Chapter I

METHODS OF ESTIMATION BASED ON RECORDS OF POPULATION GROWTH AND DISTRIBUTION BY AGE

There are many populations that have no usable direct records of births or deaths, but have been enumerated in one or more censuses in which the age and sex of each person was recorded. In this chapter methods of estimating fertility and mortality from such census data are described. The underlying rationale of such methods is that the growth and age structure of a population are determined by the mortality, fertility, and external migration to which it has been subject, and consequently the possibility exists of estimating those forces from records of the evolving size and structure of the population.

The description of estimation in this chapter proceeds from methods applicable to any closed population (or to any for which records of gains and losses by migration exist) to methods applicable to closed populations with special histories—of essentially unchanging fertility and mortality, and of unchanging fertility combined with declining mortality.

A. Estimation of mortality from census survival rates and the consequent estimation of birth and death rates

Suppose that a closed population is enumerated in two censuses at an interval of exactly ten years, and that each census contains tabulations of males and females by age, in five-year intervals. Each cohort enumerated in the first census is counted again ten years later, and it is a simple matter to calculate the apparent fraction of each cohort surviving the decade. Thus the ratio of persons 20 to 24 in the later census to those 10 to 14 in the earlier is equivalent to ${}_5L_{20}/{}_5L_{10}$ in a life table representing the mortality risks of the intercensal decade. A sequence of life table values can be based on the sequence of calculated census survival ratios, and by well-tested actuarial procedures, a life table can be constructed for ages above five—provided that the two censuses achieved accurate coverage of the population, and that ages were accurately recorded. However, this procedure does not yield estimates of survival rates in infancy and childhood unless the number of births during the intercensal decade has been recorded. Therefore, when, as is usual, adequate records of births are lacking, a complete life table can be based on census survival rates only by estimating infant and child mortality indirectly—for example, by assuming a typical relationship between mortality rates under age five and rates for persons over five.

There is a substantial literature on the construction of life tables from census survival rates, and the method has been applied to data from India, Egypt, Brazil and other countries in Latin America.[1] The demographers and actuaries using census survival rates have been forced to adjust the original census age distributions before calculating survival, or to adjust the rates after calculation, because of the usual effects of age-misreporting combined with differential omission by age. Unadjusted survival rates that are over one or that are absurdly low are common. The methods of adjusting the age distributions (or the raw survival rates) to remove the effects of age-mis-reporting are essentially arbitrary, and when the reported age distributions are seriously distorted, the age pattern of mortality embodied in the estimated life table contains a strong component of the smoothing procedure used as well as of the actual age schedule of mortality.

Once a life table—however approximate it may be in form above age five, and in level of childhood mortality—has been constructed from census survival rates, it can be combined with additional data from the two censuses to provide estimates of birth and death rates during the intercensal interval. A good approximation to the death rate (on the assumption that the censuses are accurate and the life table valid) can be obtained by calculating the average of the two age distributions and applying the m_x values from the life table. The birth rate can then be estimated by adding the average annual rate of increase to the estimated death rate.

1. *Model life tables*

The two problematic aspects of constructing a life table from survival rates are the estimation of infant and child mortality, and the determination of the age pattern

[1] For general methodological discussion, practical application and for further references to earlier writings, see Clyde V. Kiser, "The Demographic Position of Egypt", *The Milbank Memorial Fund Quarterly*, vol. XXII, No. 4 (October 1944), pp. 383 to 408; Giorgio Mortara, *Methods of Using Census Statistics for the Calculation of Life Tables and Other Demographic Measures (with Application to the Population of Brazil)*, see United Nations publication, Sales No.: 50.XIII.3; Kingsley Davis, *The Population of India and Pakistan* (Princeton, Princeton University Press, 1951) pages. 238 to 242; Hugh H. Wolfenden, *Population Statistics and Their Compilation* (University of Chicago Press, 1954), pp. 115 to 117, and Jorge Somoza, "Trends of Mortality and Expectation of Life in Latin America", *The Milbank Memorial Fund Quarterly*, vol. XLIII, No. 4 (October 1965), part. 2, pp. 219 to 233.

of mortality above childhood from data distorted by age-misreporting. A convenient solution to these difficulties is provided by model life tables once suitable models have been calculated and published. An abridged set of such tables is reproduced in annex I. The tables are briefly described in this section, and their use with census survival rates outlined in the following section.

It has often been observed that the mortality risks experienced by different age-and-sex-defined segments of a population are interrelated: i.e., if death rates are relatively high among (for example) middle-aged women in a given population, the normal expectation is that infant mortality is also relatively high. There is a great deal of statistical evidence in support of this commonsense relationship—a relationship expressing the fact that when health conditions are especially good or especially poor for one group in the population, conditions tend to be good or poor for other groups as well.

The result of a tendency for death rates experienced by different groups to be uniquely related would be that death rates for all age-sex groups but one could be precisely estimated from knowledge of the mortality of that one group. It would then be possible to construct a set of life tables that stated the proportions surviving from birth to each age under mortality conditions ranging from the highest to the lowest death rates observed in human populations, and a life table appropriate to a given population could then be chosen (with interpolation, if necessary) from this set of model tables.

Of course the mortality experienced by different populations is not in fact so perfectly uniform. Although there is a strong general tendency for relatively high rates to occur among all segments of a population if they occur in any, there are populations with especially high (or low) rates at certain ages, for one or both sexes. Various approaches have been tried in efforts to express in analytical or tabular form the variety of frequently observed sex and age patterns of mortality.[2] The most widely used model tables are those previously published by the United Nations, and these might have been (but were not) employed in this *Manual;* instead, a set of model life tables rather closely resembling the earlier United Nations tables were employed—a set based on a large body of accurately recorded national mortality experience conforming closely to a single pattern of death rates by age. This group of model tables is one of four calculated at the Office of Population Research, Princeton University.[3] The selection of this set of model tables—extracts from a more extended set appear in annex I—for use in this *Manual* was based primarily on convenience. The model life tables reproduced in annex I are accompanied in annex II by a set of model stable populations that include a wide range of useful parameters, and it is the existence of these auxiliary tables, already calculated on an electronic computer, that dictated the use of this set of model life tables rather than the earlier United Nations tables. In most instances estimates based on these model life tables are little different from those that would be obtained from the earlier United Nations life tables; however, the absence of associated model *stable populations* means that estimation from the earlier United Nations life tables would be much more laborious.

This is not the place for an extended description of alternative possible forms of model life tables, nor even of the four families of which one was chosen for use in this *Manual.* Three of the sets summarize mortality patterns characteristic of regions of Europe, and the fourth—the one partially reproduced in annex I, the so-called "West" family—expresses an age pattern of mortality common to twenty-one countries (Australia, Canada, Israel, Japan, New Zealand, South Africa, Taiwan, the United States and thirteen in western Europe). The age specific mortality rates in this set of model tables are matched quite closely at the appropriate mortality level by the published life tables of these twenty-one rather diverse populations. The 125 life tables for each sex from which these model tables were calculated were selected because they showed no systematic tendency to deviate from a preliminary set of model tables designed to express median recorded world experience. In contrast, each of the other families of model tables was based on regional patterns of consistent and persistent deviation from average world age patterns of mortality. For example, one regional set is based on life tables from contiguous populations in central Europe with a persistent tendency towards unusually high infant mortality at each level of adult mortality.

The model life tables in annex I can be logically employed to construct an approximate schedule of mortality for a population with an unknown age pattern of mortality provided there is no specific evidence of an unusual pattern. Of course the absence of specific evidence of an unusual mortality pattern does not imply that the "usual" pattern in fact prevails. It is highly probable that, if accurate records of mortality existed for all populations, patterns of deviation would be found different from and more extreme than in the four regional families. Nevertheless, the existence of high intercorrelations among mortality rates at different ages, and the existence of patterns to which many mortality schedules closely conform provide the soundest empirical basis of estimation available at present.

2. *Selection of a model life table consistent with census survival rates*

If two consecutive censuses of a closed population were perfectly accurate, and if the mortality schedule expressing average experience during the interval conformed exactly to the model life tables in annex I, the appropriate table could readily be located by comparing the values of ${}_5L_{x+10}/{}_5L_x$ in the model tables with the corresponding survival rates calculated from the two censuses. Under these hypothetical circumstances all of the survival rates would fall between the values for the same pair of model tables, and an intervening model table could be constructed by interpolation. These circumstances are, for instance,

[2] See *Age and Sex Patterns of Mortality,* Model Life Tables for Under-Developed Countries (United Nations publication, Sales No.: 55.XIII.9); *Methods for Population Projections by Sex and Age,* Manual III United Nations publication, Sales No.: 56.XIII.3); and other tabulations cited in the work listed in foot-note 3.

[3] A.J. Coale and Paul Demeny, *Regional Model Life Tables and Stable Populations* (Princeton, Princeton University Press, 1966).

rather closely approximated by the censuses of Korea before the Second World War as indicated by the comparison of female survival rates from 1925 to 1935 in that country with survival rates found in model life tables shown in figure I.

In most censuses of populations without vital statistics (so that mortality must be estimated) the recorded survival rates are highly erratic in the mortality level they indicate because of the effect of age–misreporting. As an illustration of such situations figure I also shows female survival rates calculated from Indian and Turkish censuses[4]. One way of estimating the level of mortality from a sequence of erratic rates based on erroneous age distribution would be to determine the level implied by the survival rate of each cohort, and then to take some sort of average of the levels so indicated. Specifically, the mortality levels in the model tables indicated by the census survival rates of persons 0-4, 5-9, ..., 40-44 in the earlier census could be determined, and ranked from highest to lowest indicated level, and the median selected as the final estimate. (It is prudent to avoid survival values for persons above sixty or seventy because of the prevalence of systematic age–misreporting among older persons, and the possibility of special age patterns of mortality.) In fact, as figure I indicates, the individual survival values are often so affected by age–misreporting that many of the survival

[4] Note the striking similarity in *pattern* in the sequence of apparent survival rates for Turkey and India, doubtless reflecting a basic similarity in the pattern of age–mis-statement. See section B.3 below.

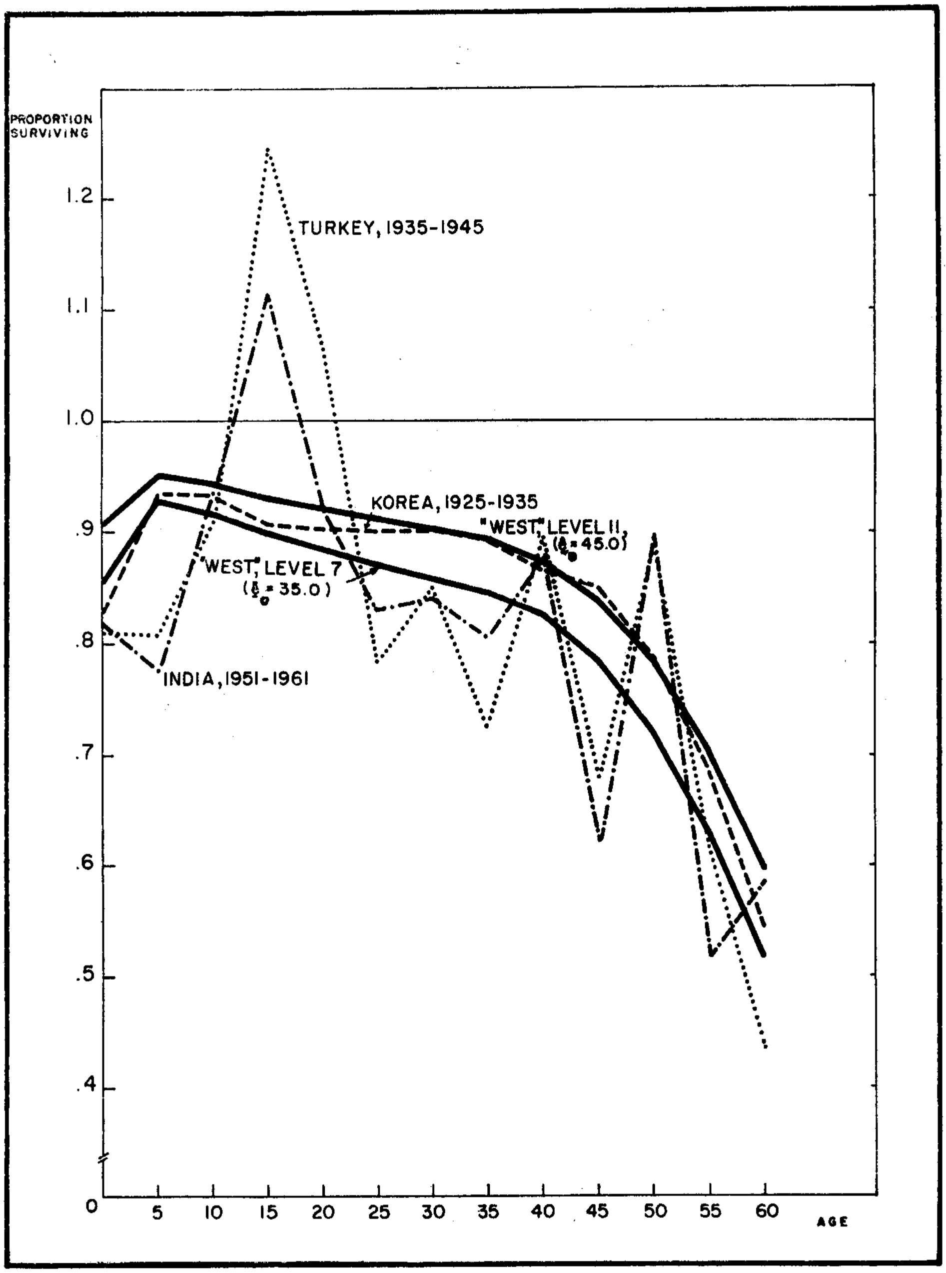

Figure I. Census survivorship rates of females from age x to $x+5$ at time t to age $x+10$ to $x+15$ at time $t+10$, according to censuses of India, Korea and Turkey, and according to selected "West" model life tables rates are outside the limits of the model tables (which extend from an expectation of life at twenty to one of seventy-five years), and little confidence could be attached to the median of such a wildly erratic sequence. It is better to take advantage of the dampening effect that cumulation has on age-misreporting, and to try to determine the level of mortality from the proportions surviving from the entire earlier population to age ten and over in the later census, the proportion five and over that survives to age fifteen and over, and so on. Survival rates of the latter type (calculated from the same census data from which the cohort survival rates plotted in figure I were obtained) are shown in figure II. Naturally, unlike simple cohort survival rates, such rates cannot be directly expressed in terms of mortality levels since their value is significantly influenced not only by mortality but also by the age distribution of the population in question. The computational procedure that permits the translation of these rates into mortality levels requires that the initial population be projected to the later date by applying the survival rates of model life tables at different levels of mortality, e.g., levels 3, 5, 7 and 9. Each projection yields a ratio of the surviving population over ten to the initial population, of the surviving population over fifteen to the initial population over age five, etc. By comparing the recorded survival ratios with those obtained by projections with alternative model tables, one obtains a series of estimates of the level of mortality—a series consistent with the recorded survival of the whole population, the population five and over, ten and over, etc., rather than

