range from $x_1 = 10+5(i)$ to $x_2 = 10+5(i+1)$, although occasionally it represents the average parity of women who were first married between $y_1 = 5(i-1)$ and $y_2 = 5(i)$ years before the survey. Furthermore, in sections where more than one survey is considered, $P(i,t)$ is used to represent the average parity of women in group i at time t.

Some symbols have different meanings in different chapters. Since this *Manual* is organized in a modular fashion, with each chapter or module being largely independent of other chapters, this misuse of symbols is not expected to cause problems. For example, it will be quite clear to the reader that in chapter VII, $C(x)$ is used to represent the proportion of the population under age x, while in chapter V, the letter C is used to represent the estimated completeness factor for death registration. Since each of these definitions of C is used consistently within each chapter, it is unlikely that confusion will arise, as long as it is remembered that chapters should be treated as modules.

On the other hand, some symbols are used consistently throughout the *Manual*. These symbols are defined below:

b: birth rate. More accurately called the "crude birth rate", b is the ratio of the number of births occurring in a population during a given period and the number of person-years lived by the population during that time period. The latter figure is usually approximated by the size of the population at the mid-point of the period, multiplied by the length of the period in years. This rate is usually expressed in births per 1,000 population; that is, its true value is multiplied by 1,000. Yet, in this *Manual*, this convention is not followed. Therefore, all birth rates quoted are numbers smaller than unity, which require multiplication by 1,000 to be comparable to figures cited elsewhere;

d: death rate. The complete name of d is the "crude death rate". It is the ratio of the number of deaths occurring in a population during a certain time period and the person-years lived by the population during that time period. An approximation of this number of person-years lived is often obtained by multiplying the mid-period population by the period length in years. Usually, death rates are also expressed per 1,000 population (that is, their true value is multiplied by 1,000), but this practice is not followed here. Hence, all death rates quoted in the *Manual* are numbers smaller than unity;

r: growth rate. More properly known as the "crude rate of natural increase," r is the difference between the birth and death rates defined above. Crude rates of natural increase are often expressed as percentages. This practice is entirely avoided in this *Manual*, but especially in this case, the reader must be aware of it so as to be able to make the correct comparisons between the figures quoted here and those presented in other sources. As an example, note that a value of $r = 0.0280$, the form in which growth rates are expressed herein, is equivalent to a growth rate of 2.80 per cent in conventional form;

$q(x)$: probability of dying. Throughout the *Manual*, the life-table probability of dying between birth and exact age x is denoted by $q(x)$. Being a probability, the value of $q(x)$ cannot be greater than one. Although its value can also be quoted as per 100 (per cent) or per 1,000, it is not the usual practice to do so. Hence, the reader will find that the $q(x)$ values given here have the same form as those generally used elsewhere;

$l(x)$: probability of surviving. The value of $l(x)$ is, in probability terms, the complement of $q(x)$, that is, $l(x) = 1.0 - q(x)$. By definition, $l(x)$ is the probability of surviving from birth to exact age x; and it is conventional practice to express the $l(x)$ values in per 1,000 or per 10,000 form. This practice is especially prevalent in the construction of life tables where $l(x)$ is also interpreted as the number of persons of exact age x in a stationary population with radix $l(0)$. The arbitrary selection of the value of $l(0)$ amounts to multiplying all the true probability values of $l(x)$ by a constant, in order to avoid using decimal points. In most of this *Manual*, $l(0)$ has the value of one; and therefore, $l(x)$ is a probability value never greater than one. However, an exception to this convention occurs in section D of chapter IX, where the radix of the estimated $l(x)$ values cannot be known, therefore making impossible their translation into true probabilities of survival from birth. The user should make a note of the general convention and of this main exception in order to make allowance for the final position of the decimal point when making calculations or comparisons with respect to other estimates;

e_x: expectation of life. The symbol e_x is used to denote the average expectation of life of a group of survivors on reaching age x. The expectation of life at birth is denoted by e_0. The value of e_x represents the average number of years that the members of the stationary (life-table) population alive at exact age x may expect to live. The reciprocal of e_x $(1.0/e_x)$ is the death rate over age x for the stationary population. Another name for e_x is "life expectancy at age x";

