614

THE ESTIMATION OF ADULT MORTALITY FROM DEFECTIVE REGISTRATION DATA

A Thesis presented for the Degree of
Doctor of Philosophy in the Faculty of
Medicine

University of London

by

Hoda Mohamed Rachad

London School of Hygiene and Tropical Medicine

1978

ABSTRACT

The availability and quality of demographic data in developing countries are far from adequate. The introduction and improvements of techniques for estimating mortality from nontraditional sources of data and for correcting the shortcomings in traditional data are indispensible.

Data on deaths in a period but with an unknown completeness of coverage is usually available through vital registration or from single or multi round household surveys. The growth balance method makes use of such data and provides an estimate of the extent of the under-registration of deaths. An extensive study of this method, regarding the effect of deviations from the underlying assumptions and possible modifications to overcome its shortcomings, is presented.

This study reveals that the method is generally robust to patterns of mortality change similar to those in developing countries and also to recent changes in fertility. Possible modifications to allow for certain types of changes in mortality and fertility are also presented.

A modification of the method to allow for the effect of migration is introduced and applied to actual data of Kuwait.

The effect of differential under-registration of deaths on the method is discussed and a procedure to estimate this differential under-registration is proposed. This procedure is applied to hypothetical data as well as to data on Iraq.

A model of age error and the general likely effect of this error on the growth balance estimate are discussed. Several practical considerations are also dealt with, such as the effect of graduating the age and death distribution

before applying the method, the appropriate method of fit and an alternative formula that may be used.

Finally, as an illustration of the interaction of several deviations from the underlying assumptions and the suitability of the technique and the adjustments procedures suggested, a general application using hypothetical data and actual data for Guinea is presented.

CONTENTS

Page

ACKNOWLEDGEMENTS

I would like to express my gratitude and debt to Professor W. Brass for suggesting the original idea, for his continuous help and guidance during the preparation of this thesis and for making his time available whenever asked.

Thanks to my husband, Mohamed, for his moral support and encouragement.

My appreciation to Mrs. Pat Reader for typing this thesis.

CHAPTER I

INTRODUCTION

1.1 INTRODUCTION

The availability and quality of demographic statistics in developing countries are far from adequate. According to a study conducted by the United Nations for 1951-55 (United Nations 1966) only about 33% of the world deaths and 42% of the world births were being registered. The situation has not changed greatly since, another more recent study (Brass et al 1968) concerned with the demography of tropical Africa pointed out that in few regions of tropical Africa there is almost no information even on the size of the population and though most countries of tropical Africa have sort of vital registration it is usually of very limited coverage and doubtful accuracy. The latter statement apply to the majority of developing countries.

The lack of accurate demographic statistics in developing countries is one of the obstacles facing their development programs. No detailed targets may be set without a realistic knowledge of the present demographic status of the population and their growth potential.

The straightforward solution of establishing new sources of basic statistics - if they are non-existent - or of improving the existing ones, may not always prove feasible. An introduction of a comprehensive vital registration system or conducting a full scale census may be too expensive as compared to the uncommitted resources in these countries, and even when such systems are available it is generally agreed that attainment of high quality data is a gradual process. In other words, the difficulties of improving the traditional sources of vital statistics lies in the cost and organizational constraints involved and it is more likely that economic development is a pre-requisite for any such improvements.

The intermediate approach where the collection procedure relies on sampling has a quicker pay-off in producing the needed data, this approach is advantageous not in terms of cost alone but also in the detailed and untraditional type of data it may supply. On the other hand, the data collected still suffer from the usual deficiencies which characterizes demographic data of developing countries in addition to sampling error.

No matter what collection procedure is used, the development and extension of methods which ensure a better utilization of the data is indispensible. The pertinent literature is quite large and a full account of such methods is not attempted but rather a brief review of some of the available methods for estimating adult mortality from defective data is presented. The emphasis is on mortality as it is one of the basic components of population change.

1.2 BRIEF REVIEW OF SOME OF THE AVAILABLE METHODS FOR ESTIMATING ADULT MORTALITY FROM DEFECTIVE DATA

The methods of estimation available differ greatly in their precision, underlying assumptions, costs and data requirements. There is no mechanical way in which any of them may be applied. The detailed procedure is sensitive to the characteristics and types of error in the data and there is always a demand imposed on the researcher's skills whether in manipulating the data or modifying the methods. The following is a presentation of the general methodological principles.

Different classification systems may be attempted but neither constitute a clear cut boundary. In this section the methods of estimation are divided into three categories; the first is mainly dependent on age distribution data. This type of data - traditionally available through censuses - reflects the cumulative results of past demographic flows and consequently offer a base for estimating them. The second category is more dependent on unconventional type information which are related to deaths in various ways. Finally, the third category include methods attempting to correct is advantageous not in terms of cost alone but also in the detailed and untraditional type of data it may supply. On the other hand, the data collected still suffer from the usual deficiencies which characterises demographic data of developing countries in addition to sampling error.

No matter what collection procedure is used, the development and extension of methods which ensure a better utilization of the data is indispensible. The pertinent literature is quite large and a full account of such methods is not attempted but rather a brief review of some of the available methods for estimating adult mortality from defective data is presented. The emphasis is on mortality as it is one of the basic components of population change.

1.2 BRIEF REVIEW OF SOME OF THE AVAILABLE METHODS FOR ESTIMATING ADULT MORTALITY FROM DEFECTIVE DATA

The methods of estimation available differ greatly in their precision, underlying assumptions, costs and data requirements. There is no mechanical way in which any of them may be applied. The detailed procedure is sensitive to the characteristics and types of error in the data and there is always a demand imposed on the researcher's skills whether in manipulating the data or modifying the methods. The following is a presentation of the general methodological principles.

Different classification systems may be attempted but neither constitute a clear cut boundary. In this section the methods of estimation are divided into three categories: the first is mainly dependent on age distribution data. This type of data - traditionally [illegible] reflects the cumulative results of past demographic flows and consequently offer a base for estimating them. The second category is more dependent on unconventional type information which are related to deaths in various ways. Finally, the third category include methods attempting to correct is advantageous not in terms of cost alone but also in the detailed and untraditional type of data it may supply. On the other hand, the data collected still suffer from the usual deficiencies which characterizes demographic date of developing countries in addition to sampling error.

No matter what collection procedure is used, the development and extension of methods which ensure a better utilization of the data is indispensible. The pertinent literature is quite large and a full account of such methods is not attempted but rather a brief review of some of the available methods for estimating adult mortality from defective data is presented. The emphasis is on mortality as it is one of the basic components of population change.

1.2 BRIEF REVIEW OF SOME OF THE AVAILABLE METHODS FOR ESTIMATING ADULT MORTALITY FROM DEFECTIVE DATA

The methods of estimation available differ greatly in their precision, underlying assumptions, costs and data requirements. There is no mechanical way in which any of them may be applied. The detailed procedure is sensitive to the characteristics and types of error in the data and there is always a demand imposed on the researcher's skills whether in manipulating the data or modifying the methods. The following is a presentation of the general methodological principles.

Different classification systems may be attempted but neither constitute a clear cut boundary. In this section the methods of estimation are divided into three categories; the first is mainly dependent on age distribution data. This type of data - traditionally available through censuses - reflects the cumulative results of past demographic flows and consequently offer a base for estimating them. The second category is more dependent on unconventional type information which are related to deaths in various ways. Finally, the third category include methods attempting to correct the deficient data on deaths.

1.2.1 The First Category

Methods in this category received wider application and more extensive discussion in the literature than the other two categories. This is probably due to their early development and to the fact that - until recently - their data requirements were more abundantly available and of better quality than data on deaths in developing countries.

Method (A): Stable and Quasi Stable Population Analysis

The method assumes that the age distribution of a given population may be approximated by a stable model. On the basis of the evidence available, a suitable model is picked and the various parameters of this stable model are assigned to the actual population.

If the actual situation is poorly approximated by the stable model or if the data available are too defective to permit a proper choice of a stable model, the estimates reached may be quite erroneous.

The term quasi-stability is used to indicate that fertility is the determinant factor in shaping the age distributions; thus the age structure of the population with declining mortality and constant fertility may still be approximated by a stable distribution.

Of course, if this principle was completely true, stable analysis would be unsuitable for supplying adequate information on mortality. Actually, the effect of the decline in mortality on deviation from stability is strongly related to the age pattern of change in mortality.

Various correction procedures have been devised to adjust for the effect of mortality change, such as Coale & Demeny (United Nations, 1967), (Zachariah, 1971) and quite recently (Abou Gamrah, 1976). Coale method requires knowledge of the duration of mortality decline and its average pace, Zachariah and Abou Gamrah discuss only the effect of deviation from stability on the birth rate.

Quasi stability played an important role in supplying some of our current knowledge on demographic trends, especially when data were scarce and nothing else could have been done. Improvements in the volume and quality of data and introduction of new techniques is reducing more and more the need for stable techniques.

Method (B): Census Survival Rates

This method presupposes the existence of two censuses or two cross-sectional demographic surveys. In this case the identification of birth cohort is easily achieved and the depletion of each cohort reflects the effect of mortality on that cohort. Thus census survival rates are calculated and life tables constructed.

This principle is true if the population is closed, the two censuses are the same in coverage and completeness and finally age errors are either non-existent or the same for any particular cohort at different points in time.

If migration occurred during the intercensal period or if the census coverage is not the same, erroneous estimates may be reached when this method is used.(unless, of course, corrections for these factors have been applied). Brass (1975) points out that: 'an analysis of intercensal mortality in Thailand between 1937 and 1970 produced estimates that suggested (unreasonably) a decrease in mortality followed by an increase. Further analysis revealed changes in census coverage over this period.

Unless adjustments can be made to allow for such changes in coverage, use of the intercensal method is unsuitable under these circumstances.'

Under the typical conditions of age misreport and differential under-registration in developing countries, survival rates show marked fluctuations which are unlikely a true feature of mortality and they are frequently higher than one. The usual procedure for dealing with these problems was either smoothing the original age distributions or the resulting survival rates. The introduction of model life tables allowed a further adjustment, a reasonable method is proposed by Coale & Demeny (United Nations 1967) in which cumulative survival rates are calculated and mortality levels corresponding to them are located through the use of model life tables, then an average level is selected as the estimate of mortality in the population analyzed.

Several problems are usually related with this method of estimation. The first is that the procedure of locating the corresponding model mortality level may become too complicated when the intercensal period is not a multiple of 5 years. The second is that it does not give a measure of mortality for the ages which are younger than the intercensal period, since a related birth cohort can not be identified in the first census. The third arises when the period between the two censuses is not a multiple of the age group length. The second problem is solved by using the model value in the average level selected as an estimate of mortality of these young ages; in other words the mortality pattern is forced to conform to one of the four patterns available in Coale & Demeny model life tables. The third problem - in the absence of detailed tabulations - may be solved through interpolation within the reported age distributions so as to form new age groups.

Brass (1972) proposes another procedure of applying this method, mainly by calculating $_nL_x$ (number of persons in the stationary population between exact ages x and x+n) using the census survival rates and initial values of $_nL_o$. Then, the values of $_nL_x$ are adjusted through the use of the logit system. In essence, Brass procedure depends on the use of the logit system. Thus, the first problem is made simpler because the logit system is more adaptable to adjustment and the second solved through imposing a more flexible model system.

1.2.2 The Second Category

When the direct recording of deaths is non-existent or greatly distorted by error; suitable questions in demographic inquiries may supply a substitute for the data traditionally provided by vital registration. In principle these inquiries may be complete census or cross-sectional surveys, but generally the latter is more appropriate when retrospective information are required, since the intensity and quality of field work needed for good quality information are more likely available for a smaller size operation.

A general characteristic of the methods in this category is their reliance on demographic models for transforming the unconceptional information into familiar types of data and for completing and filling gaps in the existing information.

One of the most famous and successful methods under this category deals with the estimation of infant and childhood mortality from proportions living among children ever born. (United Nations, 1967). The linking of this measure with adult mortality, in the absence of other information, is reached by imposing a certain pattern on the data.

Several reference sets of life tables (model life tables) are available and they may be used for linking childhood mortality with mortality at later ages. If there is no information about adult mortality, the extension of mortality is reached by using a one parameter model life table such as those of the United Nations (1955) or the Coale-Demeny system (1966). These model tables have been constructed by averaging recorded mortality patterns; thus one expects to obtain reasonable estimates of adult mortality only when the pattern of mortality studied confirm to this average pattern. Brass (1972) considers the case of Turkey where the relation of adult to childhood mortality is atypical; he points out that the extension of childhood mortality, from the 1963 retrospective survey, through the use of models results in an expectation of life at age 5 of about 47 years while other information shows the actual expectation to be in the region of 63 years.

Two parameters model life tables, such as the logit system (Brass 1964 and 1971) and some of Ledermann sets (Ledermann 1969), allow the age variations between mortality patterns to be explicitly brought out. Thus, if further information about adult mortality is available, the use of a two parameter system provides a suitable procedure for linking the available information.