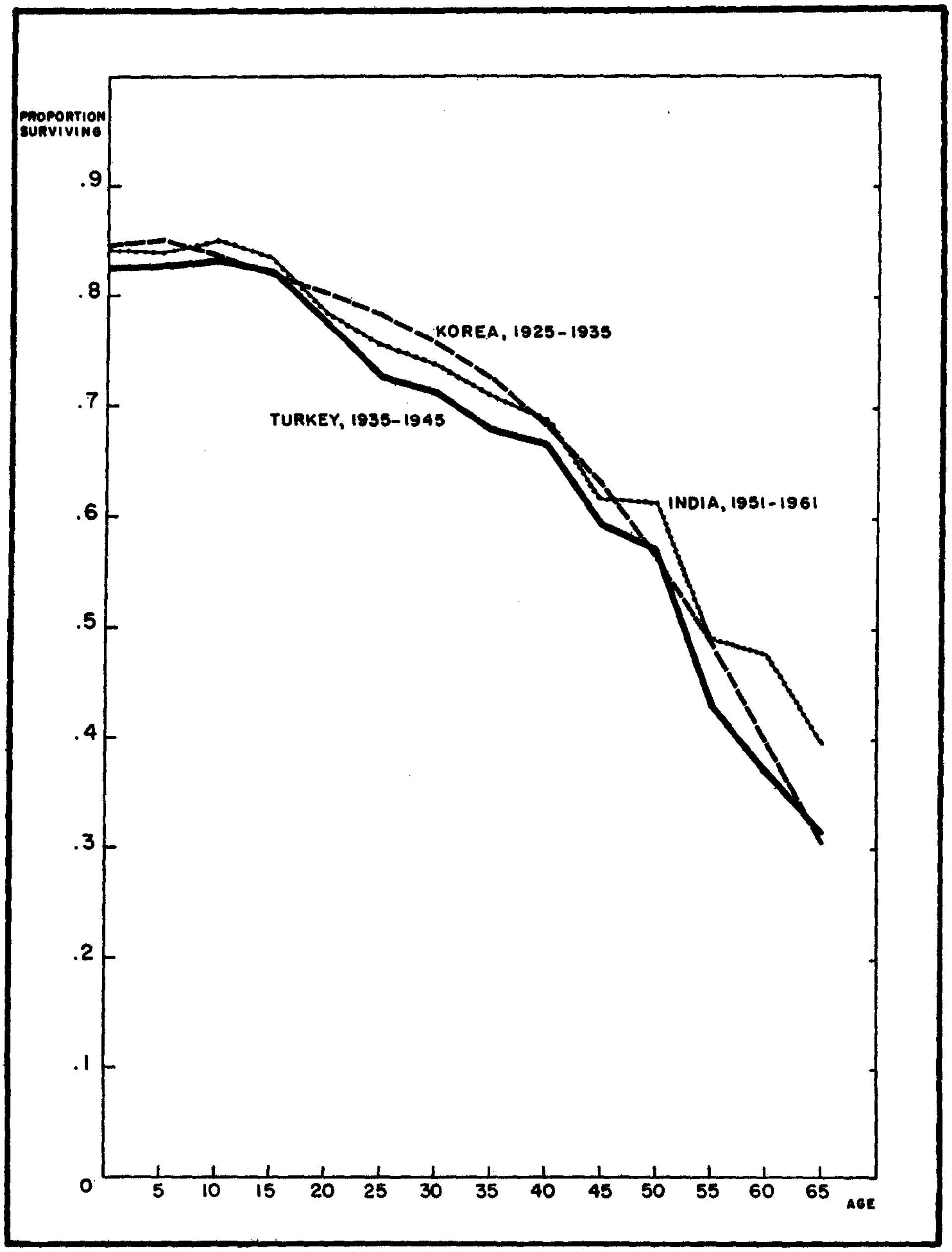

Figure II. Census survivorship rates of females from age x and over at time t to age $x+10$ and over at time $t+10$ according to censuses of India, Korea and Turkey with the recorded survival of individual cohorts.[5] The sequence of mortality levels obtained in this manner is much less affected by age-misreporting than the series based on the recorded survival of individual cohorts.

The reduced effect of age-misreporting is seen in a comparison of figure III and figure I. Determining the level of mortality by examining the survival rates of large segments of the age distribution minimizes the distorting effect of age-misreporting because the survival rates are not affected by age misreporting *within* the groups whose survival is calculated. For example, the survival ratio of the population over twenty in the later census to the population over ten in the earlier is distorted by age-mis-statements that transfer persons across twenty in the later census and across age ten in the earlier, but is unaffected by all other forms of age-mis-statement.[6]

This method of estimating the level of mortality produces a series of alternative figures that are confined to a narrow range when age-misreporting is mild, but that vary extensively when age-misreporting is extreme. In section B of this chapter stable populations methods are used to show that certain populations (including many in tropical Africa, some in northern Africa and the Near East, plus India, Indonesia and Pakistan) have census age distributions by five-year intervals that are quite substantially distorted by age-misreporting, in a pattern that has many common features. In contrast, censuses in the Philippines and Latin America have five-year age distributions that are much less distorted, and censuses in parts of Asia, including China (Taiwan), the Republic of Korea and Thailand have five-year age distributions that appear distorted only to a minor extent by age-misreporting.

The large distortions in the African-Indian-Indonesian-Pakistani censuses mean that the application of the method described above of determining the level of mortality from census survival rates produces a series of estimates with a rather wide range, and with a characteristic sequence of ups and downs. The sequence is consistent with the characteristics of age-misreporting (overestimation of the age of late adolescent girls and young women, for example) discerned by stable population analysis. This analysis suggests that certain survival rates are overstated and others understated for these populations, and that the level of mortality estimated from such rates is biased. Even when ratios with predictable biases are discarded, the remaining census survival rates may indicate levels of mortality that would lead to estimates of a death rate differing by several points. The range of mortality levels consistent with census survival rates in Latin America is typically much smaller and in censuses little disturbed by age-misreporting—such as in Korea from 1925 to 1940—the mortality levels indicated by different survival ratios are confined within a narrow range. The most satisfactory rule of thumb appears to be the selection of the median level of mortality indicated by the propor-

[5] A worked-out example of this method is given in chapter VI.

[6] Age-misreporting *can*, however, affect the number of survivors projected with a given life table, and hence influences the level of mortality that matches the reported number surviving. Specifically, overstatement of age for persons past middle age reduces the projected number of survivors, and leads to the selection of a model life table with too low mortality rates, or too high an expectation of life at birth. This effect is important only in the projection of the population that is older than forty and over in the earlier census.

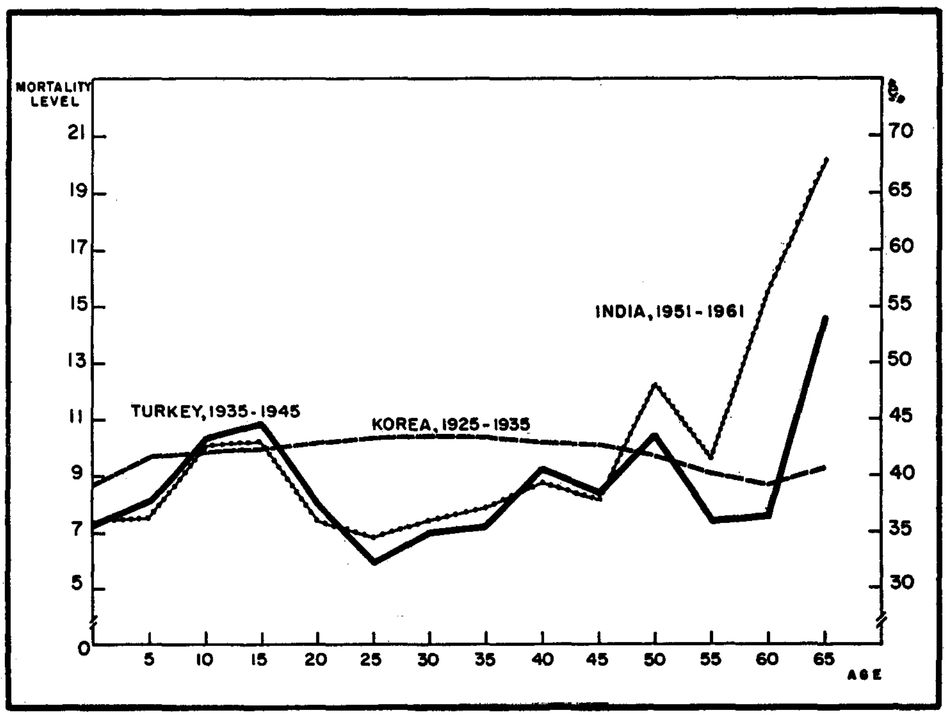

Figure III. Mortality levels in "West" model life tables implied by female survival rates from age x and over at time t to age $x+10$ and over at time $t+10$ according to censuses of India, Korea and Turkey tion surviving among the first nine groups—i.e., all persons, persons five and over, ten and over, ..., forty and over.

This procedure permits the selection of a model life table consistent with the proportions recorded as surviving from one census to the next. It is possible to use the same procedure to select a table from other families of model life tables embodying other age patterns of mortality. It is interesting to note that experimental comparisons of life tables chosen in this way show that the expectation of life at ages five, ten and fifteen in model tables selected from the four families differs very little but that the expectation of life at birth varies widely among the life tables selected. Also, the death rates estimated by applying the selected model table to an estimated mid-decade population differ widely from one family of model tables to another, while the estimated death rates for the population over age five are very nearly the same. The reason for these similarities and differences is that the census survival method establishes essentially the mortality of the non-child population—the population over age five—and that therefore the life table selected from any family of model tables must have the over-all non-child mortality indicated by census survival.

Consider the mortality indicated by a comparison of the population over ten with the whole population ten years earlier. The difference between these populations is precisely the deaths that occurred during the decade to the persons alive at the time of the earlier census. At the midpoint of the decade, this population "at risk" is the population five and over, so that the survival rate for the whole population is very closely linked to the average death rate for the population past age five. All model life tables that give a projected population ten and over equal to that recorded at the end of the decade must connote about the same death rate for the population over five. Since the expectation of life at age five is the reciprocal of the death rate over age five in the stationary population, it is not surprising that the different model life tables have nearly equal $^0e_5$ s.

In other words, the use of census survival rates to choose a model life table at a fitting level of mortality permits the estimation of the death rate over age five with some confidence. However, the death rate for the whole population is strongly affected in populations of high fertility and moderate to high mortality by death rates among infants and young children. The death rate obtained by applying the age specific mortality rates in the selected model table to the estimated mid-period population is valid only if the relation of infant and child mortality to mortality above age five in the family of model tables matches the relation in the population in question. And this relationship is far from the same in the four families of model life tables.[7] Thus census survival rates (if derived from accurate enumerations not excessively distorted by age-misreporting) establish the level of mortality for persons five and over, but leave uncertain (unless the age pattern of mortality is known) the level of child mortality, and thus the expectation of life at birth and the over-all death rate. Since this procedure yields an estimate of the birth rate by adding the intercensal rate of increase to the estimated death rate, the uncertainty surrounding the level of infant and child mortality affects the estimated birth rate precisely as it affects the estimated death rate.

[7] For a detailed discussion of this point see chapter IV.

B. Estimation of fertility and mortality by stable population analysis when fertility and mortality have been constant

Populations subject to approximately constant mortality and fertility schedules come to have the age composition characteristic of Alfred J. Lotka's *stable populations.*

This *Manual* is not the place for a summary of the extensive literature on stable populations;[8] instead it tries only to show when and how useful estimates of fertility and mortality can be based on stable population theory. The question of *when* estimates can be based on stable analysis is easily answered in principle: whenever fertility has been subject to no more than low-amplitude and short duration variations during the previous five or six decades, and mortality has changed only slightly and and gradually during the past generation. The approximate constancy of fertility is a very common, if not universal, feature of populations that are mainly agricultural, and low in literacy and income, except when fertility has been affected by wars, revolutions, major epidemics or other such episodes. The absence of major trends in mortality has also been a common characteristic of less developed areas until the past few decades when very rapid declines in death rates have been frequent. In this section the use of stable population techniques is discussed in those instances where it is clearly appropriate, namely, when there have been no major trends or fluctuations in fertility, and no sustained important changes in mortality.

Stable population analysis has also been applied by demographers to populations whose mortality has been declining, although it has been demonstrated that the resultant estimates are biased. In section C of this chapter a method of adjustment is described for altering the estimates derived from stable analysis to compensate for the effect of a history of recent decreases in mortality.

A stable population is generated by the continuation of a fixed schedule of fertility and a fixed schedule of mortality; it is characterized by an unchanging proportionate age distribution, and a constant annual rate of increase. In populations essentially closed to migration where there is no evidence of spreading use of deliberate birth control or of changing patterns of nuptiality, and no reason to believe that mortality is changing rapidly, confirmation of the conjecture that the population may be stable can be sought in comparisons of the recorded age distributions in successive censuses, and of successive intercensal rates of increase. If the census age distributions show marked differences (e.g., such as are seen in the censuses of Turkey from 1935 to 1960), it is probable

[8] The fundamental work on the subject is Alfred J. Lotka's *Théorie des associations biologiques*, deuxième partie (Paris, Hermann et Cie., 1939). For a comprehensive treatment see the United Nations Study entitled *The Concept of a Stable Population. Application to the Study of Populations of Countries with Incomplete Demographic Statistics* (Sales No.: 65.XIII.3).

that fertility has not been constant, in the cited example because the series of wars and other disturbances experienced by the Turkish population undoubtedly caused major fluctuations in fertility, and in mortality as well, at least for males. When the rate of increase rises markedly, it is likely that mortality is falling. But essentially constant age distributions and rates of increase observed in a series of censuses can be considered justification for considering the population stable.

The age distribution of a stable population is described by a well-known formula of Lotka's:

$$c(a) = be^{-ra}p(a) \tag{1}$$

where $c(a)$ is the proportion of the population at age a, b is the birth rate of the stable population, e is the base of natural or Naperian logarithms, r is the annual rate of increase, and $p(a)$ is the proportion surviving from birth to age a according to the prevalent mortality risks. The proportion $p(a)$ is an alternative expression for the survivor function (l_x/l_0) in the life table, and all that is needed to fix the proportion at every age in a stable population is the life table expressing the constant mortality conditions and the average annual rate of increase. The birth rate b is determined by the fact that the sum of the proportions at all ages must be equal to one.

