Chapter III

ESTIMATES OF FERTILITY AND MORTALITY BASED ON REPORTED AGE DISTRIBUTIONS AND REPORTED CHILD SURVIVAL

In chapter I there are described methods of estimating fertility and mortality that make use of the age distribution of a population recorded in one or more enumerations. A second enumeration provides valuable supplementary information in the form of the rate of increase and survival rates. When there has been only one usable enumeration of a population, only very rough estimation is usually possible. However, in chapter II a technique of calculation is outlined that makes it possible to determine proportions surviving from birth to age, 2, 3, 5 and sometimes to ages up to 30 or 35 from data supplied by women about the history of the children they have borne.

Knowledge of childhood mortality is an extremely useful adjunct to knowing the age composition of a population, particularly for the estimation of fertility. In fact, responses about the survival of children ever born are probably more useful in estimating recent fertility from the age composition recorded in a census or survey than is the existence of an earlier enumeration. Specifically, if $C(x)$ is known, l_2 is a more useful supplementary datum than r, the rate of natural increase, in estimating the birth rate. On the other hand, knowledge of l_2 (with or without the age distribution) gives only inferential evidence about adult mortality (i.e., mortality above age five). From data on childhood mortality, mortality above age five can be estimated only on the basis of assumed regular relations between mortality at different ages, and in different populations the relation of child mortality to adult mortality varies substantially. For example, in the four families of model life tables expressing average mortality patterns in different "regions", expectations of life were calculated at age 5 ranging from forty-six years to over fifty-three years associated with the same proportion (about 75 per cent) surviving from birth to age two.

Two enumerations spaced five or ten years apart, on the other hand, provide a good indication of the level of adult mortality, but no direct evidence on childhood mortality. Expectations of life at age five estimated from model stable populations with a given $C(30)$ and a given rate of increase vary only from 51.7 to 52.4 years when based on the different "regional" model tables; the estimated level of mortality above age five in this instance is essentially independent of variation in age pattern. But when r and $C(x)$ are known, *childhood mortality* must be approximated by assuming some kind of "normal" association between child and older age death rates. It will be seen in each of the techniques presented in this chapter that a solidly based figure for child mortality is needed for any precision in an estimate of the birth rate derived from an enumeration of a population.

A. Estimation of birth and death rates from childhood survival rates and a single enumeration by reverse projection

An accurate census that records the number of persons in each five-year age interval provides the basis for reconstructing recent birth and death rates, if migration either is known accurately or is negligible in magnitude, and if survival rates by age are known for recent periods. The method of estimation is simply to reverse the customary procedures of population projection—to reconstruct by reverse survival the births that brought into being the children recorded as under age five or ten, and to reconstruct the population among which the births occurred by reverse survival of persons over five or ten in the census.

The specific steps employed in reverse projection are to divide the population under five by ${}_5L_0/5 \times l_0$ from a life table representing the mortality of the preceding five years to obtain an estimate of births, and to divide each five-year age group by the appropriate survival factor from the same life table to reconstruct age group by age group the population five years before. The sum of all such estimated age groups is the estimated total population five years earlier. The estimated average annual number of births can then be divided by the average of the enumerated population at the end and the estimated population at the beginning of the period to give an average birth rate during the preceding five years. The average annual rate of increase ($1/5 \log_e P_t/P_{t-5}$) can be subtracted from the birth rate to estimate the death rate. Similarly, the population 5-9 can be projected back to provide the estimated birth rate in the next earlier five years, etc. The estimated total population becomes subject to increasing uncertainty as the reverse projection procedes, however, even if mortality is somehow accurately known: the oldest age group in the past population has no current survivors, and this segment of the past population must be estimated by some assumption about the nature of the past age distribution. For earlier and earlier dates, the portion of the population estimated in this way is larger and larger.

Given an accurate census, the crucial additional element for reverse projection is an appropriate life table. The Brass method of estimating child survival provides approximate values of l_2 (during the preceding four or five years) and l_3 (during the preceding six or eight). A model life table can be selected with the given l_2, and survival factors from this table employed for the reverse projection. The value of ${}_5L_0$ in the model table is very close to the correct one, if the data from which l_2 was estimated are accurate. Differences in age patterns of mortality do not much affect the ratio of l_2 to ${}_5L_0$. On the other hand survival rates above age five are estimated by assuming that the age pattern of mortality conforms to the "West" family of model life tables, and if one judges by differences among life tables based on accurate data, the actual survival rates above five may diverge from the "West" family. However, differences in mortality above age five in life tables with a given l_2 would rarely produce estimates of over-all population 2.5 years earlier differing by more than one per cent, so that the error in the estimated birth rate from this source would rarely exceed half a point (e.g., an estimate of 50.5 per thousand instead of 50).