Method (A): Orphanhood Method

This method attempts to obtain adult mortality estimates from data on the survivorship of parents. (Brass & Hill 1973). The principle of obtaining adult female (or male) mortality from maternal (or paternal) orphanhood may be presented as follows:

$$PR(x,t) = \frac{\int n(y,t-x)\ f(y,t-x)\ \frac{l(y+x,t)}{l(y,t-x)}\ dy}{\int n(y,t-x)\ f(y,t-x)\ dy} \qquad (1.1)$$

where:

PR(x,t) = proportion of children of exact age x whose mothers (fathers) were alive at time (t).

n(y,t-x): number of women (men) of childbearing age y at time (t-x).

f(y,t-x): female (male) age specific fertility rates at time (t-x).

l(y+x,t): life table survivors at age (y+x) at time (t).

Assuming that the age specific fertility and mortality rates remained constant over the required time period, then:

$$PR(x) = \frac{\int n(y)\ f(y)\ \frac{l(y+x)}{l(y)}\ dy}{\int n(y)\ f(y)\ dy} \qquad (1.2)$$

If the changes in the probabilities of survival from age y to y+x are linear - actually, the survival ratio curve is not linear but its curvature is not very pronounced - then the previous expression may be approximated as:

$$PR(x) = \frac{l(\bar{M}+x)}{l(\bar{M})} \qquad (1.3)$$

where $\bar{M}$ is the mean age of childbearing of mothers (fathers) in the population under consideration. Thus the proportion of children of exact age x whose mothers (fathers) are alive supply us with a direct estimate of the probability a female (male) aged $\bar{M}$ will survive x years.

For the last expression to be of practical value, two points need further discussion. The simplest is that PR(x) is usually available corresponding to age groups rather than exact ages x. The second is concerned with the use of $\bar{M}$ as the base age, since $\bar{M}$ may be any fractional age its direct use leads to survival rates corresponding to very irregular age intervals.

The first problem was tackled by using the following expression:

$$PR(i) = \frac{\int_{zl}^{zl+n} n(x)\ \int n(y)\ f(y)\ \frac{l(y+x)}{l(y)}\ dy\ dx}{\int_{zl}^{zl+n} n(x)\ \int n(y)\ f(y)\ dy\ dx} \qquad (1.4)$$

$$PR(i) = \frac{l(\bar{M}+e_i)}{l(\bar{M})} \qquad e_i = zl + \frac{n}{2} \qquad (1.5)$$

where (i) denotes the i^{th} age group and zl and zl+n denote the boundary of this age group. Thus PR(i) is an estimate for the survival rates between age $\bar{M}$ and $\bar{M}+e_i$, where e_i is the median age of each age group of children.

The second problem was dealt with by calculating the appropriate correction factors $c_{\bar{M}_i}$ such that:

$$PR(i) = c_{\bar{M}_i} \frac{l_{b+e_i}}{l_b} \qquad (1.6)$$

where b is an appropriate base age and $c_{\bar{M}_i}$ accounts for the difference between the use of $\bar{M}$ and b. The correction factors were calculated for different values of e_i and $\bar{M}$ by using standard models for the mortality and fertility functions and an analytical form for the age distribution.

Though the principle of estimating adult female or male mortality from orphanhood data is the same, in practise the estimation of female mortality is more rewarding. This is due to several factors, mainly the reproductive period for female is shorter and better defined, the shape and characteristics of the female age specific distribution are better known, the data for calculating $\bar{M}$ for males are generally unavailable and finally more is usually known about mothers than fathers so the data referring to female mortality are usually more accurate.

The general criticisms associated with this method are mainly directed to the underlying assumptions that there is no relation between mortality experience and number of surviving children, since those with no surviving shildren have no weight while those with several surviving children are given several weight (an approach for offsetting the latter bias may be using data on orphanhood from children with a specified birth order), also to the assumptions on which the weighting factors were based, and finally to the assumption of constant mortality and fertility.

The ultimate justification of the method lies in the plausible estimates it provides, at least concerning female mortality

Method B: Widowhood Method

A similar indirect set of measures, as in method (A), may be found in widowhood data as follows:

proportion of wives (husbands) aged y never widowed of first husband (wives) =

$$\frac{l(\text{mean age of first husbands (wives) at marriage at time } (t-a) + a)}{l(\text{mean age of first husbands (wives) at marriage at time } (t-a)} =$$

$$\frac{l(\text{mean age of first husbands (wives) corresponding to wives (husbands) aged } y \text{ at time } t)}{l(\text{mean age of first husbands (wives) at marriage at time } (t-a))}$$

where a denotes the length of marriage (exposure time).

A direct advantage of this method results from the fact that marriage takes place earlier and over a less dispressed age distribution than childbearing, thus the previous expression is more exact than expression (1·3) and also the standard error of the mean age of husbands (wives) at first marriage is much less than the corresponding standard error for the mean age of childbearing.

The previous expression is easily modified to correspond to age groups rather than fixed age y. The problem that the mean age of husbands at marriage are usually a fractional age and thus result in irregular survival ratio may be dealt with through a knowledge of the bivariate distribution of ages at marriage of men and women.

Hill (1975) used a standard mortality model and simple functions for the distribution of ages at marriage and calculated a set of correction factors similar to the one used in the orphanhood method. These factors depend on two measures, the mean age of marriage of the cohort of women (men) and the period mean age of marriage of men (women); these measures may be approximated using the available data on the proportion of persons single by age group and the age distribution.

The criticism associated with this method are basically similar to the orphanhood method. First that there is no relation between the mortality experience and marital status. Of course, the previous bias is much less in the widowhood data since married persons are usually more numerous than parents. The basic criticism is directed to the assumptions underlying the correction factors and their feasibility.

1.2.3 <u>The Third Category</u>

Data on deaths, whether through vital registration systems or other sources, are available in many developing countries but the quality of data is such that no great confidence may be placed on their direct use. The data suffer from under-registration and age misreport and there is always a need for methods which tackle these problems.

The methods available are two types, the first attempts to detect the error through a comparison of some sort, either checking for internal consistency or using a different set of data. The second type include methods directed to offset the effect of the main source of error, under-registration, on the measure of mortality.

The First Type:

One of the most sophisticated of these methods require the existence of a dual system of recording and the comparison between them is performed at the level of the smallest unit through individual matching of events. The formula for estimating the total number of events, on condition that the recording in the two systems is independent, is given by Chandra Sekar & Deming (1949) as: $\hat{N} = \frac{N1 \times N2}{N12}$, where $\hat{N}$ is the estimated number of events, N1 and N2 are the total recorded in the first and second system, and N12 those common to both as determined by matching.

The previous procedure demands substantial expenditure and high level of organization skills and naturally is only performed on a sample basis.

Another method may be the use of a sample survey to estimate the completeness of the existing vital registration system. If the comparison between the sample survey and the vital registration require matching, then this method is simply the Chandra Sekar method and the advantage gained by introducing only one new recording system is offset by the extra cost and difficulty in matching the events. If it is possible to rely on responses in the sample survey about possession of death certificates or registration of events, then a measure of the completeness of the registration may be reached.

A further method depends on an internal comparison of the data, for example if it is believed a certain area experiences higher mortality than another while the data contradict this, the data corresponding to the higher mortality area may either be neglected or modified. This method has been often used in connection with rural and urban mortality, for example E.L. Badry (1965) used this principle on Egyptian data in raising the mortality of rural areas without health bureau to match the level of mortality in similar areas but with health bureau. The difficulty in this method is that the rules for rejecting or modifying some data need to be based on close knowledge of the population under study, since if the reasonings for altering the data are not present a new source of bias is introduced.

Another method compares the data supplied by two censuses and vital registration data to adjust for the discrepancies in the registration data. Under the assumption of equal under-registration with age in the first and second census and the vital data, Bourgeois-Pichat (1957) showed that, in a closed population, the difference between the calculated population expected at the time of the second census (using the first census and the registered deaths) and the reported population at the second census is a function of this under-registration.

$$E1^{*}_{i} = P^{*}_{1i} \left(\frac{h^{*}-k^{*}}{1-k^{*}}\right) + D^{*}_{i} \left(\frac{g^{*}-h^{*}}{1-g^{*}}\right)$$

where $E1^{*}_{i}$ is the difference between the calculated and reported population for age groups i, P^{*}_{1i} the reported population in the first census in the i^{th} age group, D^{*}_{i} the reported deaths corresponding to the cohort in the i^{th} age group in the intermediate period. K^{*} and h^{*} denotes the proportionate under-registration in the first and second census respectively and g^{*} the under-registration of deaths.

If the assumptions are correct the intersection of the lines formed using the data for each age group gives an estimate for $\left(\frac{h^{*}-k^{*}}{1-k^{*}}\right)$ and $\left(\frac{g^{*}-h^{*}}{1-g^{*}}\right)$. Actually, due to the differential under-registration by age and age misreport the lines don't all intersect in one point, also this method gives only an estimate of the magnitude of the under-registration in the vital data in terms of the under-registration in the census (since more information is required to solve two equations in three unknowns) and finally the identification of deaths corresponding to each cohort may prove to be too complicated.

The Second Type

Under-registration of deaths is a common deficiency in the data of developing countries. It is generally agreed that reporting of young children in the first year of life is more strongly affected by this type of error.

Earlier attempts to correct under-registration of young children depended on comparing the mortality of age group (1-4) years with the mortality of age (0-1) through the use of model life table. Unless there is enough evidence for accepting a certain pattern of mortality, this method may be used as a rough indication of the possibility of under-recording.

Bourgeois-Pichat (United Nations 1952) presented a method for estimating mortality in the early months of life from the trend over the rest of the first year of age. This method requires a detailed tabulation for death rates by month of life to be available.

Several methods are presented to correct under-registration of deaths for adult ages, all the methods assume that after a certain age the proportion of deaths that are not reported is constant, the simple but effective idea that the proportionate death distribution is not affected by a constant under-registration is employed.

Method (a): Carrier Method

Carrier (1958) showed that in a stable population

$$\frac{\sum\limits_{X=x} \frac{dl_x}{(1-r)^x}}{\sum\limits_{X=o} \frac{dl_x}{(1-r)^x}} = \frac{l_x}{l_o} \tag{1.7}$$

where, dl_x: actual number of deaths at age x r: rate of growth

and l_x: life table survivors at age x

In case of equal under-registration of deaths after age y, the formula may be expressed as:

$$\frac{\sum_{X=x} \frac{dl_X}{(1-r)^X}}{\sum_{X=y} \frac{dl_X}{(1-r)^X}} = \frac{l_x}{l_y} \qquad x \geq y \qquad (1.8)$$

When X denotes age group rather than exact age, the formula is easily modified.

The main problem associated with this method is the requirement that r should be known; Carrier suggested that if r was not known, calculations of l_x could be made for a series of trial values of r and the resulting l_x in best accord with the United Nations model life table system may be accepted.

Carrier showed that his method is sensitive to changes in r and patterns of mortality; he accepted that the results are reasonably accurate only if r is estimated within one or two per thousand.

<u>Method (b): Fargues and Courbage Method</u>

Fargues & Courbage (1972) used the relation

$${}_nm_x = \frac{dl(x,x+n).TD.Tot}{TD.Tot.n(x,x+n)} \qquad (1.9)$$

where, ${}_nm_x$: death rate corresponding to age group (x-x+n).

dl(x,x+n): number of deaths in age group (x-x+n).

TD: total number of deaths.

Tot: total population

n(x,x+n): number of persons in age group (x-x+n).

In case of equal under-registration dl(x,x+n)/TD may be approximated by the proportionate death distribution. Using the available data on the age distribution all is needed to estimate the age specific death rates is an approximation for $\frac{TD}{Tot}$.

Using the hypothesis that in several populations with the same age pattern there is a negative correlation between the mortality level and the ratio of deaths at old ages to deaths at all ages, an approximation for TD/Tot was calculated as follows:

- adjusted death rates are calculated using the age specific death rates in several countries and the age distribution of the country studied. These countries are carefully picked as to be similar in conditions to the one studied and with reliable statistics.
- For each of the countries picked a relation between deaths at old ages divided by deaths after age 5 and the adjusted death rates is established. Deaths after age 5 rather than total deaths are used, since it is expected that deaths at young ages suffer from a higher under-registration.
- Using the available information on deaths at old ages divided by deaths at age after 5 in the country studied and the relation established in the previous step, an approximation for TD/Tot may be reached.

The main difficulty of this method lies in the choice of appropriate countries to determine the estimation relation as there may be considerable difference in the estimate according to the mortality pattern chosen as standard. Another difficulty arises from the errors of age reporting, especially for old ages; thus the ratio of reported deaths at old ages to all deaths may be different from the actual ratio.

Method (c): The Growth Balance Method

Brass (1974) showed that in a stable population:

$$N_y = r P_y + d D_y$$

where, N_y: population proportion age y

P_y: population proportion over age y

D_y: proportion of deaths over age y

r: rate of natural increase

CDR: crude death rate

Assuming equal under-registration, in the reported deaths, from a certain age; the previous relation holds from that age upward.

Thus using the reported age and death distribution the crude death rate and the rate of growth may be calculated. In essence this method is a modification of Carrier's method; it supplies the extra information needed (growth rate) through the use of the available age distribution.

The difficulties associated with this method is the effect of deviations from stability on the estimate and the possible effect of age misreport and differential under-registration.