In England and Wales in 1881 the age composition of the female population was much the same as it had been in the two preceding censuses, and the estimated intercensal rate of natural increase had been nearly constant. To be sure, the population was not closed, primarily because of a flow of emigrants to America, but between 1871 and 1881 the rate of average annual loss for females was only .00070, so that the effect on the age distribution was minor. The theory of stable populations leads us to expect, then, that the age composition of the female population of England and Wales in 1881 is closely approximated by a synthetic age distribution calculated according to formula (1), using the life table for 1871-1881 to provide $p(a)$, and equating r to the average annual intercensal rate of natural increase in that decade. Figure IV provides a comparison of this calculated age distribution with that recorded in the censuses of 1881. and gives a demonstration by example that theoretically

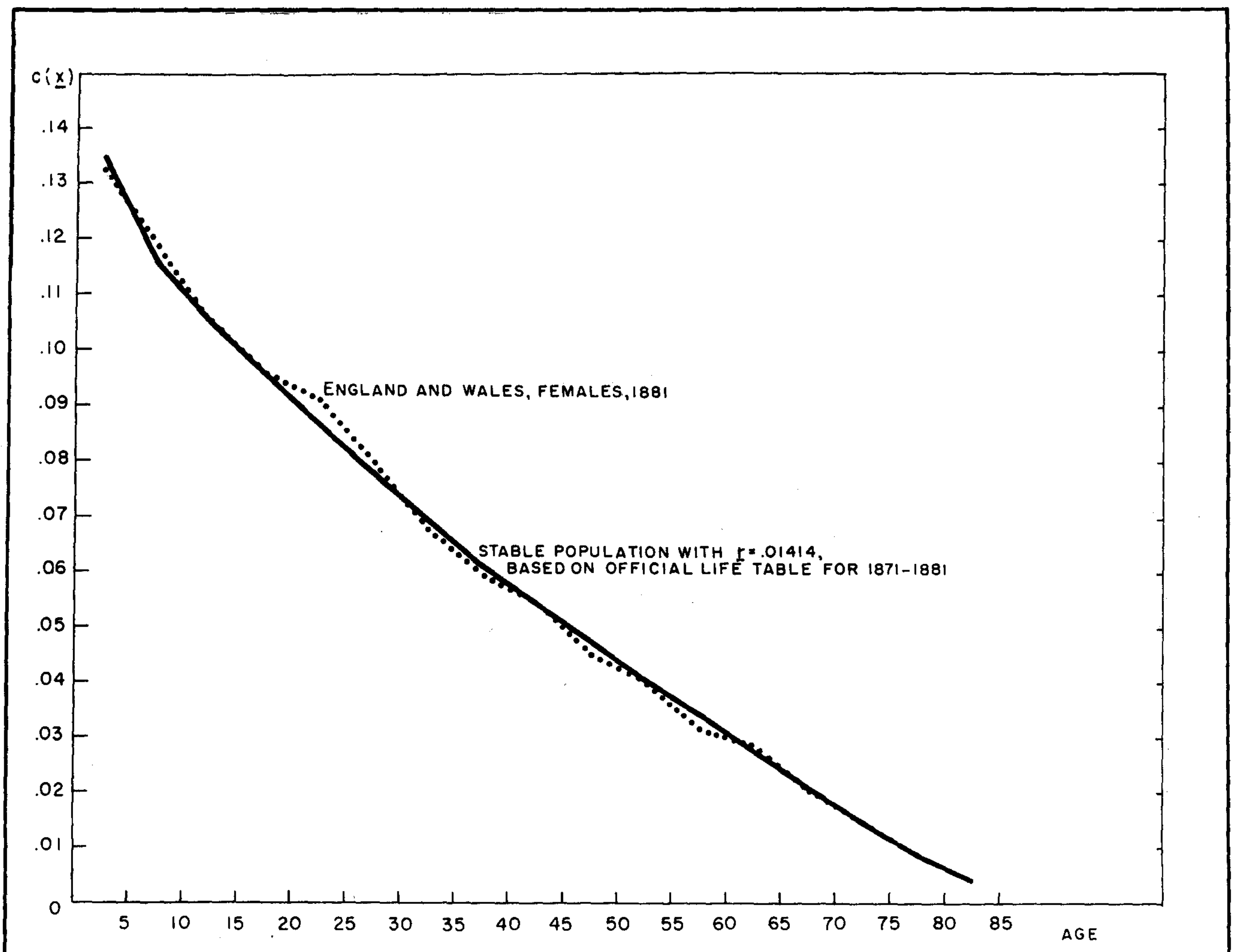

Figure IV. Female age distribution in England and Wales, by five-year intervals, as recorded in the census of 1881 and as approximated by a stable population constructed on the basis of the intercensal (1871-1881) rate of natural increase (r = .01414) and the official English life table for the same period, both for females derived stable age distributions fit actual populations very closely.

This example does not show the potential usefulness of stable populations as a way of estimating demographic variables, because when data exist to calculate an accurate life table, as in England and Wales in the 1870s, indirect estimation is not required. Suppose, therefore, that death registration had been non-existent or very incomplete. Could the theory of stable populations and the near stability of the population of England and Wales be used to estimate mortality and fertility? A positive answer is provided through the device of model stable populations based on model life tables.

1. *Model stable populations*

As is indicated by equation (1) above, the age distribution of a stable population is jointly determined by the mortality schedule (or life table) and the annual rate of increase. In the preceding section there is a description of *model life tables* which can be constructed to embody typical age patterns of mortality at different mortality levels. A set of such model tables embodying what can be considered "normal" or "typical" world mortality experience is reproduced in abridged form in annex I. Corresponding to each model life table is a set of possible stable populations at rates of increase corresponding to the highest and lowest levels of fertility that might accompany the life table. Annex II presents such a set of stable populations for each model life table, with rates of increase ranging from $-.010$ to $.050$.

This set of model stable populations includes a range of age distributions bracketing all those likely to be found in actual populations that (*a*) are themselves approximately stable (i.e., have had approximately constant fertility and mortality), and (*b*) experience mortality where the age pattern conforms more or less closely to that embodied in the set of model life tables. If an actual population appears, in the light of observed characteristics, to belong to the family of model stable populations, it is possible to locate the model stable population matching the actual one, and to estimate various demographic parameters of the actual population by attributing to it the parameters of the model stable population.

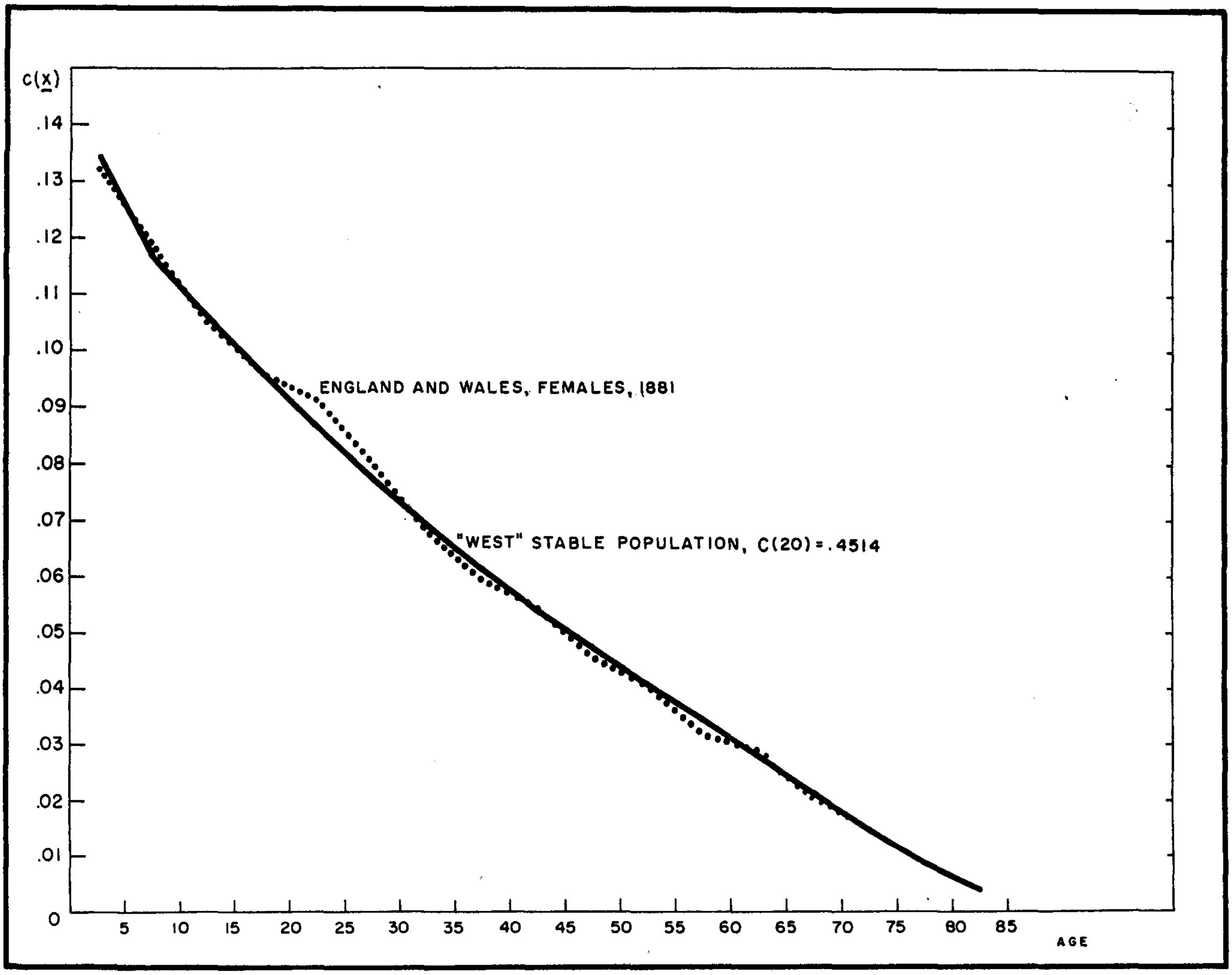

Figure V. Female age distribution in England and Wales, 1881, by five-year intervals, as recorded in census and as approximated by "West" female stable population constructed with the same proportion under age twenty as recorded in 1881 and with the intercensal (1871-1881) rate of natural increase of the female population ($r = .01414$)

In the comparison of a calculated stable age distribution and a recorded age distribution represented in figure IV, the life table and the rate of increase were known, and the stable age distribution was calculated. When stable population analysis is used as a method of estimation, the process is partially reversed. The age distribution is known, as is some other parameter such as the rate of increase, and this knowledge is used to locate the appropriate stable population among the family of stable populations, and the birth rate, death rate, expectation of life at birth and other characteristics tabulated for the model population are ascribed to the actual. (The details of the method are illustrated in the examples given in part two, especially in the example where a model stable population is fitted to the female population of England and Wales enumerated in the 1871 census.) Figure V shows a comparison of the 1881 age distribution with a model stable with a growth rate equal to the estimated average yearly rate of natural increase between 1871 and 1881 and the same proportion under age twenty as recorded in the census. It is essential to note that this is *not* the same stable population shown in figure IV. Both incorporate the same rate of increase, but one incorporates the recorded female life table of the decade and does not involve the 1881 age distribution in any direct way, so that its virtual identity with this distribution can be taken as confirming the stable nature of the actual population. The other calculated stable population, in contrast, is not based at all on the recorded English life table, but on the recorded rate of increase, the presumed stability of the population, and the assumption that the age pattern of mortality conforms to this family of model life tables. The model stable population selected in this manner has an expectation of life at birth within .3 of a year of the recorded value, and a birth rate apparently closer to the actual figure than the registered birth rate, because of a slight under-registration of births in England and Wales during the decade in question.

2. *Selecting a model stable population on the basis of an accurately recorded age distribution*

Characteristics of an actual population can be estimated by locating the model stable population that best fits certain recorded or calculated features of the population in question, and then assigning the characteristics of the model stable to the actual population. Relevant features of the actual population that are usually recorded include its age distribution and the average rate of natural increase between two censuses estimated by adjusting the average annual intercensal rate of increase for any known difference in completeness of coverage, or for any known minor rates of net gain or loss through migration[9]. The problem of selecting the appropriate model stable population usually reduces, then, to finding a stable population with a given annual rate of increase and most closely resembling the recorded population in age composition.[10]

How can one judge whether a stable age distribution "closely resembles" a recorded age distribution? If the recorded age distribution is generated by genuinely constant fertility and mortality, if the coverage of the censuses is complete, if the population is not substantially affected by migration, and if ages are accurately reported, various alternative features of the recorded age distribution would serve to locate essentially the same model stable population. The proportion under age ten, or five to fourteen years of age, or over age fifty would lead to the selection of much the same stable distribution. For example, the population of England and Wales in 1881 fits the model stable population remarkably well. But age-misreporting in the 1881 census was limited, and the age pattern of English mortality conforms quite closely to the so-called "West" family of model life tables.

In many of the less developed countries the age distribution reported in censuses or demographic surveys is affected by gross misreporting of age. The distribution by single years of age is very often conspicuously distorted by "age heaping"—by a tendency for many persons to report a preferred nearby number (one ending in zero or 5, for example) rather than the correct age. Indeed it appears likely that often many more ages are misreported than given correctly. A reported number of persons at age sixty greater than ages sixty-one to sixty-nine combined is not unusual.

On *a priori* grounds, it is clear that the effects of age-misreporting on the cumulative age distribution (the so-called *ogive*) are less than on the proportion in a particular five-year interval. In fact the proportion under age x (which will be designated $C(x)$) is affected only by those age-mis-statements that cause a net transfer of persons across age x. The proportion reported as under age thirty is not altered if children under age five are reported as five, six or seven, or if persons in their late fifties or early sixties are reported as sixty years old.

One of the advantages of the cumulative age distribution is the simple general relation that exists among the ogives of model stable populations with the same rate of increase: stable populations with higher fertility (and hence higher mortality at the same rate of increase) have greater cumulative proportions to every age than lower fertility stable populations (see figure VI). In other words, the ogives of stable populations with the same rate of increase do not cross, and each model stable population is thus completely determined by knowledge of the rate of increase and $C(x)$ for *any* value of x. Consequently, one stable population is identified by the intercensal rate of increase and $C(5)$, another by the rate of increase and $C(10)$ etc. If the reported population in fact has a stable form, and if age reporting is accurate, the series of stable populations identified with $C(5)$, $C(10)$, ..., $C(65)$, will exhibit little variation, as would be evident in a very tight cluster of ogives of the stable populations thus

[9] If gains and losses because of migration are substantial, the use of stable population analysis becomes questionable.

[10] If the proportion of children surviving to age two is known (from the methods described in chapter II) the problem is one of selecting a model stable population with the given l_2 and most closely resembling the recorded population in age composition.

selected. Deviations from constant fertility (and to a less extent from constant mortality) in the recent past would cause the stable populations identified with cumulative proportions under some ages to differ from those identified with the proportions under other ages. The set of model stable ogives consistent with C(5), C(10), ..., C(65) would then be spread out rather than tightly clustered. The same effect of diverse rather than consistent estimation of the appropriate model stable population would also result from age-misreporting of a sort that caused large net transfers of persons across ages divisible by 5 (cf. figure VII).

Suppose a population has been subject to approximately stable conditions and that its cumulative age distribution is compared with the ogives of stable populations that are consistent with C(5), C(10), ..., C(40) and the intercensal rate of increase.[11] There will be a highest and a lowest cumulative stable distribution, setting upper and lower limits to the choice of a model stable population. If the highest stable ogive is accepted as approximating the true age distribution, it follows that C(x) at all other ages from zero to forty is too low, and that therefore there must have been a net transfer of persons by age overstatement at all ages except that where the ogive agrees with the highest stable. Conversely, acceptance of the lowest ogive implies net understatement of age across all ages divisible by five except one.