If the questions on children ever born and surviving have been asked and recorded separately for males and females, separate estimates of child mortality can be made for each sex and (if the internal consistency of the data is acceptable) the model life table selected for each sex can be based on this separate evidence. However, the questions on children ever born are typically asked (or tabulated) only for both sexes combined. One is tempted to derive, from values of l_2 and l_3 for the two sexes together, estimates for each sex on the assumption that the relation of female to male child mortality in the given population is the same as in the population whose experience underlies the "West" model tables. The typical relation between male and female mortality in these populations is that male and female life tables tend to be at about the same level; i.e., when l_2 for females is .72765 (level 7, ${}^0e_0 = 35.0$), the typical l_2 for males is .69537 (also level 7, ${}^0e_0 = 32.48$). It is easy to construct a table containing the combined l_x values for $x = 1$ to 5 for both sexes (assuming a typical sex ratio at birth of 105 males per 100 females) at each "level", and then to assume that male and female mortality is expressed by the life table for each sex at the "level" indicated by the value of l_2 and l_3 for the two sexes together. However, the evidence available on sex differences in mortality in the less developed countries does not warrant the assumption that these differences always conform to the relations found in the experience—primarily from Europe, North America, and Oceania—underlying the model life tables. It is possible to find[1] many examples of female mortality higher than male mortality and this resort to male and female model tables at the same level may introduce a mortality differential opposite to the actual one.

If the age distributions of both sexes are about equally usable as a basis for estimating fertility, the uncertainty of sex differences in mortality can be ignored, and the values of l_2 or l_3 estimated for the two sexes combined can be employed as if it were a valid estimate for each sex separately. If in fact the sex differences in child mortality are substantial, estimates based on the common value of l_2 or l_3 will overstate the birth rate and death rate for one sex, and understate the rates for the other, but provide unbiased figures for the whole population. In general, however, the reliability of estimation is greater when based on age data from one sex than when based on the other. It is then necessary to make some rough estimate of the l_2 or l_3 implied for males and females by l_2 or l_3 for the two sexes combined. Often there is indirect evidence indicating the direction and approximate extent of sex differences in mortality: the sex ratio of a closed population that has not experienced large sex selective military deaths indicates the sex incidence of mortality since the sex ratio at birth can often be closely estimated; and the sex ratio of mortality in registration areas, in sample surveys, or even in neighbouring populations can be taken as relevant evidence.

Reverse projection cannot be recommended as a generally satisfactory basis of estimating birth rates even when calculation of child survival rates is possible because of the frequent unreliability of recorded age distributions. The tendency for the proportion under five to be underreported in many censuses and surveys has often led demographers to base estimates of the birth rate on the reverse projection of persons five to nine. However, the proportion of the population five to nine is frequently overstated by a wide margin (because of understatement of the ages of adolescent girls, for example, and because there is a common tendency to overstate the age of some children under five), so that this procedure cannot be endorsed as always valid. In fact, if there is evidence of substantial age-misreporting, it is not possible to make good use of reverse projection unless some means is available for identifying a valid part of the reported age distribution or of adjusting the reported figures.

B. Estimation of birth and death rates from child survival rates and a single enumeration by model stable populations

The use of model stable populations to estimate characteristics of a population requires the identification of a stable population among the tabulated age distributions that shares some of the observed or inferred characteristics of the recorded population.[2] In chapter I, the identifying features used to locate a model population were the intercensal rate of increase (assumed equal to the rate of growth of the stable population), and the cumulative age distribution or ogive up to some age that depends on the apparent pattern of age-misreporting. Under the conditions considered in this section, a model stable population is identified by the estimate of l_2, which by means of table I.2 in annex I determines the level of mortality, and by the ogive, which is used in a manner wholly analogous to the procedures outlined in section B of chapter I. For example, if the age distribution is of the Indian-Pakistani-Indonesian-African pattern, a minimum estimate of b can be obtained from C(15), and a less biased estimate from the value associated with C(35).

[1] Pravin M. Visaria, *The Sex Ratio of the Population of India* (unpublished doctoral dissertation, Princeton University, 1963). Available at University Microfilms, Ins., Ann Arbor, Michigan.

[2] As noted earlier, the identification of a model stable population does not determine unique value of the gross reproduction rate or of total fertility. To estimate these quantities, the mean age of the fertility schedule must first be estimated. For a discussion of this topic, see section B.5 in chapter I.

Two features of estimation by model stable populations selected by l_2 and $C(x)$ are worth noting: first, if an estimate of the birth rate is based on $C(5)$, the result is essentially identical with the results obtained by reverse projection, whether or not the population is stable; and second, estimates of the birth rate obtained from l_2 and $C(x)$ are insensitive to differences in age pattern of mortality, at least the differences found in the four families of model tables.