1.3 OBJECTIVES AND OUTLINE OF THE STUDY

From the previous presentation we note that the first category of methods received the wider and earlier application in demographic analysis. The use of sample surveys by many developing countries and the realization that the census need not be limited to the traditional type of data - but may include suitable questions which supply direct information on past events - shifted the importance to the second category of methods.

The accuracy of these methods depends on many factors and there is still need for more applications and checks to allow for a better judgement of their reliability and to justify more sophisticated modifications. The main advantage of these methods which is their dependence on simple type of questions may be considered one of their drawbacks; if the special questions they require were not included in the study then the road may be blocked.

The third category of methods, more precisely the growth balance method included in the second type, still needs further discussion. It makes use of data already available and is basically quite simple to apply. More exploration is required since the theoretical assumptions of this method are never fully satisfied. Mortality and fertility are changing in nearly all developing countries, migration plays an important role in some of these countries; thus it is quite important to study the effect of such changes and the possibility of modifying the method when the population is not stable. Also, this method assumes equal under-registration of deaths after a certain initial age, but since the registration of deaths of very young ages is usually different from other ages, the estimate of the death rate corresponds to certain ages only; it is extremely valuable if an allowance is made in this method for the differential under-registration. Age error is a feature of reporting in developing countries, effect of this error on the method is significant in judging its appropriateness. It is the purpose of this study to discuss all these topics as well as some practical considerations such as the best method of fit that may be used and the question of smoothing the data before applying the method.

In Chapter (2) the growth balance method is discussed in detail and a modification in applying the method suggested. This modification, though quite simple, may prove to be helpful in some cases.

Chapter (3) focuses the attention on the effect of changes in mortality and fertility on the method; it deals with this problem using two approaches. The first through standard population projection using different patterns of mortality decline and a comparison of the projected death rate and the estimated applying the growth balance method. The second approach is based on the analytical relation between the age structure and the changing schedules of mortality and fertility; several theoretical modifications are introduced for certain patterns of change in mortality and fertility and also hypothetical applications of this modifications are presented.

Chapter (4) discusses the effect of migration on the death distribution method, an adjustment procedure is presented and illustrated using actual data.

Chapter (5) deals with the problem of unequal under-registration of deaths. A method for adjusting the estimated death rate in case of unequal proportionate under-registration is proposed. Numerical applications are presented. The effect of the differential under-registration on the graph of the sets of points $(\frac{N_y}{P_y}$ & $\frac{D_y}{P_y})$ is discussed and also illustrations of the magnitude of error likely to affect the estimate as a result of different combinations of under-registration are given.

Chapter (6) tackles the problem of age reporting in developing countries. First, a model of age error is discussed in general. Then, the range of likely bias introduced in the estimate of the death rate due to age error is shown under two different assumptions. First, the type of age error is the same in both the age and death distribution; second, the error is different. Finally, the effect of graduating the data before applying the growth balance method is discussed.

In Chapter (7) several methods of fitting straight lines are presented and the best methods to be used when applying the growth balance technique suggested.

In Chapter (8) application of the growth balance method on data affected by several deviations from the assumptions are presented, and the conclusions reached in the previous chapters illustrated.

CHAPTER II

ALTERNATIVE FORMULA FOR APPLYING THE GROWTH BALANCE METHOD

2.1 INTRODUCTION

In this chapter we will discuss the growth balance method in detail. The proof of the method and some practical considerations are presented. Another formula for applying the method is suggested; numerical applications of this formula both on stable and quasi stable age distributions are illustrated and also a criterion for using the new formula is presented.

2.2 DETAILED PRESENTATION OF THE GROWTH BALANCE METHOD

An intuitive presentation of the growth balance method starts by considering the simple basic relation:

$$\text{birth rate} - \text{death rate} = \text{growth rate}$$

$$b - CDR = r$$

This relation holds from any initial age upwards, thus:

$$b_y = r_y + CDR_y.$$

If the population is stable, r_y denotes the intrinsic rate of growth and is constant for all ages. Thus, we reach the basic formula of the growth balance method:

$$b_y = r + CDR_y \qquad (2.1)$$

In other words, in a stable population the death rate and the birth rate over age y form a straight line with slope 1.

If the registration of deaths is incomplete, but is the same for all ages over y, the slope of the line is no longer 1. The ratio of actual to reported deaths may be estimated as the slope of the line formed using the reported death rate and the birth rate for ages over y. Thus, equation (2.1) may be rewritten as:

$$\frac{n_y}{P_y} = r + f\,\frac{d_y}{P_y} \qquad (2.2)$$

where, n_y: number of persons at age y

P_y: number of persons over age y

r: the growth rate

f: ratio of true to reported deaths

d_y: number of deaths over age y.

Another convenient way of writing (2.2) is:

$$\frac{N_y}{P_y} = r + CDR\,\frac{D_y}{P_y} \qquad (2.3)$$

where P_y and D_y are the proportions at risk and dying after age y and N_y the proportion of persons at age y.

A mathematical proof of equation (2.3) is presented as follows:

$$D_y = \frac{\text{number of deaths over age y}}{\text{total number of deaths}} = \frac{\int_y^w n_x\, u_x\, d_x}{CDR.Tot}$$

where, u_x: force of mortality at age x

Tot : total population.

Then,

$$CDR.D_y = \int_y^w N_x . u_x . d_x = -\int_y^w N_x \frac{l'_x}{l_x}\, dx$$

where, l_x: life table survivors at age x when $l_o = 1$

l'_x :the first derivative of l_x with respect to x

Integrating by parts, we get:

$$CDR.D_y = -l_x \frac{N_x}{l_x}\Big|_y^w + \int_y^w l_x \left(\frac{N_x}{l_x}\right)' dx$$

$$CDR.D_y = N_y + \int_y^w N_x \, d\log \left(\frac{N_x}{l_x}\right)$$

$$N_y = CDR\ D_y - \int_y^w N_x \, d\log \left(\frac{N_x}{l_x}\right) \qquad (2.4)$$

equation (2.4) is a general equation which holds for any age distribution.

If the population is stable, then $N_x = be^{-rx}l_x$ and:

$$N_y = CDR\ D_y + r\ P_y$$

$$\frac{N_y}{P_y} = r + CDR\ \frac{D_y}{P_y}$$

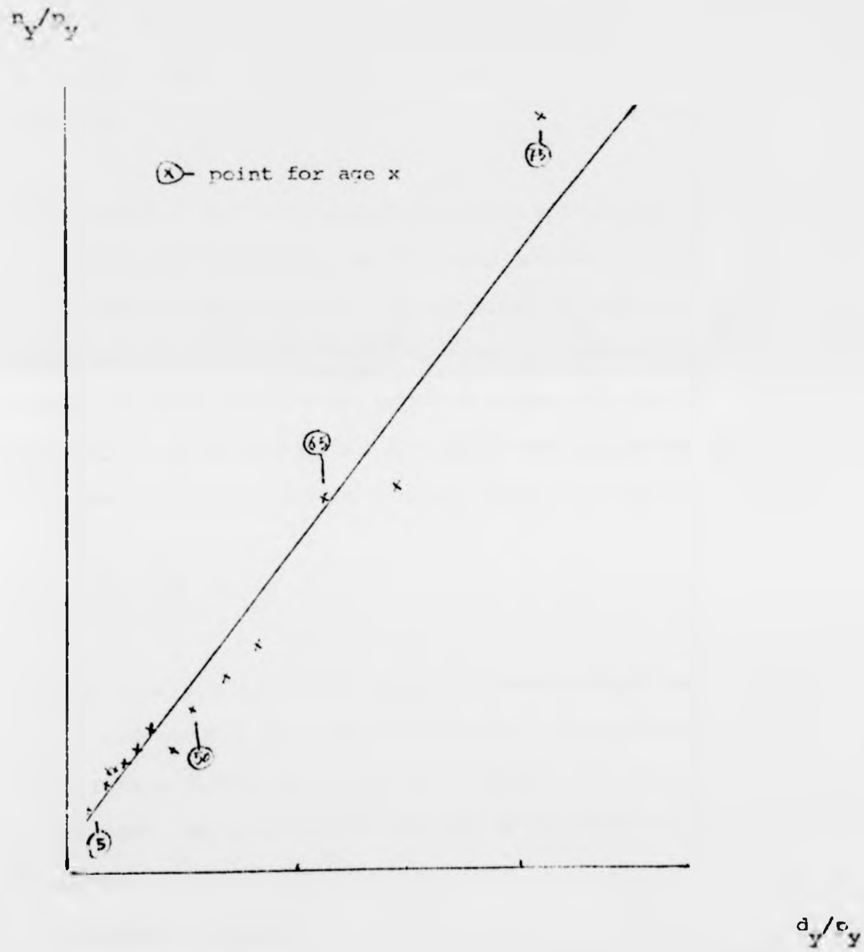
If the proportionate under-registration in the data is equal, the previous equation may be used to estimate the actual crude death rate in stable populations. Actually, this equation is valid for any fixed age upwards. Thus it is not necessary to assume that the reporting of the deaths of young children is as good as that at older ages and the equation may be used to estimate the death rate of adult ages in case of differential under-registration.

In practise the data are available corresponding to age groups, thus D_y and P_y are readily calculated; N_y has to be estimated. Sophisticated techniques may be used, but a very simple and usually accurate procedure is to take N_y as the average of the age groups on either side of y. Thus, if the length of the age group is n:

$$N_y = \frac{1}{2n} (N_{y-n,y} + N_{y,y+n}).$$

$\frac{N_y}{P_y}$ is plotted against $\frac{D_y}{P_y}$ and a straight line is fitted through the points.

Graph (2.1)* Estimation of adult mortality from the relation of death and population age distributions Jordan, 1961: females



n_y/p_y = population per year at age y/number at risk after age y

d_y/p_y = number dying after age y/number at risk after age y

*Soure: Brass (1974)

Age distributions in developing countries may not be stable and also the data are affected by age misreport and differential under-registration, thus the points corresponding to $(\frac{B_y}{P_y}$ & $\frac{D_y}{P_y})$ do not form an exact straight line. If the points lie closely on a straight line, the assumptions are given support and there can be confidence in the results; if the points show clear signs of curvature the derivation of the measure is hardly possible.

The fitting of the best straight line is not easily accomplished. Several procedures are available, but the ones usually used are least square and the average (Wald) methods. In the first method the death rate is estimated as: $CDR = \frac{\Sigma(x-\bar{X})(y-\bar{Y})}{\Sigma(x-\bar{X})^2}$, where $X = \frac{D_y}{P_y}$ and $Y = \frac{B_y}{P_y}$. In the Wald method the data are divided into two groups of size n1 and n2 respectively (n1 and n2 may be equal) and the death rate estimated as the slope of the straight line drawn through the mean points of the two groups:

$$CDR = \frac{\bar{Y}_2 - \bar{Y}_1}{\bar{X}_2 - \bar{X}_1}$$

Several applications of the growth balance method are available. Brass (1974) and (1976), applied the procedure successfully to data on Jordan and Iraq, Potter (1976) applied it to Colombia and Blacker (1977) also applied it to Chad. An illustration for the application of the method on actual data is given in Graph (2.1).

2.3 ALTERNATIVE FORMULA

In applying Brass method for mortality estimation on several stable distributions it was noticed that in some instances the estimated death rate differs considerably from the actual death rate. For example, applying the method on the stable distributions given in Coale & Demeny (1966), model west, level 1, 3, 5, 7, 9, 11, 13 and 15, we get the following results: (see Table (2.1))

Table (2.1) The actual death rate corresponding to model west, males, for different levels and growth rates and the estimated death rate using the death distribution method

level	growth rate	actual CDR	estimated CDR
1	10	59.77	67.63
3	10	45.94	49.85
5	10	36.85	38.90
7	25	32.04	33.55
9	25	25.97	26.73
11	25	21.25	21.76
13	25	17.30	17.51
15	25	14.21	14.24

Though from level 7 to 15, the estimate does not deviate considerably from the actual death rate, the estimates corresponding to level 1, 3 and 5 are greatly distorted.

It is our purpose to discuss the reason for this distortion and suggest an alteration which helps to improve the estimate.

a - The reason for the distortion:

In applying Brass method on the hypothetical data all assumptions were met, the population is stable, no differential under-registration and no age misreport. The reason for the distortion may be attributed to the method of estimating N_y. N_y was estimated by assuming that the age distribution is linear such that:

$$N_y = (N_{(y-5)-y} + N_{y-(y+5)})/10$$

Equation (2.2) where N_y is estimated assuming linearity will be denoted formula (A). Probably formula (A), though generally acceptable, does give rather distorted estimate of the actual death rate when the age distribution (especially for old ages) deviates considerably from linearity).

b - Formula (B):

Since: $\frac{N_y}{P_y} = r + CDR\ \frac{D_y}{P_y}$

integrating both sides from the start to the end of the interval:

$$\int_y^{y+n} N_y\ d_y = r \int_y^{y+n} P_y + CDR \int_y^{y+n} D_y$$

$\int_y^{y+n} N_y = A_y$ = proportion within the age group

$$\int_y^{y+n} P_y = n\ P^*_y \qquad\qquad P_{y+n} < P^*_y < P_y$$

$$\int_y^{y+n} D_y = n\ D^*_y \qquad\qquad D_{y+n} < D^*_y < D_y$$

assuming the cumulative distributions are linear within the age interval, then:

$$P^*_y = (P_{y+n} + P_y)/2$$

$$D^*_y = (D_{y+n} + D_y)/2$$

Thus:

$$A_y = n.r\ \frac{(P_y + P_{y+n})}{2} + n.CDR\ \frac{(D_y + D_{y+n})}{2}$$

$$\frac{A_y}{n.P^*_y} = r + CDR\ \frac{D^*_y}{P^*_y} \qquad (2.5)$$

This last formula will be denoted formula (B).