${}_5m_x$: central death rates, specific by age. The symbol ${}_5m_x$ is used to represent central death rates calculated directly from the number of deaths occurring to the population in the age group from x to $x+5$ and the person-years lived. When the equivalent measure is calculated for a stationary population (life table), an upper-case M is used to indicate this fact (${}_5M_x$);

P: average parity. Demographers define parity as the number of live births a woman has had. Average parity is, therefore, the average number of live births per woman, or the average number of children ever born per woman. The use of the index i to denote that average parities refer to different age groups of women has become common practice. Therefore, instead of denoting the average parity of women in age group 15-19 by ${}_5P_{15}$, as the usual actuarial notation would dictate, the symbol $P(1)$ was selected. Ease in typing $P(i)$ rather than ${}_5P_x$ has led to the widespread acceptance of the former notation, and it was adopted for this *Manual*. In fact, it has even been abused, in a sense, because i represents both age and marriage-duration groups, according to the case at hand, and it was extended by using it to index some variables other than parity. Because of the simplicity and frequent use of the indexing variable i in this *Manual* and elsewhere, the user is advised to become familiar with it;

$f(i)$: age-specific fertility rate. In this *Manual*, $f(i)$ represents the age-specific fertility rate of women whose age ranges from $x_1 = 10+5i$ to $x_2 = 10+5(i+1)$. In general, the index i used for parities is used in an identical fashion for age-specific fertility rates. Thus, $f(1)$ represents the age-specific fertility rate for women aged 15-19;

TF: total fertility. Total fertility is, by definition, the average number of children born to women who survive the entire reproductive life span; it can be either a period or a cohort measure. When total fertility is measured on the basis of period data, it is an estimate of the average number of children born to members of a hypothetical, mortality-free cohort that is subject, during its entire reproductive life, to the age-specific fertility rates for the period. Thus, if the age limits of female reproduction are taken as being 15 and 50,

$$TF = 5\sum_{i=1}^{7} f(i);$$

Exp and ln: the exponential and the natural logarithm, respectively. The symbol exp is used to denote the value of e, an irrational number whose first seven digits are 2.718282. This number is defined mathematically as the base of natural logarithms (denoted in this *Manual* by ln, though also commonly denoted by $\log_e$). The use of exponential and logarithmic functions is common in the specification of demographic models and the reader should be acquainted with their mathematical properties.

Even though this list of consistently used symbols is not exhaustive, it does highlight the most important characteristics of the relevant ones. The reader may find it useful to consult this section whenever any of the symbols listed here appears in later chapters.

Additional references

Coale, Ansley J., Lee-Jay Cho *and* Noreen Goldman. Estimation of recent trends in fertility and mortality in the Republic of Korea. Washington, D.C., National Academy of Sciences, 1980. (National Research Council, Committee on Population and Demography, report 1)

Ewbank, Douglas C. Age misreporting and age-selective underenumeration: sources, patterns, and consequences for demographic analysis. Washington, D.C., National Academy Press, 1981. (National Research Council, Committee on Population and Demography, report 4)

Hill, Kenneth. Fertility and mortality changes in Honduras, 1950-1974. Washington, D.C., National Academy of Sciences, 1980. (National Research Council, Committee on Population and Demography, report 3)

Panel on Bangladesh. Estimation of recent trends in fertility and mortality in Bangladesh. Washington, D.C., National Academy Press, 1981. (National Research Council, Committee on Population and Demography, report 5)

Panel on Data Collection. Collecting data for the estimation of fertility and mortality. Washington, D.C., National Academy Press, 1981. (National Research Council, Committee on Population and Demography, report 6)

———. Data collection: a statement for administrators. Washington, D.C., National Academy Press, 1981. (National Research Council, Committee on Population and Demography, report 7)

Panel on Egypt. The estimation of recent trends in fertility and mortality in Egypt. Washington, D.C., National Academy Press, 1982. (National Research Council, Committee on Population and Demography, report 9)

Panel on Thailand. Fertility and mortality changes in Thailand, 1950-1975. Washington, D.C., National Academy of Sciences, 1980. (National Research Council, Committee on Population and Demography, report 2)

Shorter, Frederic C. *and* Miroslav Macura. Trends in fertility and mortality in Turkey, 1935-1975. Washington, D.C., National Academy Press, 1982. (National Research Council, Committee on Population and Demography, report 8)