The birth rate in the model stable population with the same l_2 and $C(5)$ as the observed population is identical to that which would be obtained by applying reverse projection to the children under five in the stable population to obtain births, and reverse projection to the whole stable population to obtain an average number of persons (the denominator of the birth rate). The number of births estimated for the actual and the stable populations is identical, so that the only source of difference between the birth rate estimated by reverse projection and that found in the model stable population is in the denominator, which in each case is a number obtained by applying reverse projection, with the same life table, to populations with the same number of persons, and the same number over and under five, but possibly differing in the internal age structure above age five. This point loses relevance as the choice of a model stable population is made dependent on $C(10)$, $C(15)$ and ogives to higher ages, because with possibly different internal age composition in the population (e.g., under fifteen) that is implicitly or explicitly projected back to birth, and with differences in the size of the reverse projected denominator becoming more pronounced as the time period of reverse projection is extended, the virtual identity of the two estimates is lost. This feature does imply, however, that interpolation in stable populations is a convenient way of determining the birth rate implied by reverse projection of children under one, under five, one to five, five to ten or under ten.

The insensitivity of birth rate estimates from l_2 and $C(x)$ to differences in age patterns of mortality is an important advantage of such estimates. The advantage lies in the fact that the true age pattern of mortality is usually not known, and if estimates based on alternative plausible age patterns are widely different, the range of uncertainty is great. This point is discussed in greater detail in chapter IV.

C. Estimation of birth and death rates from child survival rates and age distribution in a population enumerated several times

When a population has been enumerated more than once at an interval of about five or ten years, the methods of estimation presented in chapter I can be applied, and at first thought the mortality estimates derived from data on survival among children ever born might be considered merely as a verification of the mortality estimated on the basis of $C(x)$ and r, or $C(x)$ and a model life table consistent with fractions surviving from one census to the next. However, the best use of such data is to accept the estimates they provide of mortality under age five, and of any other population parameters dependent on mortality in this age range, and to accept the estimates of mortality over age five that can be derived from an analysis of the two censuses—either by survival analysis, or by accepting the mortality above age five consistent with $C(x)$ and r. Of course such general advice is contingent on the quality of the basic figures. If two censuses are unequal in completeness of coverage, or if international migration is substantial and inadequately recorded, it would be necessary to rely more on inferences that could be based solely on age distribution and estimated child survival. Similarly, if internal inconsistencies were apparent in the reports of children ever born, the estimates of child survival might not be usable.

1. *Estimation of birth and death rates in a non-stable population*

Suppose a population is enumerated in censuses at the beginning and end of a decade, and that the second census includes data permitting the calculation of l_2 and l_3. The best estimate of the over-all life table is obtained by accepting ${}^0e_5$ (and the m_x and q_x values for age five and over) from the model life table chosen as best fitting census survival rates, and by taking l_5 (and ${}_5m_0$) from the model life table with the value of l_2 or l_3 obtained by the Brass methods. The expectation of life at birth in this hybrid model table is easily calculated. The average death rate for the decade is then calculated by applying the m_x values in this hybrid life table to a rough mid-decade age distribution obtained by averaging the distributions at the beginning and end. The birth rate can then be estimated as equal to the death rate plus the average annual rate of increase.

2. *Estimation of birth and death rates in a stable population enumerated more than once*

If the population enumerated in two or more censuses has closely similar age distributions in each census, the methods just described can be supplemented by the following procedures: (*a*) estimate the birth rate by selecting a model stable population from l_2 and $C(x)$; and (*b*) estimate the death rate as the birth rate less the intercensal rate of increase. A slightly more elaborate procedure may be applied when an estimated life table, an adjusted age distribution, and other detailed parameters are sought. This procedure recognizes the superiority of the Brass estimation procedures for determining childhood mortality, and of stable population techniques using $C(x)$ and r to estimate mortality above age five. The procedure entails: (*a*) selecting a model stable population from l_2 and $C(x)$, and accepting the proportion under five and the child death rate in this population, and (*b*), selecting a model stable population from $C(x)$ and r, and accepting the age specific death rates above five, and the age distribution within the span above five in this population. The over-all death rate is then estimated as the sum of the death rate under five in (*a*) times the proportion under five in stable population (*a*) plus the death rate over five in (*b*) times one minus the proportion under 5 in stable population (*a*).

D. Adjustment of estimates of fertility derived from child survival rates and the age distribution when mortality has been declining

If the population for which childhood survival can be estimated appears to have experienced declining mortality during the recent past, the stable estimates of the birth rate and of the gross reproduction rate (assuming that an estimate of the mean age of the fertility function has been obtained previously) can be adjusted to allow for the effects of declining mortality on the age distribution. The procedure is the same as that described in section C of chapter I, using the adjustment factors from the appropriate part (part (*c*) or (*d*)) of table III.1 in annex III. If questions on child survival were asked in two consecutive censuses, it is also possible to estimate the parameter k by means of table III.4.