Applying formula (B) on the previous stable distributions we get:

Table (2.2) The actual death rate corresponding to model west, males, for different levels and growth rates and the estimated death rate using both formula (A) and formula (B).

level	growth rate	actual CDR	estimated CDR	
			formula (A)	formula (B)
1	10	59.77	67.63	59.05
3	10	45.94	49.85	45.76
5	10	36.85	38.90	36.15
7	25	32.04	33.55	30.48
9	25	25.97	26.73	24.82
11	25	21.25	21.76	20.64
13	25	17.30	17.51	16.75
15	25	14.21	14.24	13.70

It is clear that formula (B) helps to correct the distortion to a great extent; on the other hand corresponding to level 7, 9, 11, 13 and 15 formula (A) gives a slightly better estimate. Though the difference between the two estimates in the latter is not that significant, nevertheless it is important to show that, in general:

- If formula (B) does not improve the estimate considerably it does not affect it to a great extent.
- a criterion exists to choose between both formulae in application when the actual death rate is not available.

The first point may be illustrated by applying both formula (A) and (B) on several age distributions and comparing the two estimates of the death rate.

If the formulae are applied to data subject to mortality decline rather than stable data, the illustration may be more realistic. In chapter (3) the effect of mortality decline on the growth balance method is discussed in detail; table (2.3) is an extract from the results given in chapter (3) in tables (3.4), (3.5) and (3.6). Table (2.3) presents the actual death rate and the growth balance estimate using both formula (A) and formula (B) when the data are subject to different patterns of decline in mortality (these patterns and their implications will be dealt with in detail later). Our point of concern here is that the difference between the estimates using formula (A) and (B) is big only when formula (B) is better, while the difference between the estimates is always small when formula (A) is better. Thus, as a general rule the use of formula (B) is recommended.

In actual situations, it is advisable to apply both formula (A) and (B) and to use the graph as the criterion for choosing the best estimate. For example, the following graphs illustrate the two lines formed by plotting both:

$\frac{N_y}{P_y}$ & $\frac{D_y}{P_y}$ denoted by the symbol ○

$\frac{A_y}{n\ P^*_y}$ & $\frac{D^*_y}{P^*_y}$ denoted by the symbol △

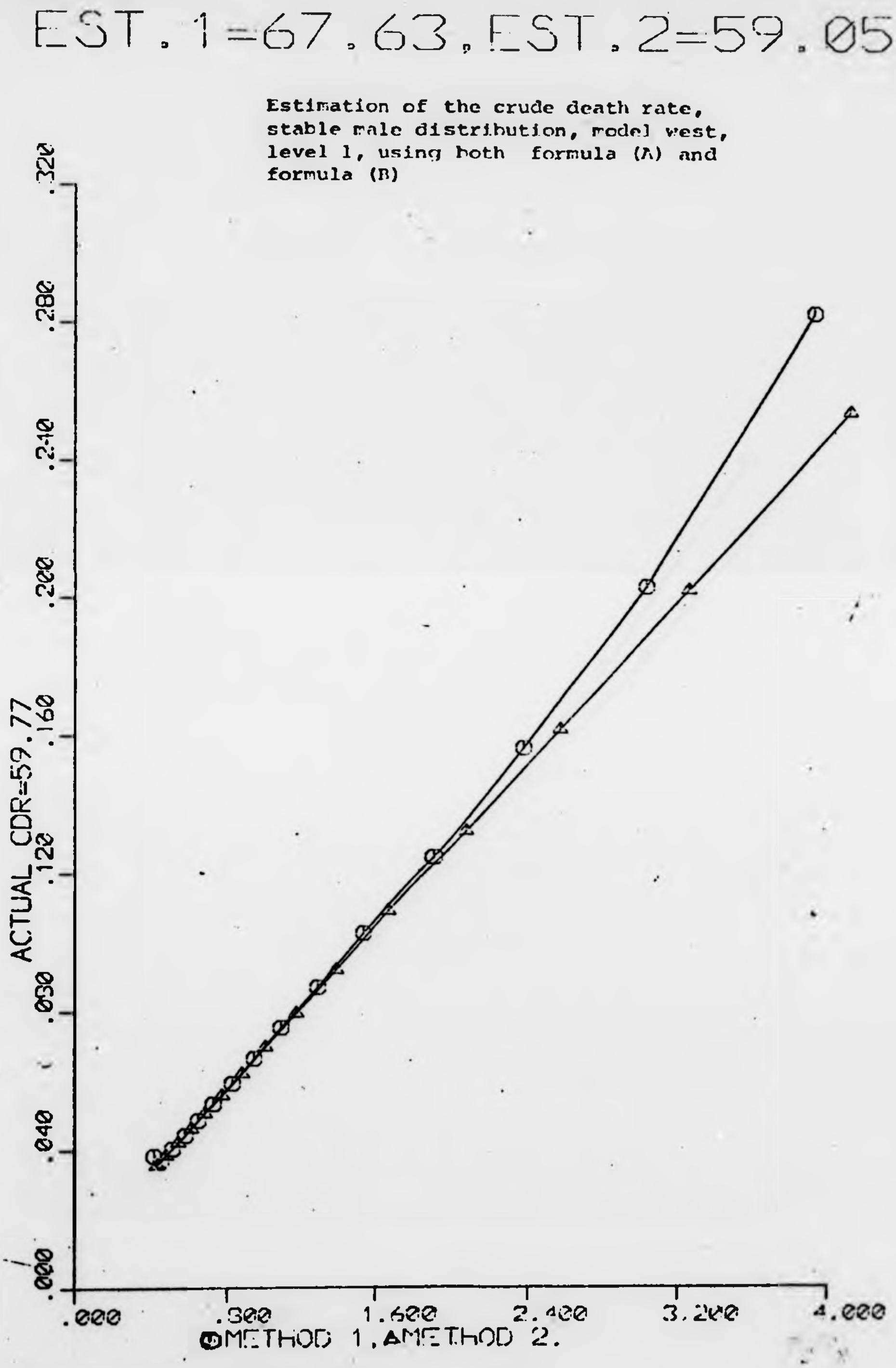
Graph (2.2), (2.3) and (2.4) represents the case of stable male distributions, model west, level 1, 3 and 5 respectively. Graph (2.5), (2.6) and (2.7) represents the first three results in the case of mortality decline according to pattern (1.6). In all the graphs the more linear the line the closer the estimate to the actual death rate.

Table (2.3)* Comparison between the estimates of the death rate using formula (A) and (B)

pattern (1.a)			pattern (1.b)			pattern (2.a)		
actual CDR	estimated formula (A)	estimated formula (B)	actual CDR	estimated formula (A)	estimated formula (B)	actual CDR	estimated formula (A)	estimated formula (B)
44.55	48.46	44.39	32.95	37.90	32.83	34.97	35.48	34.64
37.54	40.52	37.14	30.26	35.82	31.02	34.79	33.04	32.50
31.09	33.40	30.56	27.75	33.65	29.23	34.35	29.99	29.66
25.32	27.18	24.79	25.39	31.42	27.47	31.96	29.34	28.61
20.38	21.94	19.95	23.17	29.15	25.72	27.68	28.90	27.71
16.19	17.56	15.93	21.08	26.88	23.99	24.67	26.41	25.40
12.64	13.84	12.55	19.12	24.66	22.27	23.01	23.45	22.81
9.56	10.62	9.66	17.29	22.50	20.54	20.63	20.49	20.00
6.96	7.88	7.22	15.59	20.42	18.82	17.50	17.12	16.65
4.80	5.59	5.20	14.01	18.39	17.07	14.85	14.08	13.55
3.04	3.73	3.52				12.70	11.86	11.34

* extract from tables (3.4), (3.5) and (3.6)

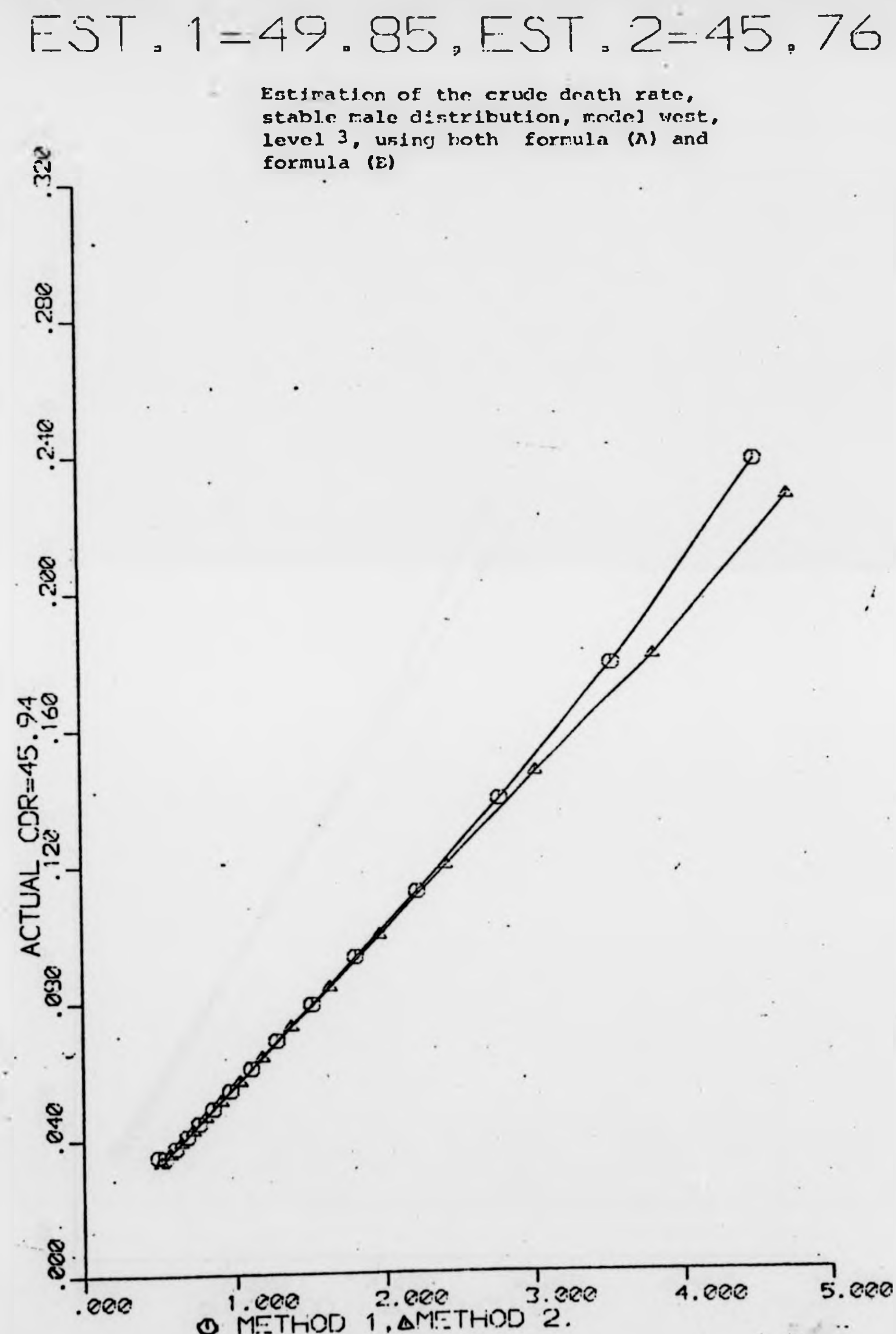
Graph (2.2)



method 1 formula (A)

method 2 formula (B)

Graph (2.3)