[11] It is wise to avoid comparisons at the older ages because differences in age-patterns of mortality have an increasing effect on the cumulative age distribution above age forty or fifty, and because of the prevalence of systematic age-misreporting at older ages.

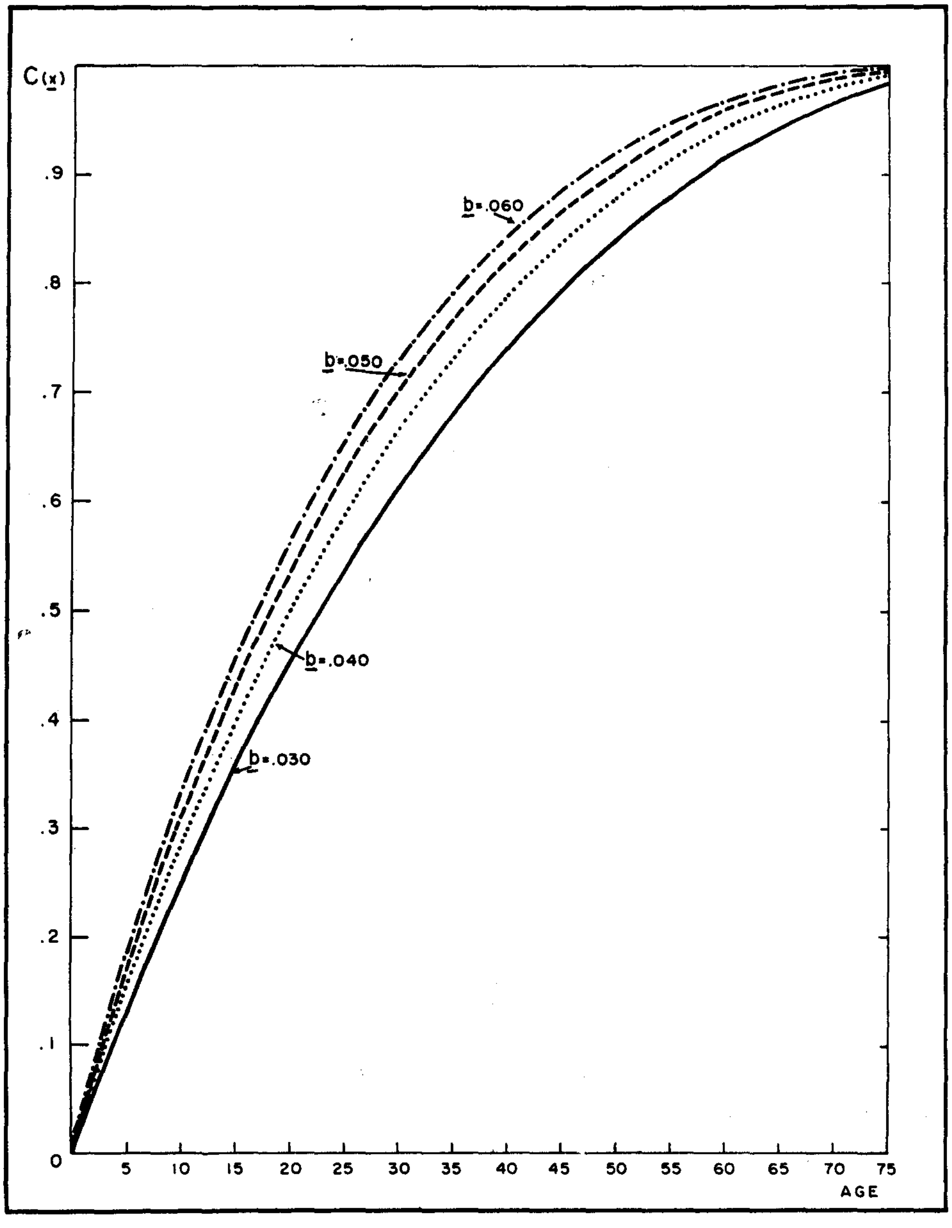

Figure VI. Ogives of age distributions (proportions up to age x) in stable populations ("West" females) with a growth rate of .02 and with birth rates (b) as indicated

These relationships are sometimes useful in detecting "patterns" of age-mis-statement, and are so employed in the next section.

3. *Characteristic forms of age-mis-statement*

We have compared a large number of recorded age distributions with stable ogives as a way of uncovering typical patterns of deviation from stable populations in certain categories of censuses or surveys. The comparison was twofold:

1. The calculation of $C(x) - \bar{C}_s(x)$, where $C(x)$ is the cumulative age distribution of the given (male or female) population, and $\bar{C}_s(x)$ is the middle (median) stable population of those with ogives that agree with $C(10)$, $C(15)$, ..., $C(40)$. The set of stables employed were the "West" model tables of annex II, with ${}^0e_0 = 40$ years.[12]
2. The calculation of

$$\frac{c(0-4)}{\bar{c}_s(0-4)}, \ldots, \frac{c(40-44)}{\bar{c}_s(40-44)},$$

when c(0-4) is the proportion aged 0-4 in the given population, and c_s (0-4) is the proportion aged 0 to 4 in the median stable population defined above.

If the given age distribution conformed exactly to the

[12] The reader may naturally suspect that the employment of ogives of alternative model stable populations with a given life table is very different from using ogives with a given rate of increase In fact, the comparisons obtained by holding ${}^0e_0$ constant at forty years are virtually indistinguishable from those that would be obtained from a fixed value of r (see figure IX.)

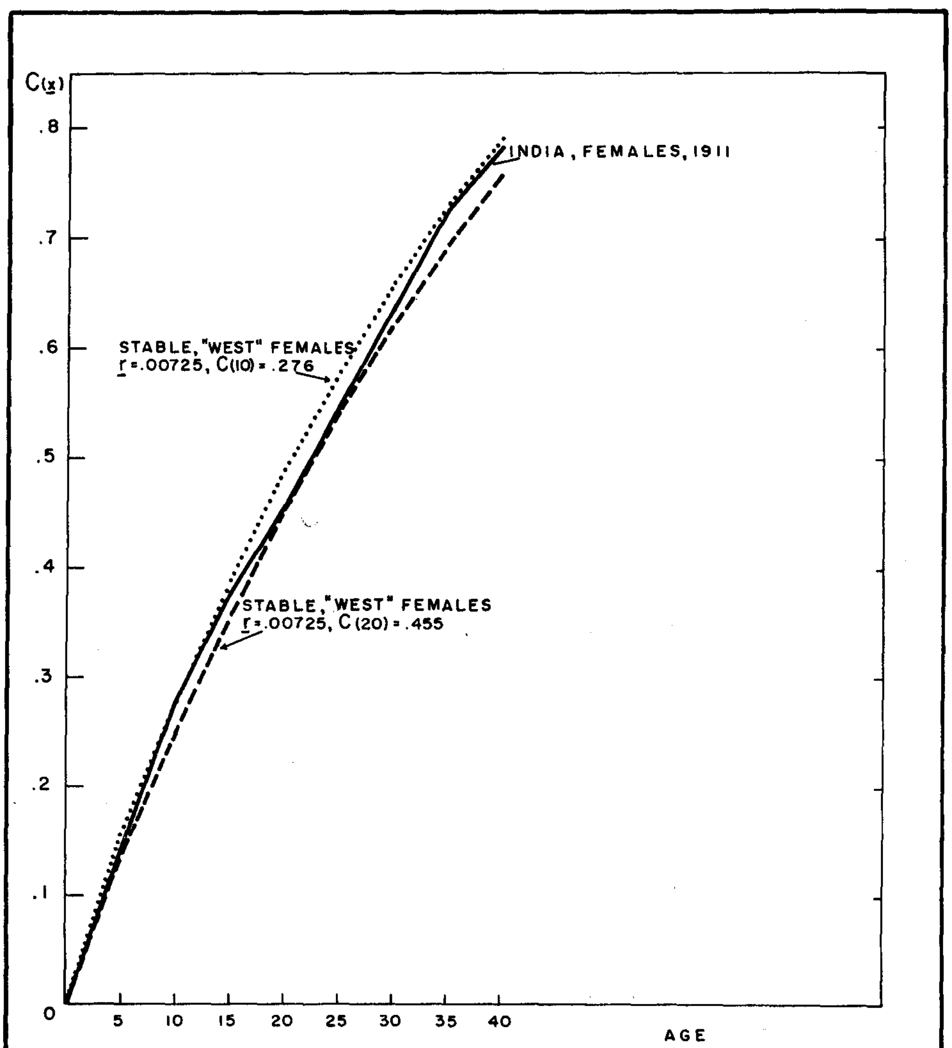

Figure VII. Ogive of the age distribution of the female population of India as reported by the census of 1911 and ogives of the age distributions in "West" female stable populations with a growth rate same as the female intercensal (1901-1911) rate of increase in India (r = .00725) and with the highest and lowest ogives consistent with values of C(5), C(10), ..., C(40) in the census population

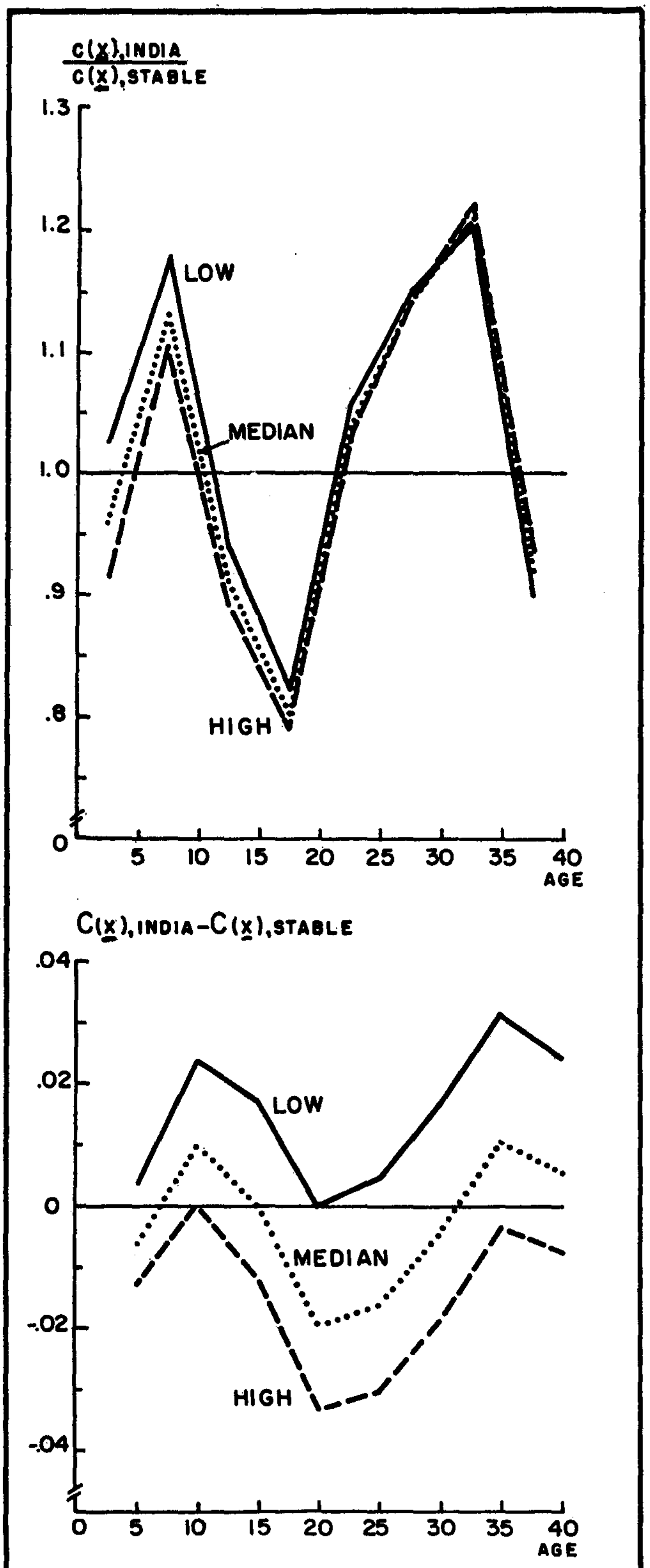

Figure VIII. Comparisons of the reported female age distribution by five-year age groups—c(x)—and its ogive—C(x)—as reported in the 1911 census of India with corresponding values in three model stable populations defined by the Indian female growth rate for 1901-1911 and by agreement with C(10), C(20) and C(15) in the census population resulting in the highest, the lowest and the median ogive respectively among those corresponding to C(5), C(10), ..., C(40)

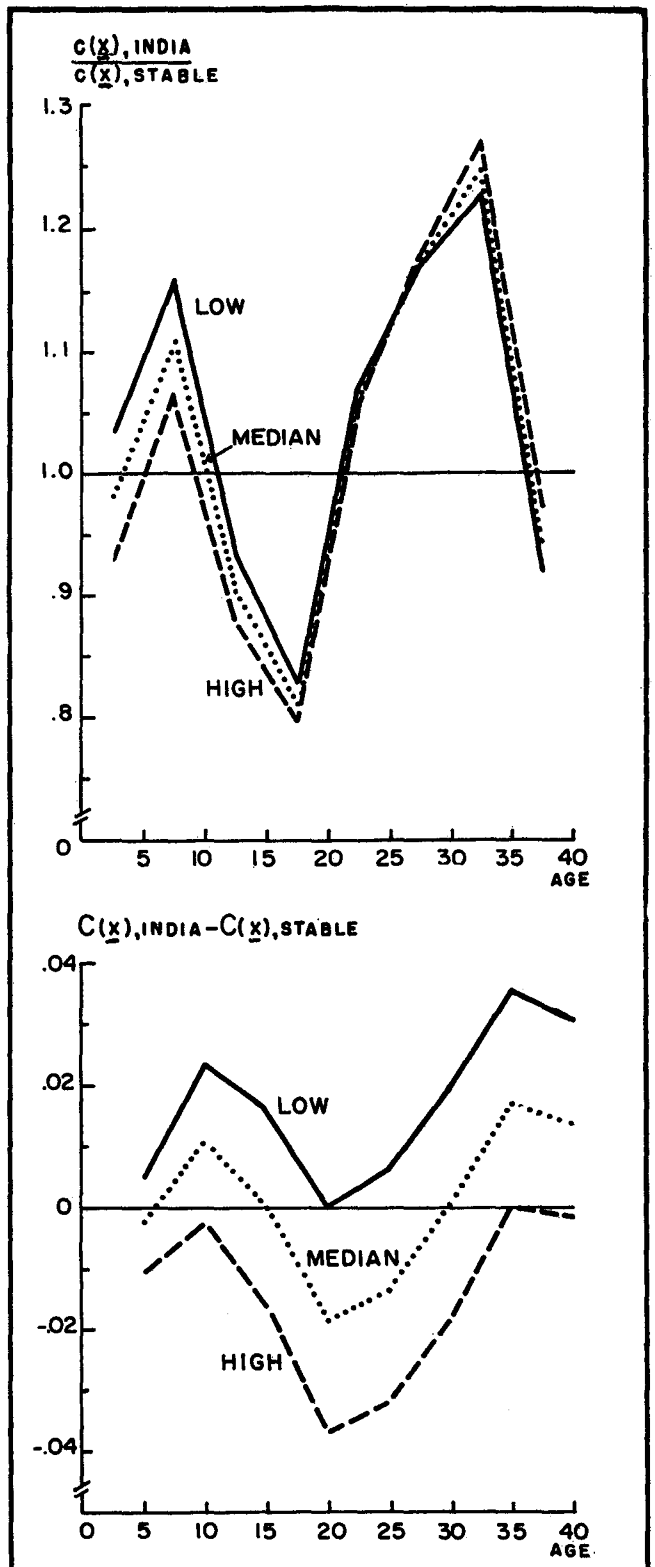

Figure IX. Comparisons of the reported female age distribution by five-year age groups—c(x)—and its ogive —C(x)—as reported in the 1911 census of India with corresponding values in three model stable populations defined by an expectation of life at birth of forty years and by agreement with C(35), C(20) and C(15) in the census population resulting in the highest, the lowest and the median ogive respectively among those corresponding to C(5), C(10), ..., C(40) model stable, $C(x)-\bar{C}_s(x)$ would be zero at each age, and

$$\frac{c(x-y)}{\bar{c}_s(x-y)}$$

would be one in each age interval. A positive value of $C(x)-\bar{C}_s(x)$ implies age understatement that shifted persons across age x, and a

$$\frac{c(x-y)}{\bar{c}_s(x-y)}$$

greater than one implies that age-misreporting has inflated the reported number of persons in the given age interval—both implications following if $\bar{C}_s(x)$ is accepted as a valid estimate of the true age distribution. But figure VIII shows that the sequential pattern of ups and downs of $C(x)-C_s(x)$ and

$$\frac{c(x-y)}{c_s(x-y)}$$

is maintained whether the comparison stable is the highest, the lowest or an intermediate ogive. It is evident in figure IX that the apparently arbitrary choice of stable ogives with ${}^0e_0 = 40$ years has no important influence on the sequence.