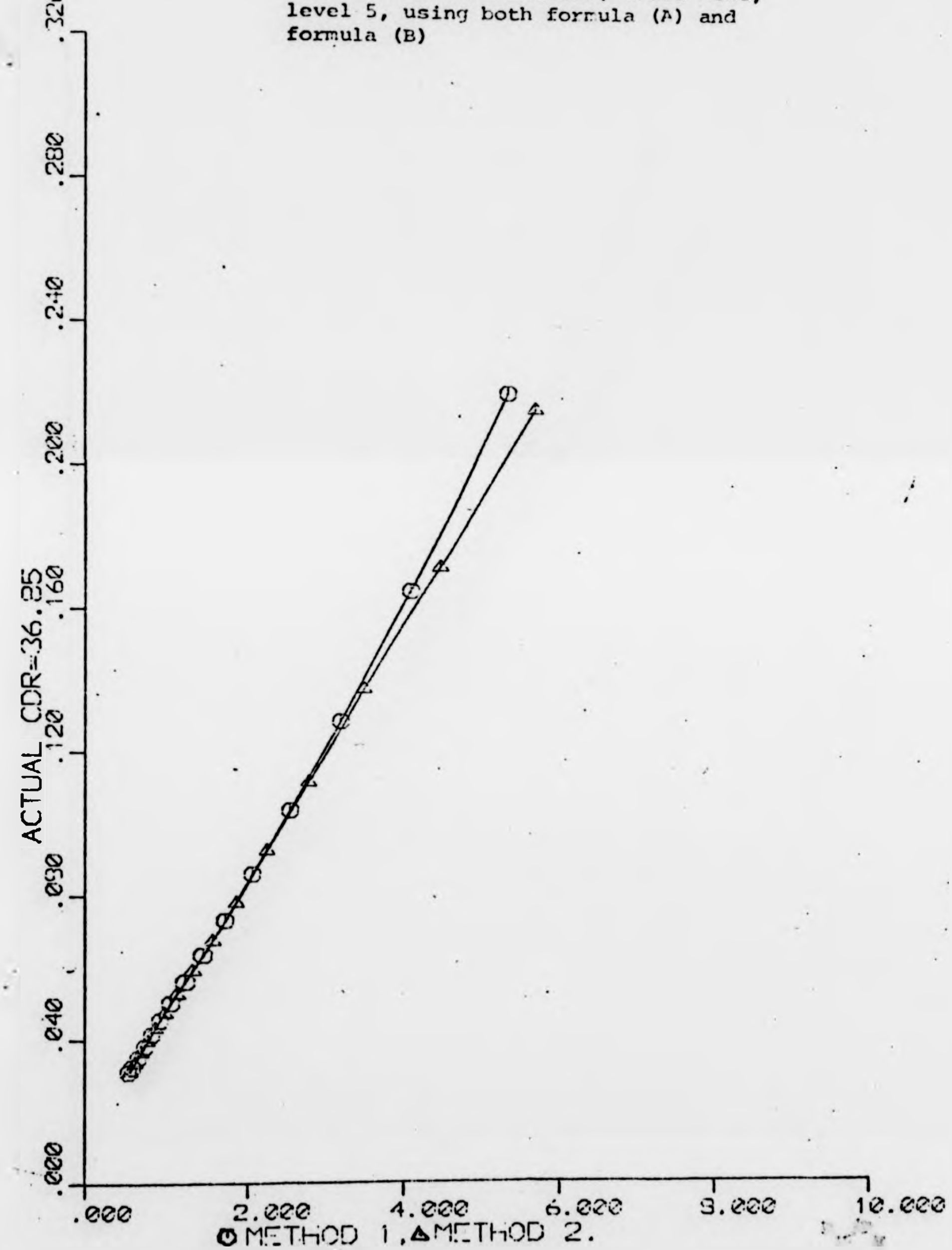
method 1 formula (A)

method 2 formula (B)

Graph (2.4)

EST.1=38.90,EST.2=36.15

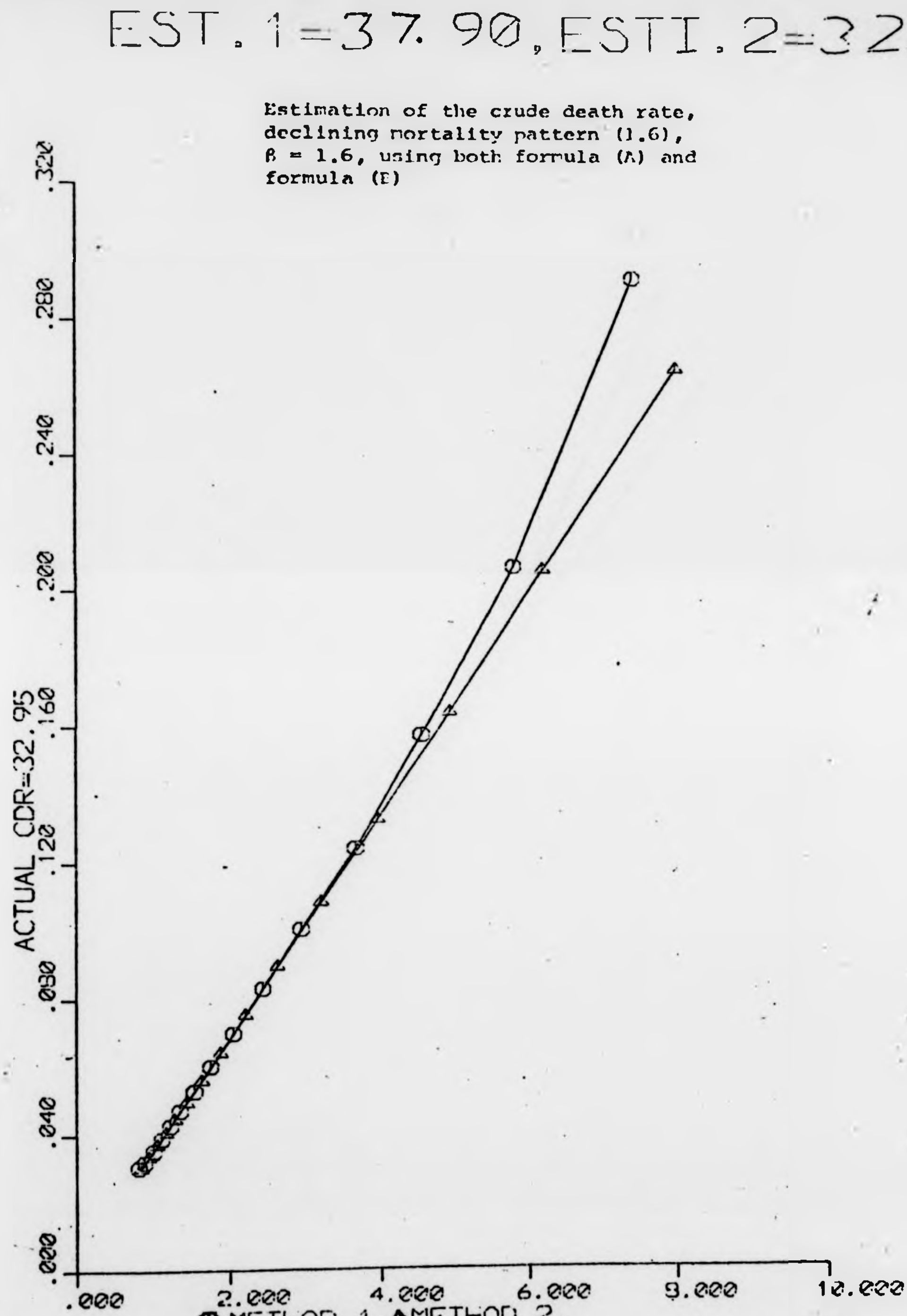
Estimation of the crude death rate, stable male distribution, model west, level 5, using both formula (A) and formula (B)



method 1 formula (A)

method 2 formula (B)

Graph (2.5)

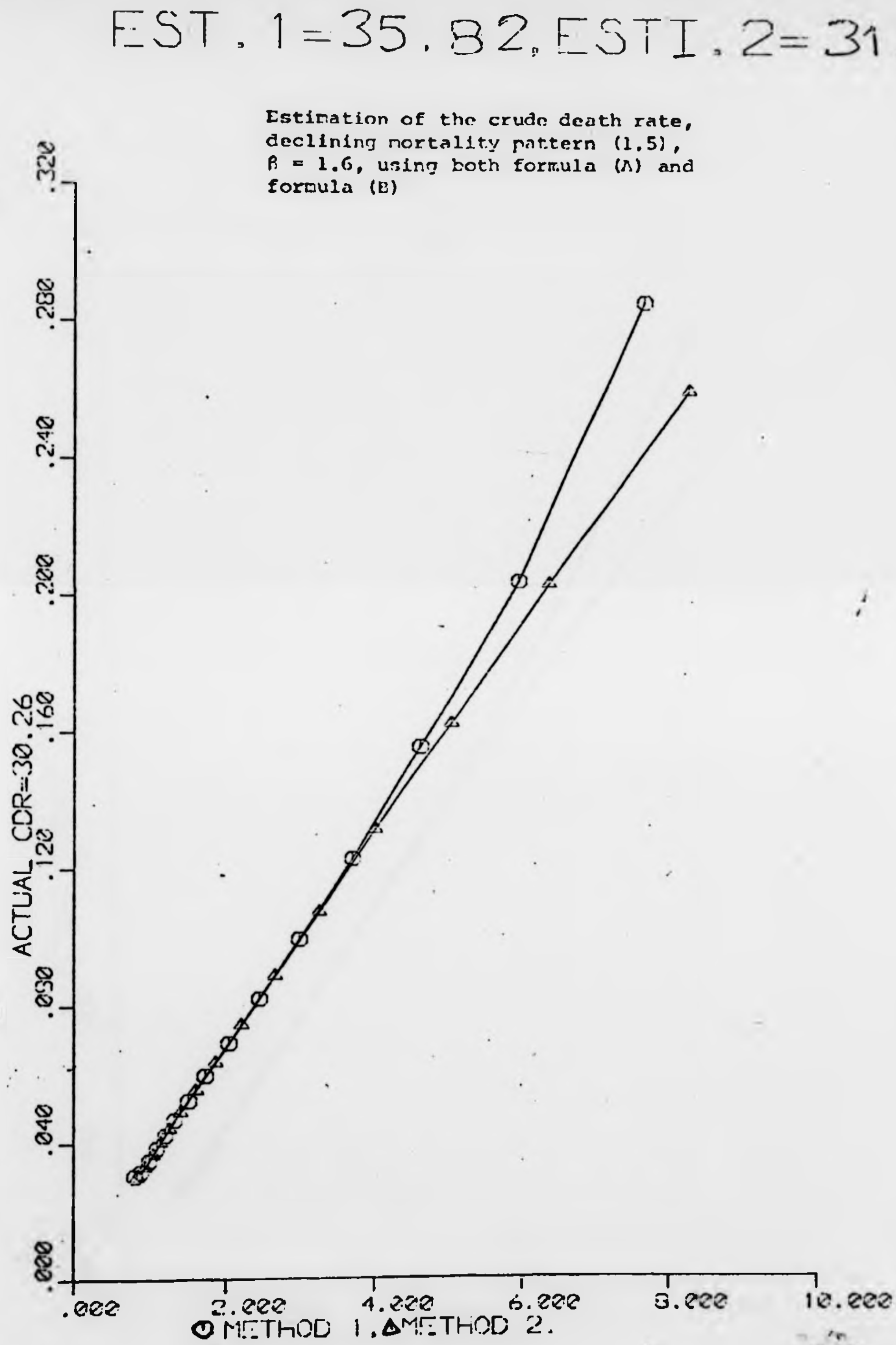


Estimation of the crude death rate, declining mortality pattern (1.6), β = 1.6, using both formula (A) and formula (B)

method 1 formula (A)

method 2 formula (B)

Graph (2.6)



Estimation of the crude death rate, declining mortality pattern (1,5), β = 1.6, using both formula (A) and formula (B)

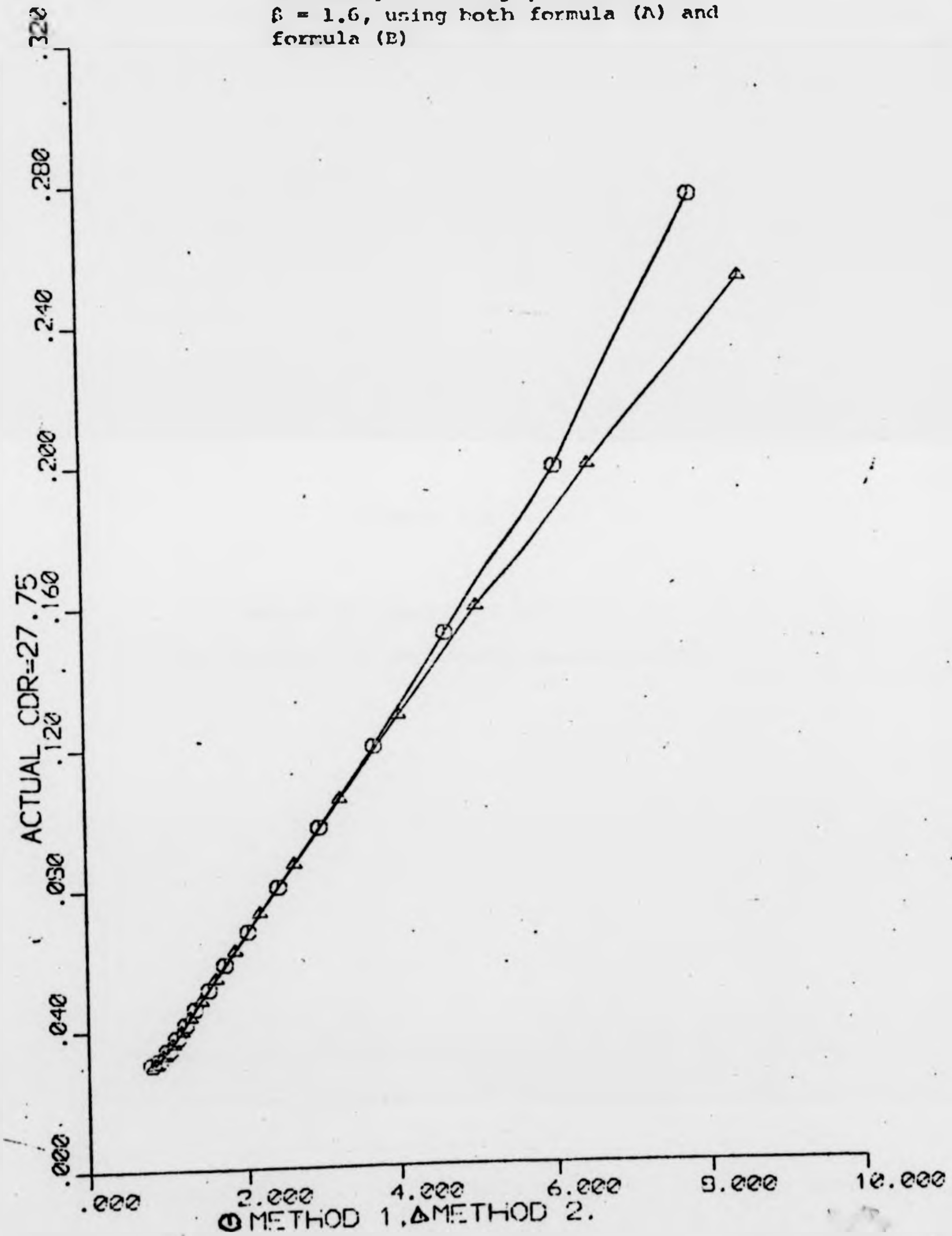
method 1 formula (A)

method 2 formula (B)

Graph (2.7)

EST. 1=33.65, ESTI. 2=29.23

Estimation of the crude death rate, declining mortality pattern (1.4), β = 1.6, using both formula (A) and formula (B)



method 1 formula (A)

method 2 formula (B)

CHAPTER III

EFFECT OF CHANGES IN MORTALITY AND FERTILITY ON THE GROWTH BALANCE METHOD

3.1 INTRODUCTION

It is our purpose to study the effect of changes in mortality and fertility on the growth balance estimate for the crude death rate. Several studies concerned with the effect of mortality changes on the age distribution are already available, but these studies are either devoted to the effect of changing mortality schedules on different stable age distribution or restricted to a special pattern of mortality change. Our concern is of a more practical nature, which is to assess the effect of changes in mortality and fertility - experienced by developing countries - on the applicability of the growth balance method for estimation. We should point out that we are not interested in the effect of mortality and fertility changes on the age structure, because this is merely a sufficient not necessary condition. In other words, cases may arise - as will be illustrated - when the age distribution deviates from a stable model but still the growth balance method is applicable.

Two approaches are used in this chapter. The first approach depends on standard component methods of population projection in illustrating the effect of changes in mortality and fertility, of the nature found in developing countries, on the applicability of the growth balance method. A general pattern of mortality and fertility change does not exist except in very broad terms, because within the developing countries the degree and rapidity of the change has been quite uneven; also the availability and quality of data makes any attempt to find accurate patterns almost impossible. Much of the discussion - under the first approach - is directed to finding several patterns of mortality and fertility change, which as a whole embody a range of feasible trend patterns for developing countries. The second approach is of a more theoretical nature; it is an attempt to analyse the relation between the growth balance estimate and certain features of mortality and fertility change. It explains and justifies the results reached in the first approach and also proposes possible modifications on the growth balance method.

3.2 THE FIRST APPROACH

3.2.1 Review of the Available Information on Changes in Mortality and Fertility in Developing Countries

(A) Mortality Change

The most trustworthy feature of mortality change in developing countries is the rapidity of its decline. The speed of this decline is unprecedented and has not been matched in the now advanced countries. Davis (1956b) illustrates this fact as follows: 'analyzing the data for fifteen under-developed countries we find that the crude death rate dropped by 53% during the thirty years from 1920-24 to 1950-54. The diminuation has been accelerating ... over a five year period (from the average for 1945-49 to the average for 1950-54), the death rate in eighteen underdeveloped countries declined by 20%. The history of the northwest European countries back to the eighteenth century shows no thirty year period in which a consistent decline occurred between half-decade averages, or in which the total thirty year change was anywhere near the 53% record of our backward areas.' The reason for using the crude death rate as the measure of mortality decline, was justified by its being the only index of mortality levels readily available for a number of backward countries and that it is unlikely the age structure of the populations considered has changed much due to the short period studied and the fact that fertility remained constant.

Stolnitz (1965) - using the available information from life tables on changes over 10 to 20 year periods ending in the 1950's - showed that the average increases in expectation of life at birth exceeds 0.5 years per annum and more often than not they are closer or above one year per annum. To compare this magnitude of change to the one experienced by developed countries; the same study indicated that before 1900 the average rise in life expectancy in developed countries amounted annually to about two-tenths of a year and that the increases since 1900 - although larger - have been only four-tenths. The highest recorded short term increase in life expectancy between 1840 and 1940 was 0.63 years, in the Netherlands between 1915 and 1926.

It should be noted that though the decline in mortality has occurred nearly everywhere in non-industrial areas, the degree and rapidity of this decline has been quite uneven; thus in Asia, Africa and Latin America there is at present great disparity in mortality levels in developing countries. The comparison within the underdeveloped areas is faced with the problem of inaccurate statistics, it may generally be stated that the level of mortality in Latin American countries is closer to the level in the underdeveloped countries of Asia than it is to that in countries of Africa. Most of the African countries still have a very high mortality and as pointed out by Arriaga Davis (1969) 'in the 1960's 13 out of 20 countries with information still had a life expectancy no higher than 40'.

The age pattern of mortality change for developing countries can only be documented with great caution. This is due to the limited and sketchy nature of the data available; thus it cannot be claimed that the few data we can rely on are representative. Also, the quality of the information available may distort the changes by age and one may doubt if some of the age pattern characteristics are a true feature of mortality or the effect of these errors.

The main findings of the available attempts to analyse the age pattern of mortality change may be summarized as follows: in a United Nations study (United Nations, 1962) using data for developing countries it was stated 'The data for these countries show a greater diversity of age patterns of recent mortality reductions than seems to have been characteristic of the earlier reductions in countries of Europe, Northern America, Australia and New Zealand. Also it seems that in these under-developed countries, adults up to middle age have shared with children in the benefit of the recent mortality reductions, to a more nearly equal extent than they previously did in the countries which led in the movement of declining death rates.' Illustrations showing these diversity of patterns of mortality declines are given in Table (3.1).

Table (3.1) Percentage annual decline in age specific death rates and the annual increase in e_o (males)

Country / Period / Age	Chile[1] 1920-1959	Taiwan[1] 1920-1960	Mauritius[2] 1944-1952
0-	1.396%	2.167%	3.700%
1-	2.034	2.175	5.975
5-	2.134	2.260	6.795
10-	1.820	2.160	8.137
15-	1.941	2.143	9.250
20-	1.995	2.078	9.337
25-	1.747	2.213	9.200
30-	1.631	2.211	8.912
35-	1.621	2.188	8.212
40-	1.477	2.098	7.662
45-	1.220	1.984	6.812
50-	1.173	1.804	6.075
55-	.547	1.632	5.837
60-	.845	1.499	5.825
65-	.725	1.297	4.825
70-	.653	.845	2.850
75-	.440	.779	2.675
80-	.849	.950	1.862
e_o	.607	.89	

[1]source: Preston & Keyfitz & Schoen (1972)

[2]source: United Nations (1962)

Stolnitz (1965) indicated that in a general way 'absolute declines in age specific probabilities of dying (q_x) or rises in their complements, the probabilities of surviving resemble reversed J's or even U's, ..., a related result, since age specific survivorship falls in the age beyond the childhood years, is that the corresponding percentage changes have also been reversed J's or U's'.

To summarize the previous review, we may stress that to imitate the pattern of decline in developing countries - using population projection - one should be careful to reproduce a rapid decline in mortality and to allow for several speeds and age patterns of decline. Some of the age patterns of change in survivorship ratios should resemble a reversed J or U's.

(B) Fertility Change

Fertility is almost universally high in developing countries. Kirk (Behrman & Corsa & Freedman 1969) indicated that 'There is a sharp dichotomy between the natality of the "developed" and the less developed countries. No "less developed" country of major consequence yet has a birth rate under 30. No developed country has a birth rate over 25, and few have a birth rate over 20.'

Detailed study for the historical pattern of fertility change are available for many developing countries. In contrast to the common picture of mortality decline, fertility change seems to have taken different forms. In some cases, such as the carribean islands (Jamaica, Trinidad and Guiana), there seem to be an upward movement in fertility up till the 1960's. In others, including several countries of Latin America in the 1940's to the 1960's (El Salvador, Guatemala, Mexico, Brazil ...), the picture is more of a constant fertility. As recently as 1960 Taiwan, Singapore and Puerto Rico were almost the only countries with a significant declining trend of fertility; the list has been growing since and is including countries such as

Ceylon, Korea and Chile. More and more countries are joining in the decline of fertility, but the general downward trend has not been firmly established yet.

3.2.2 Patterns and Data used

(A) Mortality Patterns

Pattern 1

The ideal situation - in such a study - is to find actual representative patterns of mortality decline, which will provide a direct check on the applicability of the growth balance method. The lack of actual historical data necessitates the use of a model system. A suitable model needs to incorporate the basic characteristics found in the age structure of mortality rates at a certain point in time besides being flexible enough to show the general tendencies revealed in the path of change.

The most appropriate model system for our purpose seems to be the logit model because its standard life table represents the basic characteristics common in the mortality schedules of developing countries, also the logit model allows - through the use of two parameters - more freedom in choosing the required path of change among the multitudes of paths available and finally the model is readily suited for application on computer.

The logit model describes the relation between mortalities in different countries - or at different periods in the same country - using a mathematical function as:

$$\text{logit } l_x = \alpha + \beta \text{ logit } ls_x$$

where $\text{logit } l_x = \frac{1}{2} \log \frac{1-l_x}{l_x} = -\frac{1}{2} \log \frac{l_x}{1-l_x} = -\text{logit } (1-l_x)$, and l_x denotes the survivors at age x in the life table. ls_x denotes the survivors at age x in an arbitrary life table chosen as a basic standard pattern. The standard life table used in all the following applications is based on Brass general standard life table (Brass 1971), which is an average representation for published life tables of moderate and high mortality. α and β are constants which vary among life tables; their implications are not fully defined but in general α may be regarded as a measure of the level of mortality while β describes the age pattern of mortality. The limits of α and β, that cover the range observed for recorded mortality schedules of developing countries, is within $-1.0 \rightarrow 0.5$ for α and $0.6 \rightarrow 1.6$ for β.

The decline in mortality may be achieved through changing one or both of the parameters as follows:

a - changing α and fixing β.

b - changing α and β simultaneously.

The option of changing β and fixing α was disregarded beforehand, since a change in β only accomplishes small changes in e_o. For example:

$\alpha = 0.5$	$\beta = 0.6$	$e_o = 24.7$
$\alpha = 0.5$	$\beta = 1.6$	$e_o = 31.0$
$\alpha = 0$	$\beta = 0.6$	$e_o = 43.1$
$\alpha = 0$	$\beta = 1.6$	$e_o = 45.0$

Pattern (1.a):

In this option α is decreasing and β is kept constant. The reason for using this pattern of decline is that the age pattern of change in q_x, due to a change in α, resembles some of the available information which suggest that the absolute decline in ${}_nq_x$ (${}_nq_x$ denotes probability of dying between x and x+n) follow a U shape. Table (3.2) shows the decline in ${}_nq_x$ when β is constant and α declining.

Table (3.2)* The decline in ${}_nq_x$% from α = 0.5 to α = 0.9 and β constant

β / age	0.6	1.0	1.6
0	22.88	17.43	8.63
1	6.61	8.99	8.98
5	1.47	2.21	2.63
10	1.05	1.62	2.02
15	1.65	2.59	3.38
20	2.03	3.24	4.47
25	1.91	3.09	4.50
30	1.83	3.00	4.52
35	1.88	3.11	4.79
40	2.02	3.31	5.15
45	2.29	3.69	5.65
50	2.66	4.17	6.07
55	3.03	4.53	6.01
60	3.57	4.85	5.40
65	3.91	4.58	3.90
70	4.24	3.87	2.18

*source: values of q_x were given in Brass (1971), table 5.

Pattern (1.b):

In this option, where α and β are changing, the relation between childhood and adult mortality changes while mortality is declining. This kind of decline would affect the age composition considerably and allow us to test the death distribution method under the worst circumstances.

To avoid unnecessary complications - due to the several combinations of change in α and β that may be tried - we restricted our attention to a constant decrease in β while α has been changed indirectly through fixing l_1.

Data used in the projection:

For each case we started with a stable age distribution resulting from a constant schedule of mortality and a certain growth rate. The schedule of mortality used corresponds to the initial values of α and β and the standard life table for each sex separately. The stable age distribution was subjected to constant fertility and to the hypothetical mortality decline; the fertility schedule used was based on model fertility, pattern 6 of the United Nations (United Nations 1963) corresponding to the crude birth rate of the stable age distribution.

Pattern (1.a):

Different constant values of β were used, ranging from $\beta = 0.7$ to $\beta = 1.3$; for each constant value of β, α was changed to achieve a specified rate of increase in expectation of life at birth. Two rates of increase in e_o were applied; a rate of annual increase of 0.5 years and a faster annual rate of 1.0 years. The initial value of α was 0.7. The projection stopped when e_o exceeded 70.0 or after 20 periods equivalent to 100 years.