The purpose of the comparisons is to bring to light common patterns of deviation from the stable form when approximate conformity to a model stable population might be expected. Over 150 censuses or surveys of populations of each sex thought to have a history of approximately constant fertility were analyzed in this way.

The analysis reveals the existence of certain general patterns of deviation from approximately appropriate model stable populations—patterns in each instance shared by censuses and surveys of several different populations. One pattern is clearly the product primarily of age-misreporting, and others appear to result in large part from past variations in fertility and mortality—the result of departures from the prerequisites of a genuinely stable population, in other words.

Censuses and surveys with extreme age heaping evident in the single-year age distributions could be expected to have ogives and distributions by five-year intervals that do not conform closely to model stable populations. But in fact some censuses and surveys obviously subject to poor age-reporting (by single years) deviate only slightly, and some quite strongly from the expected form of ogives and five-year age distributions. Figure X (upper panel shows

$$\frac{c(x-y)}{\bar{c}_s(x-y)}$$

for females in the Philippines (1960), Colombia (1951), Venezuela (1950) and Ecuador (1950) on the one hand, and in India (1911), Morocco (1960), Ghana (1960) and Indonesia (1961) on the other. All are populations in which age heaping is extensive. The lower panel of the figure compares $C(x)-\bar{C}_s(x)$ for the same censuses. The most conspicuous contrast in the upper panel of figure X is between the large deviations from the stable, with two or more consecutive age intervals deviating in the same sense in one set of censuses, and the more modest deviations, usually with a saw-tooth quality (positive deviations followed by negative ones) in the other set of censuses. This contrast is manifested in a U-shaped sequence of large differences between the ogive of the reported distribution and the model stable in one group of censuses, and a more or less alternating pattern of small differences in the other group. The implication of these contrasting patterns is obvious: in the first group of censuses, there is a systematic form of unidirectional age-misreporting over a broad range that distorts the reported age distributions even as an ogive; in the second group, although age-misreporting causes pronounced age heaping by single years, the distribution by five-year intervals tends to alternate excesses and deficits, and the cumulative distribution is not much distorted.

(a) *Female age distributions with large distortions*

An examination of twenty-nine female age distributions of the sort affected by large-scale misreporting shows the following common characteristics in the pattern of the cumulative age distribution, as revealed by $C(x)-\bar{C}_s(x)$:

1. The cumulative age distribution rises (relative to the stable) from age 5 to 10;[13]
2. It falls from age 10 to 15[13] and from 15 to 20;[14]
3. It rises from 25 to 30,[13] and from 30 to 35.[15]

The proportion in five-year intervals shows the following characteristics, relative to the stable:

1. The proportion 5-9 is above the stable;[13]
2. The proportions 10-14[13] and 15-19[14] are below the stable;
3. The proportions 25-29[13] and 30-34[15] are above the stable.

In other words, the age distributions have a surplus at 5-9, and a deficit in the adolescent age intervals (10-14 and 15-19) followed by a surplus in the central ages of child-bearing (25-34). Censuses and surveys in all of the countries of tropical Africa, in India, Indonesia, Morocco, and Pakistan show this pattern. It is repeated in all of the censuses of India before partition (except the census of 1931, where the published age distribution was smoothed) as well as in both Pakistan and India in 1961. The underlying pattern of age-misreporting can be detected in the censuses of most Near Eastern and North African countries. It is *not* evident in Latin American countries, Ceylon, Taiwan (China), Malaya, the Philippines, the Republic of Korea, or Thailand.

Why should African censuses and surveys show the same kinds of misreporting of age as censuses of India, Indonesia and Pakistan ? A plausible explanation of the similar patterns is that in these enumerations the age entered on the interview schedule was often an estimate made by the interviewer, rather than the transcription of a number supplied by the respondent. In other words the common form of distorted age distributions is caused by the common biases in the estimation of women's ages

13 No exception in twenty-nine censuses or surveys.

14 Only one exception in twenty-nine cases.

15 Two exceptions in twenty-nine cases.

by another person. Unfortunately, this explanation of the similar patterns of distorted age distributions does not by itself indicate what the true age distribution is. It suggests that the age distributions affected are distorted in a similar way, but to determine where there is net overestimation of age and where net underestimation it is necessary to find by some other means which of the alternative model stable populations does in fact resemble the actual age distributions.

There are a few instances where there is an independent basis for selecting the appropriate stable population, and by considering these it is possible to obtain an insight into the typical age-mis-statements prevailing in these instances—and presumably in other censuses and surveys with the same pattern of distortion in the age distribution. In the sample census of the Congo of 1955-1957, the ages of all but a minority of young children were verified by the interviewer through the examination of birth certificates or of entries in the mother's identity booklets, and the stable population consistent with the proportion

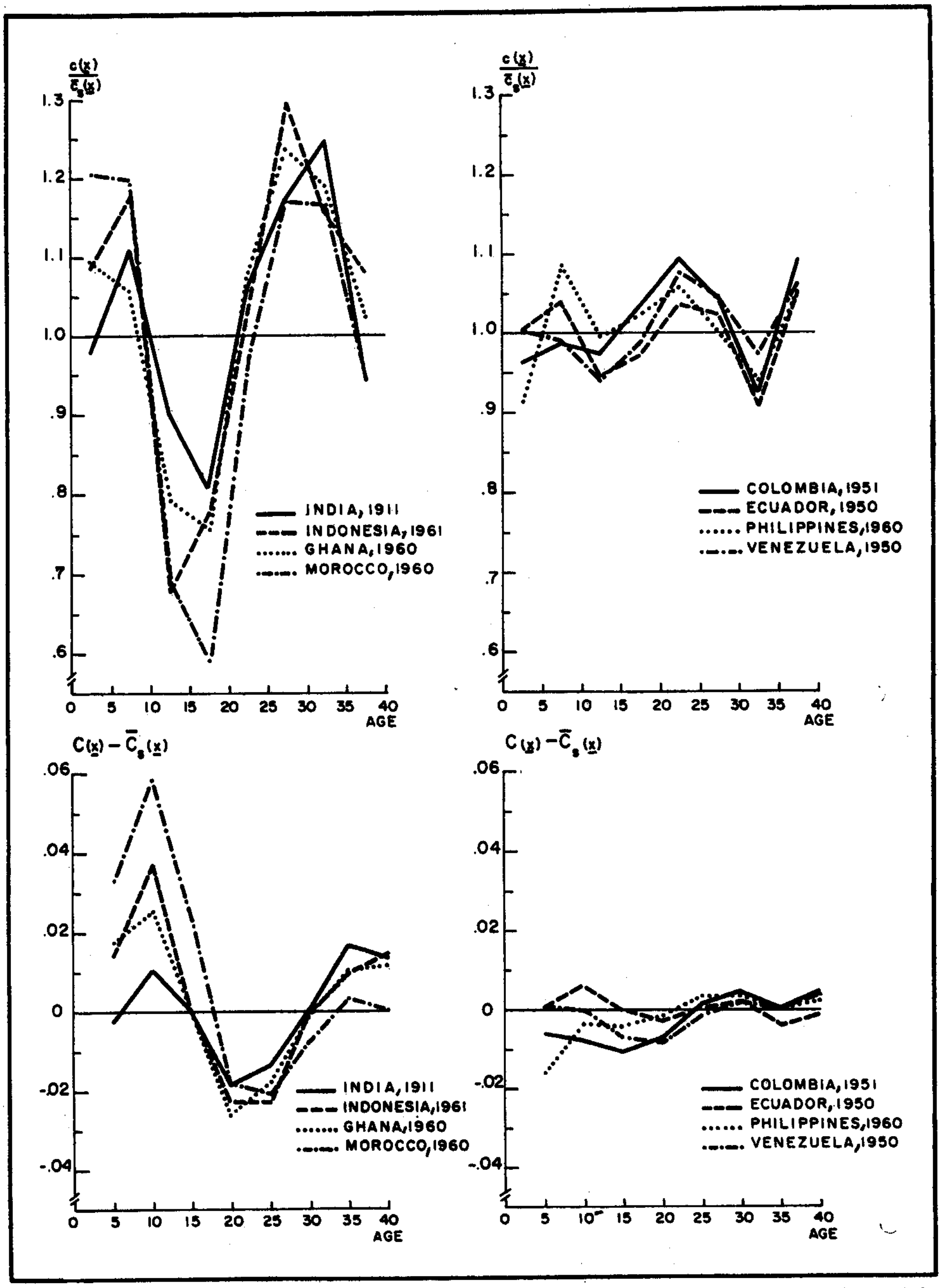

Figure X. Comparisons of the female age distribution by five-year age groups—$c(x)$—and its ogive—$C(x)$—as reported in various censuses, with corresponding values in stable populations median among those defined by an ${}^0e_0$ of forty years and by agreement with C(5), C(10), ..., C(40) in the census populations. The comparisons illustrate two typical patterns of age-misreporting: the "African-South Asian" pattern (left-hand diagrams) and the "Latin American" pattern (right-hand diagrams) of children can be accepted as a valid fit. In several other African territories fertility and mortality could be estimated on the basis of retrospective data on children ever born, children surviving, and births and deaths in the year before the census, and a stable population chosen to conform to these fertility and mortality estimates. The stable populations selected in this way in some instances agreed with the reported cumulated distribution up to age ten, and in other instances had a lower proportion under age ten but a larger proportion under age 15 than the census or survey. If these examples are representative, the distorted age distributions of Africa, India, Indonesia and Pakistan are the result of the following typical errors in estimating the ages of females:

1. A tendency to overestimate the age of young children, contributing to the typical excess proportion at ages 5-9, and the relative deficit at 0-4.

2. A tendency to overestimate the age of girls 10-14 who have passed puberty, especially if they are married, combined sometimes, but not universally, with a tendency to underestimate the ages of girls 10-14 who have not reached puberty, causing a net transfer downwards across age ten, and contributing to the peak at ages 5-9.

3. A tendency toward overestimation like that affecting some of the 10 to 14 year olds, for females 15-19, 20-24, and 25-29, causing net upward transfers across ages 15, 20, 25, and 30, and causing deficits at 10-14 and 15-19, and excessive proportions at 25-29 and 30-34. This overestimation of the age of young women may be caused by an unconscious upward bias associated with marriage and child-bearing, or from a mechanical assumption that women were married at some alleged conventional age at marriage and have then experienced an allegedly typical passage of time between marriage and first birth, and in each subsequent interbirth interval.

(b) *Female age distributions with smaller distortions*

When, on the other hand, the recorded age is supplied by the respondent, the age distribution is naturally less distorted— first of all because when most respondents supply a plausible figure to the enumerators, the maximum error is generally below the level of the ridiculous, and the average error is thus diminished. That is, when respondents are usually prepared to supply an age on request, and then the figure given is acceptable in the sense of being only rarely absurd (e.g. seven for a grown man with a beard), it is plausible that members of the population know their approximate age, and that the broad outline of the age distribution is approximately correct. Rough knowledge of age on the part of most persons does not require a high level of literacy but merely a culture in which numerical age has importance. Apparently ages were accurately known in Sweden in the middle of the eighteenth century, for example. Distortions in the cumulative age distributions are relatively minor in the Philippines and Latin America—even in Honduras, where the proportion of persons over fifteen recorded as illiterate is nearly as high as in Indonesia.

Some of the features of the less pronounced age–misreporting in the populations where apparently ages were supplied by the respondents are just what would be expected from respondent's errors. For example, there may be a reluctance to pass certain milestones, such as age thirty or age forty. The relatively inflated age groups of 25-29 and 35-39 found in several of the censuses of Latin America and the Philippines perhaps include some women who have really passed thirty and forty, respectively. This possibility is reinforced by the relative deficit found at ages 30-34 and 40-44.

(c) *Distortions in male and female age distributions*

Estimations of population parameters can be based on the age distribution of either sex. The stable population methods of estimation described in this section can be applied either to males or females, as can the census survival techniques outlined earlier. The ratio of male births to female births is confined to the limits of 1.04 to 1 to 1.07 to 1 in almost all populations where birth registration is essentially complete, the consistent exceptions being in populations of African origin, where the ratio varies from 1.02 to 1.04. Because of the approximate constancy of the sex ratio at birth, it is possible to check the consistency of estimates derived from male age distributions on the one hand, and female on the other: the estimated male births should exceed the female by about 5 or 6 per cent in non-African populations, and by 2 or 3 per cent in African populations.

The approximate constancy of the sex ratio at birth also makes it possible to base the estimates for both sexes on the analysis of the age distributions of only one. For example, if female births are estimated by stable population techniques, male births can be estimated as 6 per cent more than the female. This estimate divided by the recorded male population gives an estimated male birth rate. The male birth rate less the intercensal rate of increase of the male population gives the male death rate.

The results of analysing age–misreporting in female populations with approximately stable age distributions by comparisons with model stable populations have already been described. The corresponding male populations have been examined in precisely the same manner. In general, age distortions among male populations show similarities analogous to those found among females: one pattern of major distortions in both ogives and five-year distributions is found to characterize surveys in most of Africa and southern Asia while another pattern of substantial age heaping, but relatively minor distortions in ogives and five-year distributions, is characteristic of Latin America and the Philippines. However, among the male populations in Africa-South Asia, the similarities in distortion from country to country are less for males than for females, and the distortions themselves appear larger on the average. In the Latin American populations, the distortions of the female age distributions are almost always larger than in the male; moreover there is usually a slight systematic bias found among the female age distributions (ogives that yield continuously increasing estimates of fertility as age increases from ten to forty) not found in the males.