Pattern (1.b):

The initial value of $\beta = 1.6$. l_1 was fixed to the value of ls_1. Thus the implied initial value of α for males = $(1-\beta)$ logit $ls_1 = 0.486$. β was decreased from 1.6 to 0.6 in steps of 0.1 and a slower step of 0.05. The projection stops when $\beta = 0.6$.

Pattern (2):

In constructing this pattern, and the following one, we depended heavily on two studies: (Keyfitz & Flieger 1968) and (Preston & Keyfitz & Schoen 1972), in which detailed information on mortality and fertility for several countries are given. Two countries were picked and their mortality decline patterns calculated. These countries are: Taiwan and Chile. The reasons for choosing them are that their data goes further back than other published data on developing countries and that we are not only interested in a common pattern (which has no significant effect on the age structure) but rather on some extreme patterns (extreme in a sense of a very big and/or fast decline in mortality).

Data used in the projection:

For each country we started with a stable age distribution, resulting from a constant schedule of mortality - corresponding to the actual death probabilities given by (Preston & Keyfitz & Schoen 1972) for the earliest period - and a certain growth rate as a measure of probable value which may have applied in the base year of projection. The stable age distribution is subjected to constant fertility and to a mortality change illustrating the actual pattern of decline in each country; the fertility schedule used is based on model fertility pattern 6 of the United Nations (United Nations 1963) corresponding to the crude birth rate of the stable age distribution. The mortality decline is based on the actual percentage increase in the male survivorship ratios; this increase is calculated for each two consecutive published data and it is assumed to be uniformly distributed within the given period.

a) Chile

Starting with the published probability (q_x) corresponding to the base year 1909 and a hypothetical growth rate .009, Chile is projected for 50 years, assuming the decline in mortality to follow the annual increase in male survivorship ratios.

b) Taiwan

Similarly, Taiwan is projected for 40 years starting with the base year 1920 and a growth rate = .009.

Pattern (3):

Though our concern is mainly with mortality decline of under-developed countries, this third pattern deals with European data. The justification for this is the existing possibility that some under-developed countries did follow a European pattern of decline.

The data of three countries are used; mainly Sweden, Portugal and Italy. Sweden represent a special case due to the availability of fairly reliable data on mortality from the 18th century upwards; so the Sweden trend of mortality has been accelerated to agree more with the fast reduction that has been observed in under-developed countries. Thus two cases are applied here:

a) The actual Swedish, Italian and Portugese pattern;

b) the actual Swedish pattern from 1783 till 1863 (corresponding to e_o = 43.347) is applied, but starting at 1863 the speed of mortality decline is doubled. Thus it is assumed that a gain in life expectancy equal to 24.282 years is achieved in 40 years rather than 80 years. This is done by considering each 10 year decline to happen in 5 years only.

The same procedure - as in the previous pattern - has been followed here starting with the base year 1783 and a growth rate of .009 for Sweden, the base year for Italy is 1881 and the growth rate used is .005 and finally Portugal is projected for 40 years starting 1920 and with a growth rate .009.

(B) Fertility Patterns:

For each of the countries considered; the annual pattern of decline in age specific fertility rates is calculated using actual data and starting from the earliest published data on age specific fertility rates. The female age distribution is subjected to mortality and fertility decline. Mortality decline started from the same base year as in the previous patterns; fertility is held constant till period 10 for Italy, period 7 for Chile, period 7 for Taiwan, period 3 for Portugal and the first period for Sweden. Each period is equivalent to 5 years of mortality decline.

3.2.3 Results

(A) Mortality decline and fertility constant

Pattern (1):

Considering the results of pattern (1.a) - given in tables (3.3) and (3.4) - we note that generally the estimate is close to the actual death rate.

Comparing the actual death rate with the best estimate whether estimate (A) or estimate (B) - where estimate (A) is calculated using formula (A) and least square fit, and estimate (B) is calculated using formula (B) and least square fit - we get: for $\beta > 1$, whether the annual increase in e_0 is .5 or 1.0, the maximum deviation does not exceed 1%. For $\beta < 1$ only in the slow decline the deviation exceeds 1%; when $\beta = 0.7$ the maximum deviation is 2.62% and this deviation only occurs once corresponding to period 12; for other values of β the maximum deviation decrease reaching 2.21, 1.78 and 1.16 for $\beta = 0.8$, 0.9 and 1.0.

Considering the results of pattern (1.b) - given in table (3.5) - and comparing the actual death rate and estimated (A), there is a considerable amount of deviation in the neighbourhood of 5%. Since this deviation starts from the first period when the population is stable and mortality decline has not started yet, one is quite suspicious that the source of deviation is due to the method of estimation rather than the effect of mortality decline. The second estimate does confirm this point, so we will assume that the difference between the actual death rate and estimated (B) is due to mortality decline while the difference between estimated (A) and estimated (B) is the effect of the method of estimation.

There is a difference between the actual death rate and the estimated (B).

Table (3.3) Fixed β, changing α.
The actual and estimated death rate, using formula (A) and (B), corresponding to annual increase in e_o = 0.5.
Pattern (1.a)

Beta period[1]	0.7			0.8			0.9			1.0		
	actual CDR	Est. (A)	Est. (B)	actual CDR	Est. (A)	Est. (B)	actual CDR	Est. (A)	Est. (B)	actual CDR	Est. (A)	Est. (B)
1	54.76	54.95	54.55	53.01	53.70	52.82	51.25	52.48	51.07	49.51	51.35	49.33
2	50.78	50.93	50.62	49.10	49.66	48.90	47.41	48.45	47.19	45.72	47.29	45.48
3	46.87	46.81	46.60	45.18	45.52	44.89	43.55	44.34	43.23	41.89	43.17	41.54
4	42.85	42.53	42.44	41.29	41.32	40.82	39.72	40.19	39.24	38.16	39.11	37.66
5	39.15	38.41	38.44	37.63	37.31	36.94	36.12	36.26	35.46	34.64	35.25	33.99
6	35.89	34.73	34.89	34.40	33.69	33.46	32.93	32.71	32.06	31.50	31.78	30.68
7	33.00	31.37	31.66	31.51	30.39	30.30	30.03	29.44	28.93	28.65	28.57	27.64
8	30.25	28.14	28.56	28.75	27.20	27.24	27.30	26.33	25.97	25.94	25.52	24.75
9	27.57	25.00	25.53	26.10	24.17	24.31	24.69	23.38	23.14	23.34	22.62	21.99
10	24.95	22.04	22.62	23.50	21.28	21.49	22.17	20.06	20.45	20.89	19.94	19.43
11	22.45	19.31	19.92	21.08	18.68	18.93	19.80	18.08	17.99	18.61	17.51	17.09
12	20.11	16.90	17.49	18.82	16.35	16.61	17.60	15.82	15.77	16.49	15.33	14.99
13	17.93	14.82	15.34	16.71	14.33	14.56	15.58	13.87	13.82	14.54	13.42	13.11
14	15.91	13.13	13.48	14.76	12.68	12.78	13.70	12.23	12.11	12.74	11.82	11.49
15	14.00	11.93	11.94	12.94	11.46	11.30	11.97	11.02	10.69	11.10	10.60	10.13
16	12.24	11.23	10.85	11.28	10.70	10.22	10.40	10.19	9.62	9.61	9.71	9.06
17	10.62	10.10	10.04	9.75	9.54	9.35	8.96	9.00	8.71	8.24	8.51	8.13
18	9.16	8.42	8.41	8.38	7.96	7.85	7.66	7.53	7.33	7.02	7.12	6.85
19	7.79	7.07	7.05	7.08	6.67	6.56	6.44	6.29	6.12	5.88	5.94	5.71
20	6.49	5.95	5.88	5.88	5.59	5.47	5.33	5.26	5.09	4.84	4.94	4.74
e_o	68.62			68.94			69.35			69.83		

[1] each period denotes a 5 year interval.

Table (3.3) (continued)

Beta period	1.1			1.2			1.3		
	actual CDR	Est. (A)	Est. (B)	actual CDR	Est. (A)	Est. (B)	actual CDR	Est. (A)	Est. (B)
1	47.80	50.29	47.63	46.14	49.33	45.98	44.55	48.46	44.39
2	44.03	46.18	43.77	42.45	45.22	42.17	40.91	44.32	40.62
3	40.32	42.13	39.95	38.80	41.16	38.39	37.31	40.26	36.88
4	36.67	38.13	36.16	35.22	37.21	34.69	33.82	36.35	33.28
5	33.20	34.31	32.55	31.84	33.47	31.20	30.53	32.69	29.90
6	30.13	30.91	29.35	28.81	30.10	28.06	27.58	29.40	26.87
7	27.30	27.74	26.38	26.04	27.02	25.20	24.83	26.33	24.06
8	24.65	24.78	23.60	23.42	24.08	22.48	22.27	23.46	21.44
9	22.11	21.96	20.94	20.95	21.34	19.94	19.87	20.79	19.00
10	19.71	19.34	18.47	18.64	18.81	17.59	17.63	18.33	16.75
11	17.50	16.98	16.24	16.49	16.51	15.45	15.55	16.07	14.70
12	15.46	14.87	14.23	14.52	14.44	13.53	13.66	14.07	12.88
13	13.59	13.01	12.46	12.72	12.62	11.83	11.94	12.28	11.26
14	11.87	11.43	10.90	11.09	11.07	10.35	10.38	10.75	9.84
15	10.32	10.22	9.60	9.61	9.86	9.10	8.97	9.54	8.63
16	8.89	9.28	8.54	8.26	8.90	8.07	7.70	8.55	7.63
17	7.62	8.08	7.61	7.05	7.69	7.13	6.56	7.36	6.72
18	6.45	6.75	6.41	5.96	6.43	6.02	5.54	6.15	5.68
19	5.39	5.63	5.36	4.96	5.35	5.04	4.60	5.11	4.76
20	4.42	4.66	4.44	4.06	4.43	4.19	3.76	4.22	3.96
e_o	70.36			70.93			71.56		

Table (3.4) Fixed β, changing α.
The actual and estimated death rate, using formula (A) and (B), corresponding to annual increase in e_o = 1.0.
Pattern (1.a)

Beta period[1]	0.7			0.8			0.9			1.0		
	actual CDR	Est. (A)	Est. (B)	actual CDR	Est. (A)	Est. (B)	actual CDR	Est. (A)	Est. (B)	actual CDR	Est. (A)	Est. (B)
1	54.76	54.95	54.55	53.01	53.70	52.82	51.25	52.48	51.07	49.51	51.35	49.33
2	47.31	47.44	47.20	45.58	46.07	45.42	43.91	44.82	43.69	42.21	43.59	41.94
3	40.17	40.05	39.96	38.57	38.77	38.30	36.98	37.55	36.65	35.41	36.38	35.03
4	33.55	33.12	33.14	32.07	31.98	31.66	30.60	30.87	30.17	29.21	29.85	28.75
5	27.81	27.08	27.21	26.45	26.10	25.91	25.12	25.15	24.62	23.82	24.23	23.35
6	23.05	22.06	22.29	21.76	21.20	21.13	20.52	20.37	20.00	19.35	19.59	18.90
7	18.94	17.73	18.06	17.71	16.98	17.03	16.56	16.27	16.04	15.47	15.58	15.09
8	15.20	13.85	14.24	14.07	13.23	13.37	13.01	12.64	12.54	12.04	12.08	11.75
9	11.78	10.42	10.82	10.79	9.94	10.13	9.87	9.48	9.47	9.03	9.04	8.85
10	8.72	7.48	7.86	7.89	7.14	7.34	7.14	6.81	6.86	6.45	6.48	6.39
11	6.05	5.06	5.37	5.40	4.83	5.01	4.81	4.60	4.66	4.28	4.36	4.34
e_o	71.43			71.78			72.21			72.71		

[1] each period denotes a 5 year interval.

Table (3.4) (continued)

Beta period	1.1			1.2			1.3		
	actual CDR	Est. (A)	Est. (B)	actual CDR	Est. (A)	Est. (B)	actual CDR	Est. (A)	Est. (B)
1	47.80	50.29	47.63	46.14	49.33	45.98	44.55	48.46	44.39
2	40.62	42.50	40.30	39.05	41.48	38.69	37.54	40.52	37.14
3	33.89	35.29	33.46	32.46	34.31	31.98	31.09	33.40	30.56
4	27.84	28.88	27.36	26.56	28.01	26.06	25.32	27.18	24.79
5	22.62	23.41	22.16	21.46	22.64	21.02	20.38	21.94	19.95
6	18.22	18.85	17.85	17.19	18.19	16.87	16.19	17.56	15.93
7	14.46	14.96	14.19	13.51	14.37	13.34	12.64	13.84	12.55
8	11.15	11.56	11.01	10.33	11.07	10.32	9.56	10.62	9.66
9	8.28	8.63	8.27	7.58	8.23	7.72	6.96	7.88	7.22
10	5.84	6.16	5.96	5.29	5.87	5.56	4.89	5.59	5.20
11	3.82	4.14	4.04	3.40	3.93	3.77	3.04	3.73	3.52
e_0	73.27			73.87			74.54		

This difference increases as mortality decreases, but the magnitude of difference is not very disturbing (bearing in mind that we expect this sort of decline to produce the biggest change on age-dist.). For example, while expectation of life at birth for males increased from 30.79 years to 51.33 years the estimate deviates only 3.06%.