A summary comparison of the effect of distortions in age distribution on estimation is shown in figure XI. Birth

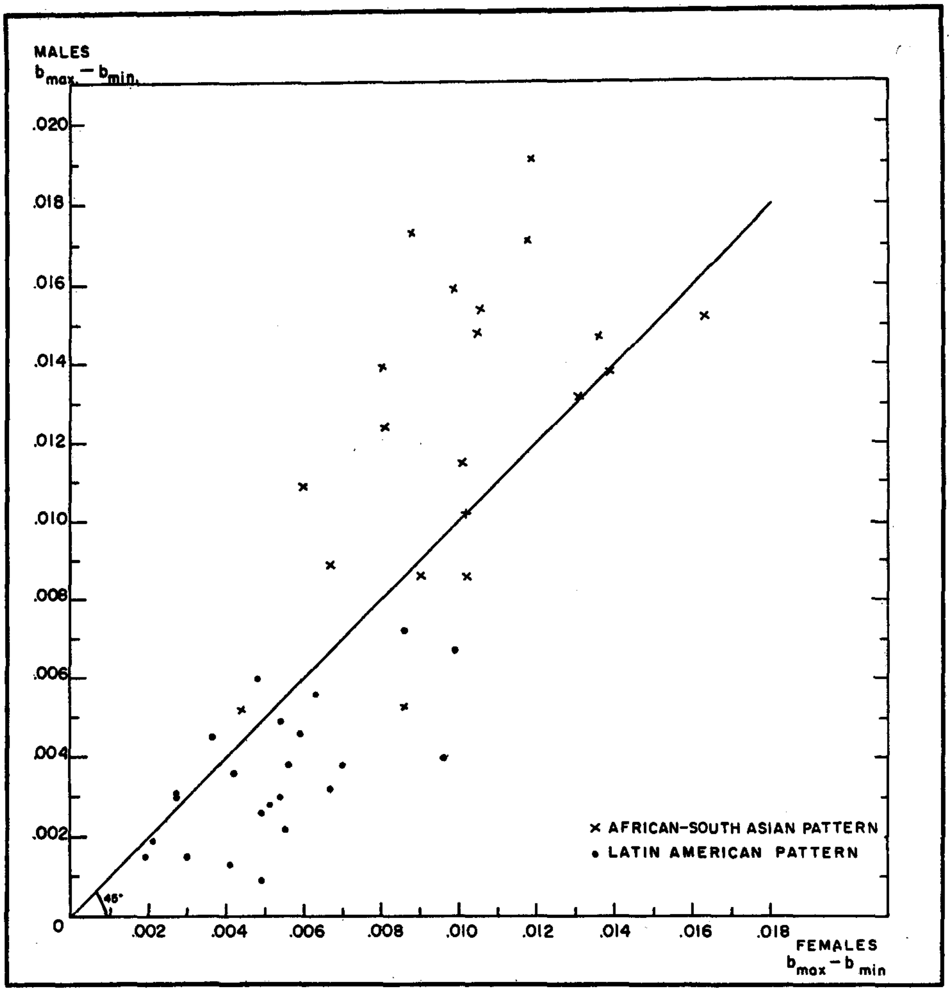

Figure XI. Ranges of stable population estimates of the female birth rate (b) versus those of the male birth rate in various censuses. Ranges are the difference between the highest and the lowest birth rate in stable populations having an $^0e_0$ of forty years and agreeing with C(5), C(10), ..., C(40) in the census populations. Estimates shown were obtained from two groups of censuses, each characterized by a typical pattern of age-misreporting

rates were estimated (assuming $^0e_0 = 40$) on the basis of ogives to age 5, 10, 15, ..., 40 for males and females, employing "West" model stable populations. These calculations were limited to age distributions not obviously affected by rapid changes in fertility or mortality. The range (highest estimate minus lowest estimate) for males and females in the African - southern Asian and Latin American censuses so selected were then plotted. Note that the range is much less for both sexes in the Latin American group; that in a large proportion of Latin American censuses the male age distribution leads to a smaller range than the female; but that in the African-southern Asian populations, the opposite is seen—a larger range of estimates when based on male age distributions.

The smaller gross displacements in female than in male age distributions in the African - southern Asian censuses is consistent with the hypothesis that in these surveys age is often estimated by someone other than the person in question. The estimation of female ages is assisted by clues—bodily changes associated with menarche, number and age of children, and a fairly well-defined upper limit to the ages of childbearing—that do not exist, or are less evident, for males. The smaller distortions affecting male age distributions in Latin America are also consistent with the surmise that age in these populations is usually supplied by the respondent. The more extensive education and greater worldliness of the male would lead to a better average knowledge of age.

4. *Selecting a model stable population on the basis of a distorted age distribution*

The observations on age-misreporting in the preceding section imply that the pattern of distortion must be taken into account when determining what model stable population best fits the population in question. As a preliminary step, the female age distribution may be compared with a model stable population with the same proportion under age thirty five (and $^0e_0 = 40$) to see if it has the characteristics (African - southern Asian) listed under sub-section (*a*) above. If the age distribution *does* have these characteristics it is likely that the female age distribution is a better source of estimated population characteristics than is the male. The choice of a female stable population is made and the consequent estimate of the birth rate is obtained as follows:

(1) The female model stable populations with a rate of increase equal to that of the female population and with the same C(5), C(10), ..., C(40) are selected and the birth rate of each is recorded.

(2) The sequence of estimated birth rates will typically have a peak at age ten, fall at ages fifteen and twenty, and rise to a second peak at thirty-five or forty, if the age distribution is of the African - southern Asian type. On the basis of the analysis of this pattern given earlier, the birth rate estimate based on C(10) may be about right, or may be too large, because there is sometimes, but not always, understatement of the ages of enough 10-14 year-olds to inflate the proportion under ten. The birth rate estimated from C(15), C(20), C(25) and C(30) would be too low because of the almost universal overstatement of age among young women in these populations. The estimate based on C(35) is expected to be at about the right level, except for the possible effects of recent mortality or fertility trends.

(3) A minimum estimate of the birth rate in the female population is obtained from *r* and C(15).

(4) If evidence in favour of stability is convincing (i.e., little change in intercensal growth rate or in age distribution for two or three decades), the estimate based on C(35) can be accepted, and with more confidence if confirmed by the figure derived from C(10). If growth has been accelerating, the estimate based on C(35) can be adjusted by methods described in section C of this chapter.

(5) The birth rate in the male population is estimated as

$$\text{sex ratio at birth} \times \text{female birth rate} \times \frac{\text{female population}}{\text{male population}}$$

In the absence of reliable direct information on the sex ratio at birth a value of 1.05 should be taken in populations other than those of tropical Africa. In African populations a multiplier of 1.03 should be employed.

Now suppose that the comparison of the female age distribution with the model stable population with the same C(35) shows that the recorded population has a substantial deficit under age five relative to the stable, and a tendency for the proportions under age ten, fifteen, twenty and twenty-five also to fall somewhat short of this stable, but to a minor and decreasing extent, and the proportion in successive five-year age groups does not depart by more than perhaps 5 to 10 per cent from the stable from age ten to age forty. It may then be assumed that the age distribution is of the Latin American pattern and the choice of a stable population is made as follows:

(1) Find the male model stable populations with a rate of increase equal to that of the male population, and with the same proportions under age 5, 10, 15, ..., 40 as in the recorded male population, and note the birth rate in each. Perform the analogous calculations with the female population;

(2) The sequence of male birth rates will typically rise from age five to ten, and then fluctuate mildly until age forty or forty-five. The female birth rate will follow a similar sequence except that there will be a tendency toward mildly increasing estimated birth rates from age ten to forty or forty five.[16]

(3) Select the model stable population for males that produces the *median* value of the birth rate among those agreeing with C(5), C(10), ..., C(45). The birth rate of the female stable population can then be taken as

$$\frac{\text{male birth rate}}{\text{sex ratio at birth}} \times \frac{\text{male population}}{\text{female population}}$$

This estimate can be checked by comparing it with the average of the birth rates in the female model stable populations that agree with the recorded female populations in growth rate and C(20), and in growth rate and C(25).

5. *Assigning the characteristics of a model stable population*

Once a model stable population has been selected as a close approximation of the actual population, the characteristics of the stable population can be ascribed to the population in question. At a later point, in chapter IV, we shall comment on the accuracy of correspondence between the model stable population and the actual with regard to mortality and fertility. At this place we wish only to indicate what characteristics of the population can usually be attributed to the actual.

When circumstances warrant the assumption of approximate stability of the population (i.e., when there are indications of a past history of approximately constant fertility and mortality), the age distribution of the stable population can be attributed to the actual population. This attribution is useful in those populations with large systematic errors in age reporting. For example, an appropriately chosen stable population undoubtedly has an age distribution closer to reality than the reported age distribution for the population of India in 1911. In fact the stable population age distribution can be used to provide a base population for population projections that will provide a more valid basis for estimating the future evolution of the population of schoolgoing age or of the ages of labour force participation or the like than would be obtained by basing a projection on the often highly distorted age distribution recorded in a census or survey. The principal purpose of the stable population analysis described in this *Manual* is, however, to provide estimates of fertility and mortality.

As will be seen later, the estimates of fertility derived from stable analysis are ordinarily more trustworthy than estimates of mortality. Strictly speaking, information on age composition and growth which permit the choice of a matching model stable population provide an estimate of the crude birth rate as a measure of fertility and do not enable us to estimate total fertility (the number of children

[16] Frequent exceptions are found in the Latin American censuses taken since 1950 because of the effect on the age distributions of declining mortality, combined in some instances with a slight rise in fertility. The effect is to produce estimates of the birth rate among males that fall from age ten to forty or forty-five, and among females that are about constant, or declining less sharply than the male estimates. The normal sequence of approximate constancy for the males and slightly rising estimates for the females is restored by the adjustments described in section C of this chapter.

born per woman passing through the fertile part of life) or the gross reproduction rate (the number of daughters born per woman in the same span). The reason is that two populations with the same fertility in the sense of the same birth rate and with the same age composition can have quite different numbers of children per woman passing through the childbearing span, depending upon whether, on the average, births occur relatively early or relatively late in the span. When fertility is as high as it is in almost all of the populations for which stable analysis is appropriate, a low mean age of child bearing permits women with a smaller total fertility to produce the same birth rate as achieved by a population in which women of higher total fertility produce their children somewhat later in life. This relationship – i.e., that low fertility and early child-bearing produces the same birth rate as higher fertility and late child-bearing – is represented in our tabulation of model stable populations by the presentation of four gross reproduction rates with each population; each gross reproduction rate being associated with a particular mean age of the fertility schedule.[17] This multiple tabulation of gross reproduction rates with each stable population of course implies that the identification of a model stable population as essentially identical with a given actual population is not a sufficient basis for estimation of total fertility or the gross reproduction rate. In addition, the analyst must have some knowledge of the mean age of the fertility schedule characterizing the women in the population in question.

The mean age of the fertility schedule can be calculated from tabulated responses to a question about births occurring in the preceding year (with due allowance for the fact that women who report a birth during the preceding year would on the average have been six months younger at the time of birth than at the time of the survey). In many populations no such direct evidence on the age pattern of fertility is available. In such populations the mean age of the schedule can sometimes be estimated by various indirect approaches. One possibility is to estimate the age pattern of fertility by assuming that marital fertility follows a standard pattern. Table 1 shows a pattern of marital fertility—fertility rates expressed in terms of the rate for age 20 to 24—that can serve this purpose.

TABLE 1. STANDARD AGE PATTERN OF FEMALE MARITAL FERTILITY RATES

Age	*Index of marital fertility rate*
15-19	$1.2 - .7m(15\text{-}19)^a$
20-24	1.000
25-29	.935
30-34	.853
35-39	.685
40-44	.349
45-49	.051

[a] $m(15\text{-}19)$ = proportion of married females at ages 15-19.

The rates shown in table 1 for ages 20-49 are based on average experience of a number of populations in which little or no birth control is practised.[18] The value for age 15-19 is a rough approximation according to which if, in the absence of birth control, marital fertility for age 20-24, i.e., f (20-24), and the proportion of married females at age 15-19, i.e., m(15-19), are known, f(15-19) may be estimated as $1.2 f(20\text{-}24) - .7\ f(20\text{-}24)\ m(15\text{-}19)$. In a population in which it is known that the fertility of non-married women is a negligible factor in the over-all birth rate, and where the census includes a tabulation of marital status by age, the age pattern of fertility can be approximated by multiplying this standard marital fertility schedule by the proportion married among the women in each age group and from this approximated schedule the mean age of the fertility schedule calculated.

The method described in the previous paragraph is not applicable in populations with high proportions of births outside of marriage, or in which there is a variety of sanctioned sexual unions, including for example widespread consensual unions.

When women have been asked about the number of children ever born, and the responses have been tabulated by age of woman, it is possible to estimate the mean age of the fertility schedule from the relation of average parity at age 20-24 to average parity at ages 25-29. If the earlier average parity is called P_2 and the later P_3, the ratio P_3/P_2 (in a population not practising birth control) depends primarily on the ages at which women begin their child-bearing, a high value of P_3/P_2 indicates a late start, and a low value an early start. It may be assumed that the decline of fertility with age in populations not practising birth control follows a fairly common pattern, so that the mean age of the fertility schedule is determined primarily by the rising portions. The relationship between the mean age of the fertility schedule and the value of P_3/P_2 has been calculated on the basis of the recorded schedules of a number of populations in which there is little practice of birth control. The mean age of each schedule was calculated, and the values of P_2 and P_3 implied by the schedules computed. The relation of $\bar{m}$ to P_3/P_2 is very close, as is shown in figure XII, and is well expressed by the following linear expression:

$$\bar{m} = 2.25 \frac{P_3}{P_2} + 23.95$$

When, due to the lack of required data, neither of the approximations described above is applicable, the best alternative is to assign to the population in question the mean age of the fertility schedule of another population presumed to have similar factors affecting the age of the fertility schedule. Thus in a Latin American country a possible expedient is to utilize the mean age of child-bearing from a nearby neighbour, if the two countries in question are both known to have a roughly similar prevalence of consensual unions and of marriages performed by the State or the church. The reader's attention is directed to the expedient used in examples worked out in part two of this *Manual*.

[17] The arithmetical mean of the schedule itself, unweighted by the age distribution.

[18] They are adopted from Louis Henry, "Some Data on Natural Fertility", *Eugenics Quarterly*, vol. 8, No. 2 (June 1961), pages 81-91.

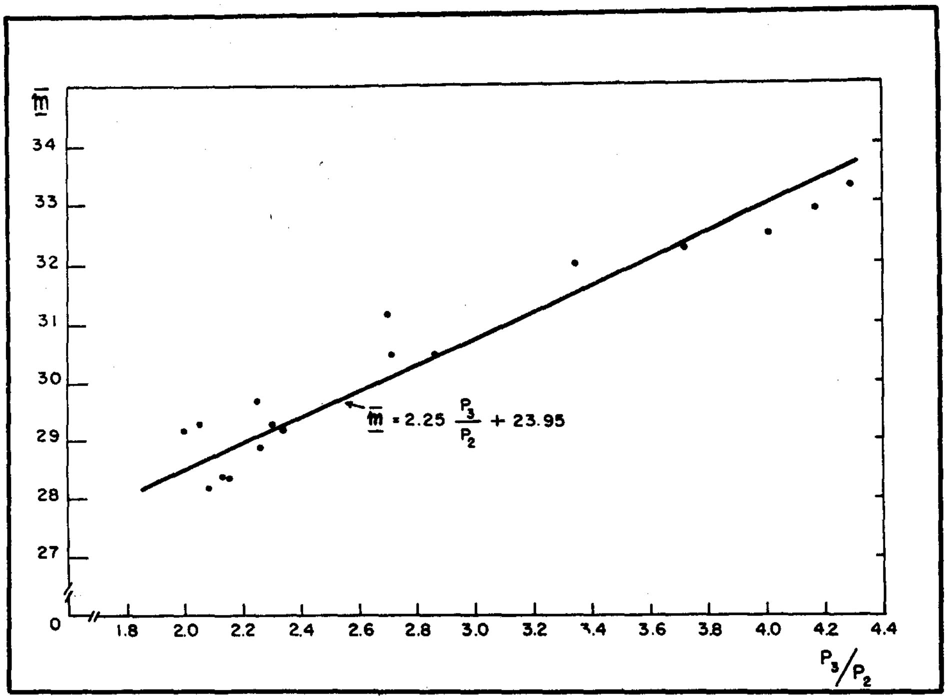

Figure XII. Mean age of the fertility schedule in populations not practising birth control *versus* ratio of P_3 (average parity at age 25-29) to P_2 (average parity at age 20-24)

C. Adjustment of estimates based on model stable populations for the effects of recent decreases in mortality

It goes without saying that many populations have age distributions that do not conform at all closely to any of the model stable populations. Major fluctuations in fertility create unusually large or unusually small birth cohorts that stand above or below the corresponding age group in any stable population during their lifetime. Persistent trends in fertility can create an age distribution far different from any stable population. Because the populations of all highly industrialized countries of the world have experienced sustained decreases in fertility amounting in almost every instance to a 50 per cent reduction of total fertility, and since most of these populations have experienced substantial fluctuations in fertility in addition to the downward trend, age distributions of industrialized populations cannot in general be matched by those of model stable populations. Events in the history of some of the less developed countries have also caused irregularities in the age distribution that would find no match in one of the model stable populations. Some populations have age compositions strongly influenced by age-and-sex-selective migration. Such migration is typical of the urban populations of developing countries and a non-stable form of age distribution is therefore to be found at certain census dates in such populations as that of Singapore and Hong Kong. The prolonged mobilization of a large fraction of the young adult male population into military service can have a pronounced effect on fertility and create a lasting notch in the age distribution. Military casualties of course affect the age distribution of the male population. Invasions and revolutions can leave similar traces. Finally, major epidemics such as the world-wide infleunza epidemic of 1918 to 1920 cause a temporary reduction in fertility and an excess of infant and child mortality that again produce a small cohort evident in subsequent age distributions.

On the other hand, in developing countries where the population is little affected by international migration, and in the absence of major catastrophes such as wars or great epidemics, fertility tends to remain at a fairly level plateau. The only apparent exception is in areas where relatively late marriage is the custom, such as in western Europe before the systematic decline in marital fertility began. In these populations, for example, in Tuscany during the nineteenth century, changes in total fertility amounting to 20 or 25 per cent can occur caused by long intervals in which the average age of marriage was increased and other intervals in which it was reduced. It appears possible that in Latin America changes in age at marriage, and perhaps differential recourse at different times to marriage on the one hand and to less stable concensual unions on the other have raised and lowered the average level of fertility.

The conclusion that emerges from these observations is that there are populations in less developed countries that cannot be analysed by stable population techniques, and others in which the precision of the estimates may be degraded, even though the general approach remains useful. Among the contemporary populations for which the method does not appear useful are those strongly affected by migration and those that have suffered a sequence of wars or other major disturbances. There remains the majority of populations in the less developed countries, where the assumption of a history of more or less constant fertility is warranted, but where mortality has followed a strong and sustained downward trend in recent years.

The prevalence of rapidly falling death rates in the less developed countries is well known and need not be described in detail in this *Manual.* Falling mortality has followed as many different courses (if these are considered in detail) as there are identifiable populations in the developing countries. It is of course not possible to describe how every imaginable decline in mortality would affect the age composition of a population. Many demographers have noticed that different mortality schedules produce only slightly different stable populations, and that populations experiencing approximately constant fertility and changing mortality show only restricted alterations in age composition. Finally, it has been noted that the age composition produced at each moment of time during a prolonged period of declining mortality bears a closer relationship to the stable implied by the current fertility and mortality conditions than to the age distribution of an earlier period when mortality was higher. Such observations have been used to justify the application of the methods of stable population analysis to populations experiencing approximately constant fertility and steadily declining mortality. Indeed a special designation of quasi-stable has been invented for these populations. Coale and Demeny have analysed the ways in which populations that have experienced declining rather than constant mortality differ from the stable population that would have resulted had current mortality and fertility conditions prevailed throughout the past.[19] They noted that in spite of the close visual resemblance between the age composition of a stable population and of a so-called quasi-stable population, an estimate of total fertility or the birth rate based on the ogive of the quasi-stable population and the current rate of increase can be in error by as much as 10 to 15 per cent. In this section a method is described by which the demographic analyst can adjust estimates of population parameters extracted from model stable populations to compensate for the distorting effects of a history of recently declining mortality.

On the basis of previous analytical work it is known that the principal effect of declining mortality on age composition closely resembles the influence of steadily rising fertility. Over a specified interval of falling mortality, it is possible to find a proportionate change in fertility—say a 7-per cent increase—that is equivalent to the recorded decline in mortality so far as the effect on age composition is concerned. It is possible, by an extension of this idea, to determine what sequence of annually rising expectation of life in the "West" model life tables given in annex I would be equivalent to an annual increase in fertility of one per cent. A series of population projections were constructed along these lines in which the initial population was a model stable population with a total fertility of about 5, 6 and 7 respectively, and with various levels of the initial expectation of life at birth. These populations were then projected for forty years with steadily rising expectation of life at birth (i.e., with steadily falling mortality) at a pace which in each year was equivalent to a one per cent increase in fertility in the principal influence on changing age distribution. These projections were performed on an electronic calculator. The programmed computations included the calculation of the average rate of increase during each five-year period and also during each ten-year "intercensal" interval, birth rates, death rates, age distributions in five year groups at the end of each five-year time interval, and ogives of these distributions. Birth rates and total fertility rates were then "estimated" at five-year intervals by choosing a model stable population with the same rate of increase as the "intercensal" rate of increase during either the preceding five years or ten years in the population projection, and the values of C(5), C(10) etc. in the projected population. The difference between the estimated rate calculated in this fashion and the average value of the "true" birth rate or of the "true" total fertility during the intercensal period (calculated as part of the projection programme) was then obtained. These differences, expressed as a proportion of the true value of the birth rate or of total fertility, turned out to be very nearly the same whether the population projection was for a fertility of 5, 6 or 7 births per woman. Thus these proportionate differences can be used as the basis for adjustment factors to reconstruct the true birth rate from the birth rate estimated from the model stable population—if the duration of the decline in mortality is known and if the falling death rates are at a pace equivalent to a one per cent annual increase in fertility so far as age composition effects are concerned. That is so since the calculations also showed that the adjustments needed, e.g., at fifteen years after mortality decline began, were for practical purposes, independent of the initial level of mortality, i.e., they were substantially the same whether the initial expectation of life at birth before the onset of declining mortality was, for example, twenty years or thirty years. Finally, sets of population projections were programmed that differed with respect to the speed of mortality decline, e.g., projections in which the decline in mortality was only half as fast as specified above; in short, in which mortality, changes were equivalent in their principal age compositional effects to an annual increase in fertility of one-half of one per cent. These projections showed that the differences in the estimated and actual values of birth rates and total fertility were proportionate to the speed of mortality decline, i.e., in the example just given they were almost exactly half as great as with the more rapid decline in mortality.

This *Manual* includes as a result of these calculations a table of adjustments (see annex III, table III.1) to be applied to birth rates and to gross reproduction rates (or total fertility rates) derived from the model stable populations— adjustments that are appropriate for ogives up to age 5, 10, ..., 40 and which assume different values at $t = 5, 10, \ldots, 40$, where t is the time after the decline in mortality begins. These adjustments should be applied when the value of k equals .01, where k is a parameter indicating the rate of mortality change expressed in terms of the equivalent proportionate annual increase in

[19] Ansley J. Coale, "Estimates of Various Demographic Measures Through the Quasi-Stable Age Distribution", in *Emerging Techniques in Population Research*, Milbank Memorial Fund,1963, pp. 175-193, and Paul Demeny, "Estimating Vital Rates for Populations in the Process of Destabilization", *Demography* (Chicago), vol. 2, 1965, pp. 516-530.

fertility in so far as age distribution effects are concerned. For values of k other than .01 – i.e., in the general case when the age distribution effects of the changing mortality are equivalent to an average annual fertility change that is greater or smaller than one per cent per year—the tabulated adjustments are to be scaled up or down in the same proportion as the actual value of k differs from .01. For instance if the value of k is .012 the appropriate adjustment factors are to be increased by 20 per cent. Annex table III.1 thus enables the analyst who has made a preliminary set of stable population estimates of the birth rate and of the gross reproduction rate (from the age distribution cumulated to age 5, 10 etc. in conjunction with the intercensal rate of increase) to adjust these preliminary estimates in order to correct the bias present in the stable estimates due to the fact that contrary to the assumptions underlying the stable estimates mortality in fact has been declining. A quasi-stable estimate of the death rates is obtained by subtracting the intercensal growth rate from the adjusted birth rate. A quasi-stable estimate of the expectation of life at birth is finally derived by finding the ${}^0e_0$ of the *stable* population characterized by this death rate plus any of the other parameters (GRR, b or r) for which the quasi-stable estimates have previously been calculated.

The application of the method described in the preceding paragraph assumes that estimates of the duration and average pace of mortality decline—i.e., estimates of the parameters t and k—have already been obtained. Preferably such estimates should be based on information concerning the rate of population growth in the decades preceding the census that is being analysed— information that is sufficient to locate approximately the time when the departure from the stable state has occurred and which indicates the tempo at which destabilization has taken place. Specifically, on the basis of the same calculations that were outlined above, it was found that k can be estimated as $17.8 \times \Delta r/\Delta t$ where Δr is the absolute change in the rate of growth as compared to the original stable rate, and Δt is the number of years that have elapsed while that change took place.[20]

The acceleration of population growth is not the only evidence from which the parameters k and t can be estimated. What is needed is *any* approximate indication of the duration and pace of recent mortality declines. In chapter II a method of estimating child mortality from special census data is described, and in chapter III the combination of estimated child mortality and the age distribution to select a model stable population is outlined. At the end of chapter III it is noted that *changes* in the estimated level of child mortality from census to census can be used to determine approximate values of k and t[21].

Another basis for estimating k and t is the changing age composition of deaths. The proportion of deaths by age changes when mortality declines and fertility remains constant. It is possible to obtain a very rough estimate of k and t from even an inexact record of the changing age composition of deaths. The reader should note that the method outlined in the next paragraphs is clearly imprecise, but that it is used only to indicate the magnitude of an adjustment for the effects of declining mortality, and that an approximate adjustment usually produces an estimate superior to the unadjusted stable value.

The estimation procedure rests on the following supporting facts and relationships:

(1) When initially stable populations are projected with declining mortality but constant fertility, certain changes in the age composition of deaths are very closely correlated with the associated change in expectation of life at birth. The most usable index of the changing composition of deaths is perhaps the ratio of deaths to persons over sixty-five to deaths to persons over age five. This index is as closely related as any other to $\Delta {}^0e_0$, and has the merit of omitting infant and child mortality, which could have an overiding and possibly misleading effect, especially when completeness of registration changes.

(2) The change in the proportion of deaths over sixty-five may be faithfully represented in registered deaths, even when these are so incomplete in coverage that neither the level nor the trend of the crude death rate can be derived directly.

(3) An approximate value of $\Delta {}^0e_0$ can be obtained from annex table III.2 as the difference between two *stable population* estimates of ${}^0e_0$ referring to different periods of time, each based on a value of the index deaths 65+/deaths 5+ and on a rough measure of fertility

$$\left(\frac{\text{births of a given sex}}{\text{persons 15-44 of the same sex}}\right)^{22}$$

(4) The value of $k \,.\, t$ associated with a given $\Delta {}^0e_0$ depends rather strongly on the base level of ${}^0e_0$, and an approximate value of the latter is needed before $k \,.\, t$ can be estimated from table III.3. A crude but usable indication of the terminal level (which less $\Delta {}^0e_0$ gives the base level) of ${}^0e_0$ can again be obtained by the provisional assumption of stability. It is recommended that the provisional terminal level be taken as the average of the ${}^0e_0$'s in the stable populations associated with r and C(10), and r and C(15). The same procedure is to be used in obtaining the rough measure of fertility that is needed in estimating $\Delta {}^0e_0$, as explained in point (3) above.

[20] This formula assumes that the acceleration of growth is attributable in its entirety to a change in mortality, i.e. that fertility has remained constant. It may be noted at this juncture that in a formal sense the procedure adjusting stable estimates for the effects of changing mortality may be extended to the case where destabilization has been brought about by changing fertility, or a mixture of changing fertility and mortality, rather than by changing mortality alone. Assume, for example, that fertility has been changing for t years (following an original stable situation) while mortality has remained constant. The adjustment factors tabulated in table III.1 would still be applicable—with proper attention to sign, i.e., multiplied by -1 in case of declining fertility—the value of k to be used being simply the average annual change in fertility. Note that the value of the multiplier connecting k and $\Delta r/\Delta t$ would then be about twice as large as in the case of mortality change; approximately 36 instead of 17.8. In view of the fact that sustained changes in fertility are less regular and, under the conditions necessitating the application of quasi-stable techniques for estimating vital rates, are less common, no systematic discussion of this topic is offered in this *Manual*.

[21] See chapter II, section B, chapter III, section C.2 and annex table III.4.

[22] This ratio is one of the stable population parameters tabulated in annex II.

It must be noted that the recommended procedure for estimating kt from the changing age composition of deaths (i.e., approximating Δ^0e_0 and the level of $^0e_0$ from the provisional assumption of stability before using table III.3) generates provisional data on mortality—namely the level of $^0e_0$—that are useful only in determining a correction factor, and that must not be confused with a final estimate of that parameter.

D. Concluding remarks on estimates adjusted for the effects of recent declines in mortality

Whether adequate direct or indirect information on past changes in mortality are available or not it is to be expected that with reasonably good reporting of ages such a deviation from stability will be discernable from the age distribution itself. The nature of the effect of a history of declining mortality on the stable age distribution is indicated in figure XIII which shows $C(x)-C_s(x)$ where

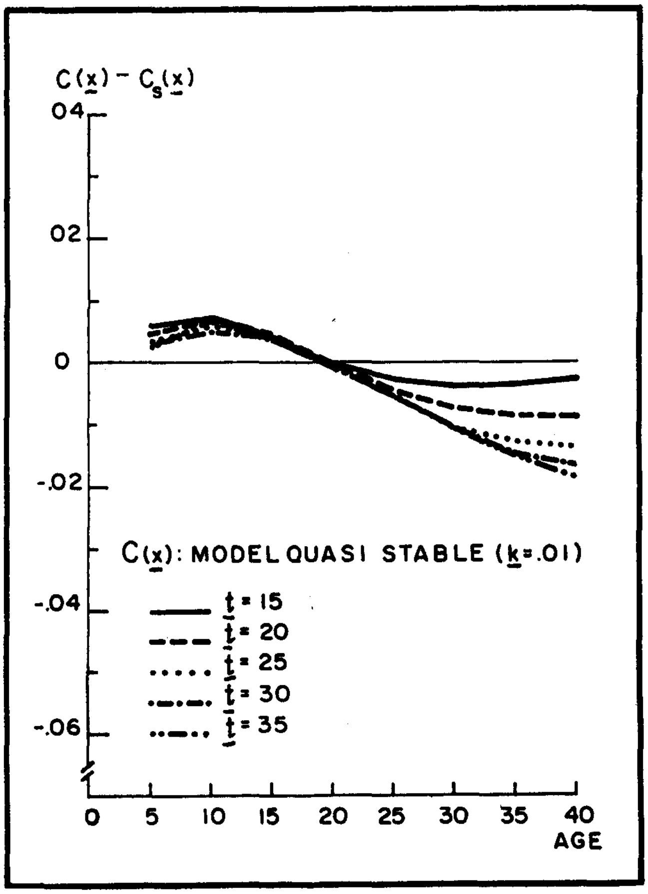

Figure XIII. Comparisons of the ogive—$C(x)$—of model quasi-stable populations (populations which were originally stable but which have experienced a decline of mortality for t years), with that of stable populations—$C_S(x)$—having the same proportion under age twenty and an $^0e_0$ of forty years

$C(x)$ is the ogive of the projected population with mortality declining for 15, 20, ..., 35 years and $C_s(x)$ is that of the stable populations having the same $C(20)$ and $^0e_0$ of 40 years. In figure XIV the plot of $C(x)-C_s(x)$ of several Latin American populations with censuses in the early 1960s is shown in comparison with the projected population where mortality has been declining for twenty years. The imprint of the effect of declining mortality on the age composition of these populations is clearly evident.

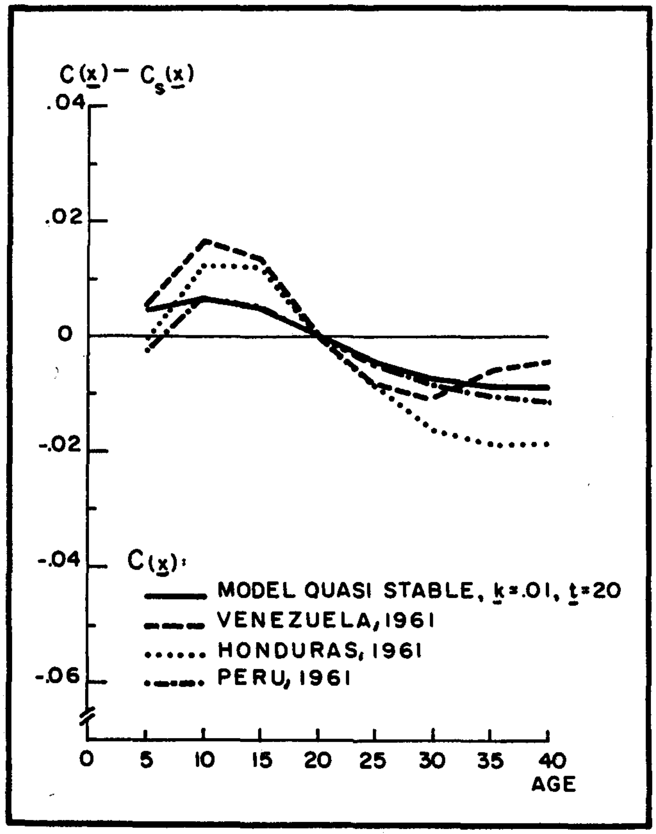

Figure XIV. Comparisons of the ogive—$C(x)$—of a model quasi-stable population and of male populations as recorded in recent Latin American censuses, with that of stable populations —$C_S(x)$—having the same proportion under age twenty and an $^0e_0$ of forty years

The age distributional effects of declining mortality are also present in recent censuses of such populations as that of India and Pakistan, but because of the nature of the distortions in the reported age distribution the effects are not so easy to discover as in the Latin American censuses. However it is interesting to note that when stable population analysis is applied to the Indian age distribution (females) in 1911 there emerges from the intercensal rate of increase and the proportion under age ten and under age thirty-five estimates of total fertility in approximate agreement—estimates of 6.5 children per woman in the former case and 6.3 in the latter case. When stable population analysis is applied to the Indian census of 1961 the total fertility estimated from the proportion under age ten and the intercensal rate of increase is 6.6 and from the intercensal rate of increase and the proportion under thirty-five is only 5.8. However, when the appropriate adjustments are made to allow for the influence of declining mortality in the preceding forty years at a pace estimated from the average rate of acceleration of growth the resultant estimates are 6.6 and 6.4, respectively.

E. The estimation of fertility from the age distribution recorded in one census

What is the minimum information that permits demographers to form approximate estimates of fertility or mortality? In the earlier sections of this chapter methods are described for basing estimates on a series of two or more population censuses, and in the next chapter there is an outline of techniques of estimation to extract the maximum of reliable inference from special questions consciously inserted in a census or survey to measure fertility and mortality. What can be learned from a *single* census (so that no intercensal rate of increase can be calculated) that provides no special fertility or mortality tabulations?

Age composition is more strongly affected by fertility than by mortality, so that with minimal information more reliable estimates can be made of the birth rate than of the death rate. For example, if a closed population contains a very high proportion of young children, it must be a population that has recently experienced high fertility—an inference that is valid whether the expectation of life at birth is high or low. However, because high infant and child mortality diminish the fraction under age five or ten, a population with a large proportion of children would be a moderately high fertility population if mortality were low, and a very high fertility population if mortality were high.

Consider a specific example: suppose the proportion under ten years of age in a population was 30 per cent. If mortality were assumed to be that of the model life table with ${}^{o}e_0 = 30$ years, births in the preceding ten years could be estimated as 30 per cent of the current population times a factor derived from the model life table expressing the reciprocal of the proportion surviving, and the guess may be hazarded that with such a low expectation of life, the population grew moderately—at perhaps 1.5 per cent per year so that it was about 7.5 per cent larger than it was at the midpoint of the preceding ten years. On these assumptions the average birth rate during the decade preceding the census would be 50 per thousand. On the other hand, if mortality were assumed to match that of the model life table with ${}^{o}e_0 = 40$ years, and if a more rapid rate of population growth—say, 2 per cent annually—were viewed as a sensible guess, the resultant estimate of the birth rate would be 44.3. Often the conditions in which the population lives are known well enough, and the mortality of other populations in similar circumstances are well enough recorded to give some confidence that a broad estimated range of the level of mortality probably encompasses the actual figure. The use of model stable populations can be used to translate a recorded proportion under age five, ten or fifteen and an assumed level of mortality into an estimated birth rate without elaborate calculations. If the basis of estimation is confined to ages under fifteen, the procedure is closely equivalent to reverse projection and the estimate is valid even if the enumerated population in fact departs from the stable age distribution.

In the hypothetical example just described, the implication that the demographer would know that the proportion under age 10 is .3 is unrealistic, since if a population has had its age distribution recorded only once, it is likely that the recorded distribution is seriously distorted by age-misreporting. A more realistic example is provided by considering the problem of estimating vital rates for Indonesia on the basis of the age distribution information provided by the 1961 census of that country. As shown in figure XV combination of various indices of this age distribution (proportions up to age 5, 10, ...etc.) with specified hypothetical growth rates or levels of mortality imply widely varying birth rates. Unless some particular indices of the age distribution can be accepted as more reliably reported than others, thus leading to a narrower range of uncertainty, the information provided in figure XV would be of little value. Inspection of this figure as well as comparisons of the Indonesian age distribution in cumulative form with the ogives of model stable populations reveals the pattern of age-mis-statement characteristic of censuses in India, Pakistan and many African populations. It is therefore appropriate to use the rules of estimation devised for such age distributions. Table 2 highlights the figures that are relevant in applying these rules.

Table 2. Stable estimates of the female birth rate of Indonesia derived from the 1961 female age distribution

Assumption about mortality or rate of increase	*Birth rates based on assumed levels of mortality or rate of increase, and proportion under age:* 10	15	35
${}^{o}e_0 = 30$	.0571	.0492	.0502
$= 35$	.0534	.0460	.0472
$= 40$	.0505	.0434	.0450
$r = .010$	.0817[a]	.0559	.0565
$= .015$	.0675	.0491	.0503
$= .020$	.0594	.0435	.0462

[a] Extrapolated figure.

The estimates based on C(15) are a series of minimum estimates, those based on C(10) appear in this instance to be a series of maximum estimates, and those based on C(35), to be the best obtainable. While the interpretation of the Indonesian age distribution in the light of outside experience does drastically reduce the width of the interval within which the birth rate is likely to be located, lacking further information it would be unwarranted to go beyond the cautious assertion that the birth rate in Indonesia is probably not less than 45 per thousand and not more than 56 per thousand.

The Indonesian population was enumerated in 1930 in a census that did not record the distribution by chronological age. Strictly speaking, then, Indonesia is not a proper example of a single census. For example, over-

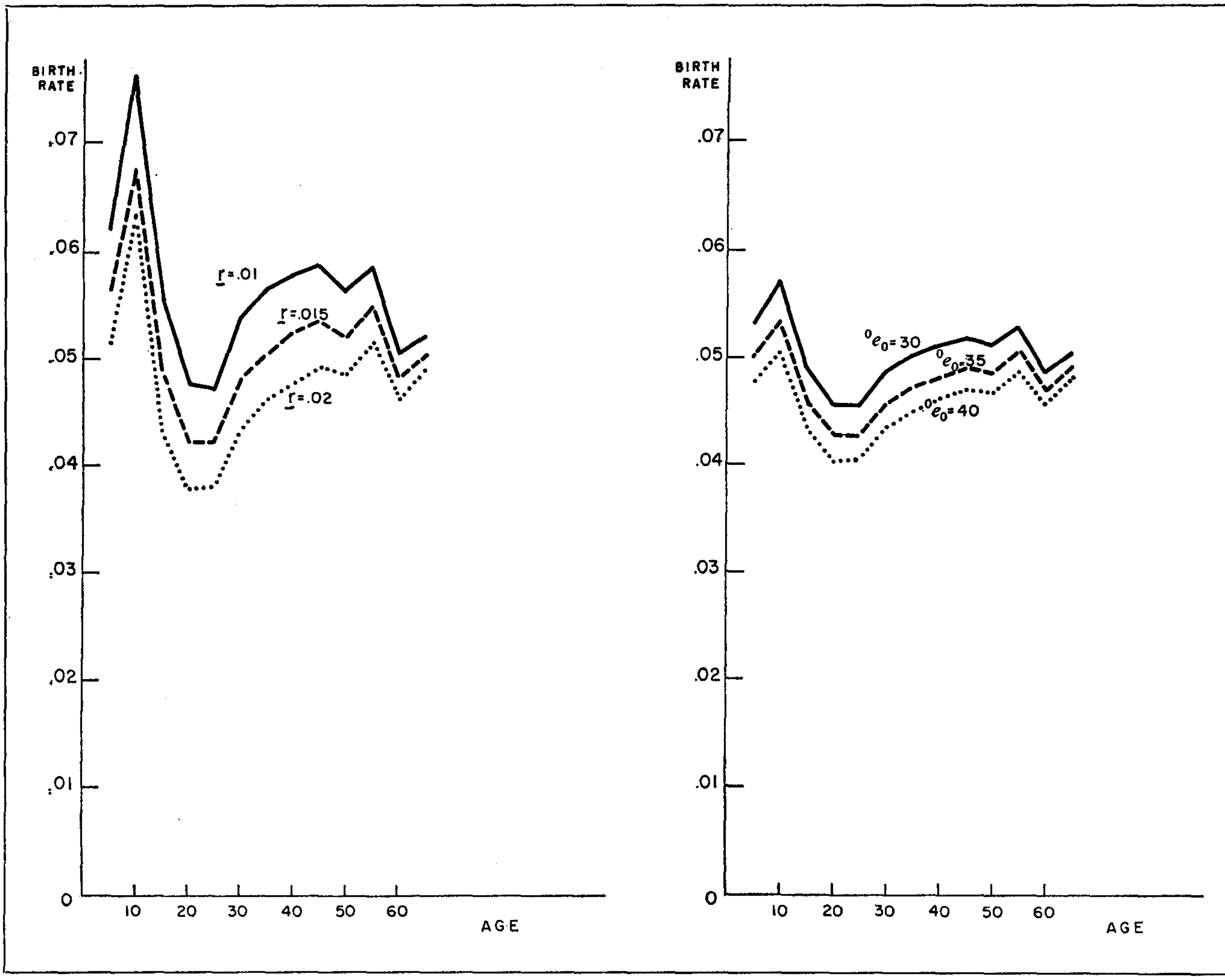

Figure XV. Stable population estimates of the female birth rate in Indonesia derived from $C(x)$ as reported by the census of 1961 and from hypothetical rates of population growth (left panel) or hypothetical levels of mortality (right panel)

looking differences in geographical coverage and completeness of enumeration, one can calculate the intercensal rate of increase of the female population (for thirty-one years) as 1.63 per cent per annum. This information would lead us to guess that the growth rate just before 1961 was 2 per cent (or more), and that the estimate of 46.2 per thousand for the birth rate was therefore to be preferred to the 56.5 and 50.3 associated with lower rates of increase. But if a rate of increase above the intercensal average is accepted for the period just before the census, population growth must have been accelerating, primarily, one assumes, because of falling mortality. If mortality had been falling for fifteen years at a rate causing an acceleration in the growth rate of about 5 per thousand each decade, the estimated female birth rate based on C(35) and r = .02 should be increased by 6 to 10 per cent—to 50 per thousand, or a little higher. The male birth rate is probably about 8 per cent higher than the female, because the number of males is about 2 per cent less, and the number of male births is normally about 5 per cent greater. As a result of the above arguments then, it is possible to conclude that the birth rate in Indonesia was probably at least 2 or 3 points above 50 per thousand in the years before the census. A corresponding minimum estimate for total fertility, assuming that the mean age of the fertility schedule is twenty-nine years, is 6.5.