The maximum deviation in the fast decline reaches 3.25% while it is only 1.68% in the slow decline.

Table (3.5) Changing α and β.
The actual and estimated death rates using formula (A) and (B) corresponding to fast and slow decline in β.
Pattern (1.b)

Beta	fast decline in β			slow decline in β		
	actual death rate	estimated (A)	estimated (B)	actual death rate	estimated (A)	estimated (B)
1.6	32.95	37.90	32.83	32.95	37.90	32.83
1.55				31.60	36.85	31.92
1.5	30.26	35.82	31.02	30.31	35.71	30.98
1.45				29.08	34.49	30.04
1.4	27.75	33.65	29.23	27.90	33.21	29.06
1.35				26.76	31.90	28.11
1.3	25.39	31.42	27.47	25.66	30.57	27.13
1.25				24.59	29.24	26.16
1.2	23.17	29.15	25.72	23.56	27.93	25.19
1.15				22.55	26.63	24.21
1.1	21.08	26.88	23.99	21.56	25.34	23.24
1.05				20.60	24.05	22.26
1.	19.12	24.66	22.27	19.67	22.78	21.27
.95				18.76	21.55	20.28
..9	17.29	22.50	20.54	17.88	20.40	19.32
.85				17.03	19.35	18.42
.8	15.59	20.42	18.82	16.21	18.34	17.56
.75				15.42	17.34	16.68
.7	14.01	18.39	17.07	14.67	16.37	15.82

Pattern (2):

Generally, the estimated death rate agrees quite well with the actual death rate; the maximum deviation for Taiwan does not exceed 97% (the deviation is calculated with respect to the best estimate); for Chile also, the estimate is quite good except for the first 15 years of mortality decline where the deviation reaches 4.36%.

Table (3.6) The actual and estimated death rates, using formula (A) and (B) corresponding to actual patterns of mortality decline. Pattern (2)

CDR period[1]	a) Chile			b) Taiwan		
	actual death rate	estimated (A)	estimated (B)	actual death rate	estimated (A)	estimated (B)
(0)	34.97	35.48	34.64	37.44	39.90	37.62
(1)	34.79	33.04	32.50	33.61	36.66	33.73
(2)	34.35	29.99	29.66	25.96	29.66	26.93
(3)	31.96	29.34	28.61	21.66	24.85	22.54
(4)	27.62	28.90	27.71	20.72	23.07	21.05
(5)	24.67	26.41	25.40	18.77	20.75	18.98
(6)	23.01	23.45	22.81	15.81	17.56	16.10
(7)	20.63	20.49	20.00	12.85	14.26	13.08
(8)	17.50	17.12	16.65	9.86	10.94	10.03
(9)	14.85	14.08	13.55			
(10)	12.70	11.86	11.34			

[1]each period denotes a 5 year interval.

Table (3.7) The actual and estimated death rates, using formula (A) and (B), corresponding to actual patterns of mortality decline. Pattern (3)

CDR period[1]	Italy			Portugal			Sweden (actual pattern)			Sweden (accelerated pattern)		
	actual	estimated (A)	estimated (B)	actual	estimated (A)	estimated (B)	actual	estimated (A)	estimated (B)	actual	estimated (A)	estimated (B)
(0)	30.17	30.20	30.00	28.87	29.13	28.56	29.36	30.15	29.04			
(1)	28.67	28.60	28.33	26.92	27.53	26.95	28.47	29.20	28.01			
(2)	25.59	25.53	25.15	22.63	23.95	23.43	26.69	27.51	26.30			
(3)	23.06	23.17	22.76	19.84	20.54	20.16	25.76	26.59	25.59			
(4)	21.17	21.33	21.02	18.50	18.28	18.00	25.66	26.10	25.34			
(5)	19.49	19.49	19.25	16.77	16.50	16.27	25.60	25.55	24.79			
(6)	18.06	18.02	17.75	14.74	14.67	14.51	25.66	25.47	24.63			
(7)	16.64	16.70	16.50	12.95	12.62	12.54	24.78	24.87	23.94			
(8)	15.21	15.14	15.11	11.34	10.68	10.66	22.86	23.32	22.40		the same	
(9)	13.35	12.73	12.82				22.09	22.43	21.63			
(10)	10.83	9.79	9.94				22.57	22.39	21.71			
(11)	8.75	7.88	8.01				22.65	21.93	21.34			
(12)	7.42	6.84	6.94				22.35	21.21	20.66			
(13)	6.01	5.53	5.56				22.57	21.55	20.81			
(14)	4.40	4.02	4.01				23.29	22.99	21.89			
(15)	3.08	2.81	2.79				22.72	23.37	22.11			
(16)	2.10	1.89	1.89				20.86	22.14	20.98			
(17)							19.86	21.27	20.19	19.76	21.35	20.26
(18)							19.74	21.15	20.12	18.13	20.07	18.99
(19)							18.99	20.58	19.49	15.69	17.82	16.78
(20)							17.60	19.59	18.39	14.05	16.03	15.09
(21)							16.43	18.67	17.56	12.66	14.18	13.43
(22)							15.47	17.27	16.60	10.90	11.81	11.41
(23)							14.61	15.31	15.11	9.34	9.44	9.37
(24)							13.84	13.60	13.51	7.90	7.46	7.46
(25)							13.02	12.49	12.31			
(26)							12.19	11.72	11.51			
(27)							11.18	10.83	10.65			
(28)							10.03	9.72	9.56			
(29)							9.16	8.76	8.64			
(30)							8.57	7.98	7.92			
(31)							7.76	7.13	7.09			
(32)							6.70	6.23	6.16			

[1] each period denotes a 5 year interval

Pattern (3):

For all three countries the agreement between the estimate and actual death rate is very good, for example the maximum deviation for Italy does not exceed .89% and this only occurs in one period (period 10), also for Portugal the maximum deviation is .80% For Sweden, whether the actual or accelerated pattern, the maximum deviation is 1.5%

(E) Mortality and Fertility Decline

The results of the effect of mortality and fertility decline on the estimate of the death rate are given in Table (3.8) and (3.9). We note that the pattern of fertility decline used hardly affect the estimate.

Table (3.8) The actual and estimated death rates, using formula (A) and (B), corresponding to actual patterns of mortality and fertility decline.

country	a) Chile			b) Taiwan		
CDR period[1]	actual CDR	estimated (A)	estimated (B)	actual CDR	estimated (A)	estimated (B)
(0)	31.35	31.53	31.05	34.26	34.85	33.91
(1)	31.24	29.30	29.12	30.30	31.59	30.41
(2)	30.87	26.51	26.53	22.37	24.58	23.22
(3)	28.46	25.72	25.39	18.02	19.85	18.70
(4)	24.17	24.99	24.19	17.24	18.34	17.42
(5)	21.26	22.58	21.91	15.38	16.24	15.48
(6)	19.76	19.94	19.58	12.46	13.25	12.64
(7)[2]	17.51	17.21	16.96	9.50	10.11	9.63
(8)	14.17	13.76	13.47	6.54	7.01	6.64
(9)	11.98	11.29	10.95	4.62	5.09	4.82
(10)	10.10	9.33	9.04			
(11)	7.66	7.50	7.28			

[1]each period denotes a 5 year interval
[2]the beginning of fertility decline for Taiwan and Chile

Table (3.9) The actual and estimated death rates, using formula (A) and (B), corresponding to actual patterns of mortality and fertility decline.

Country	Italy			Portugal			Sweden (actual)			Sweden (accelerated)		
CDR period[1]	actual	estimated (A)	estimated (B)	actual	estimated (A)	estimated (B)	actual	estimated (A)	estimated (B)	actual	estimated (A)	estimated (B)
(0)[2]	29.38	29.39	29.20	25.22	25.09	24.95	26.64	27.06	26.36	26.64	27.06	26.36
(1)	27.87	27.80	27.54	23.20	23.54	23.36	25.91	26.25	25.47	25.91	26.25	25.47
(2)	24.79	24.75	24.37	19.09	20.11	19.88	23.83	24.25	23.44	23.83	24.25	23.44
(3)[3]	22.27	22.39	21.98	16.27	16.76	16.59	23.04	23.54	22.90	23.04	23.54	22.90
(4)	20.39	20.54	20.24	14.60	14.34	14.90	22.87	23.07	22.63	22.87	23.07	22.63
(5)	18.72	18.73	18.48	13.25	13.09	12.96	22.88	22.65	22.19	22.88	22.65	22.19
(6)	17.29	17.28	17.01	11.64	11.82	11.70	23.04	22.69	22.16	23.04	22.69	22.16
(7)	15.88	15.94	15.74	10.21	10.24	10.17	22.35	22.28	21.67	22.35	22.28	21.67
(8)	14.46	14.39	14.34	8.90	8.68	8.65	20.67	20.96	20.34	20.67	20.96	20.34
(9)	12.59	12.04	12.11				19.89	20.06	19.55	19.89	20.06	19.55
(10)	10.07	9.16	9.28				20.21	19.87	19.48	20.21	19.87	19.48
(11)[4]	7.91	7.16	7.26				20.09	19.27	18.94	20.09	19.27	18.94
(12)	6.30	5.86	5.91				19.59	18.42	18.11	19.59	18.42	18.11
(13)	5.64	5.29	5.28				19.74	18.71	18.24	19.74	18.71	18.24
(14)	4.49	4.25	4.19				20.47	20.10	19.32	20.47	20.10	19.32
(15)	3.51	3.36	3.30				20.02	20.56	19.60	20.02	20.56	19.60
(16)	2.75	2.62	2.59				18.33	19.04	18.57	18.33	19.04	18.57
(17)							17.36	18.14	17.26	17.25	18.18	17.29
(18)							17.17	18.17	17.62	15.53	16.98	16.37
(19)							16.40	17.58	16.66	13.11	14.72	13.86
(20)							15.00	17.03	15.89	11.57	13.50	12.61
(21)							13.93	16.54	15.41	10.38	12.18	11.40
(22)							13.13	15.49	14.77	8.80	10.11	9.67
(23)							12.43	13.44	13.45	7.43	7.75	7.80
(24)							11.81	11.35	11.59	6.12	5.66	5.82
(25)							10.98	9.93	10.09			
(26)							10.09	9.11	9.15			
(27)							9.11	8.54	8.40			
(28)							8.07	7.96	7.76			
(29)							7.48	7.57	7.33			
(30)							7.26	7.40	7.19			
(31)							6.90	6.99	6.88			

[1]each period denotes a 5 year interval
[2]the beginning of fertility decline for Sweden
[3]the beginning of fertility decline for Portugal
[4]the beginning of fertility decline for Italy

3.3 THE SECOND APPROACH

The previous approach illustrates the magnitude of error likely to affect the growth balance estimate of the crude death rate when applied to data of developing countries. Generally, it may be stated that the estimate, except in few cases, is not significantly affected by the pattern of mortality and fertility change that prevailed in developing countries.

In this approach, an attempt to explain this apparent robustness in the growth balance estimate and a justification for the few cases when a deviation appeared is presented. Also, possible modifications of the growth balance method are proposed.

3.3.1 Justification for the Effect of Mortality Decline

The important role of the age pattern of mortality change in shaping the age distribution has been sufficiently stressed in demographic literature; though most of this literature was devoted to studying the effect of changing mortality schedules on different stable age distributions.

A general rule - which applies whether we are discussing actual or stable age distributions - is that an equal difference between two mortality schedules (implying the same relative change in age specific survival rates) does not affect the proportionate age distribution. Of course, the absolute age distribution is increased by equal percentages at all ages corresponding to the lower mortality schedules.

This rule does clarify why in some cases big mortality differences only affect the age distribution slightly. For example, we note that in pattern (1.a) α is decreasing while β is kept constant; this means that though mortality is declining the relation between childhood and adult mortality is constant. In pattern (1.b), the relation between childhood and adult mortality changes while mortality declines. This implies that the relative change in age specific survival rates is nearly constant under pattern (1.a) as compared to pattern (1.b). This fact is illustrated in Table (3.10) which shows the relative change in age specific survival rates when α is changing from 0.5 to -0.5 and β is constant with value 0.6 and changing from 1.6 to 0.6.

Table (3.10)* % relative change in age specific survival rates from $\alpha = 0.5$ to $\alpha = -0.5$

x	$\beta = 0.6$ pattern (1.a)	β changing from 1.6 to 0.6 pattern (1.b)
5	2.653	4.110
10	1.898	3.233
15	2.958	5.791
20	3.883	8.354
25	3.708	9.027
30	3.601	9.774
35	3.794	11.[illegible]
40	4.17[illegible]	13.741
45	4.899	17.7[illegible]7
50	5.955	24.084
55	7.234	32.749
60	9.311	47.980
65	11.365	67.417
70	14.623	105.548
75	16.878	157.648

*source: using values of α_x given in Brass (1971), table 5.

Examining the results of pattern (1.a) and (1.b), we note that - as expected - bigger deviation in the estimate appear in the latter.

We now turn our attention to a detailed investigation of the factors age specific survival rates is nearly constant under pattern (1.a) as compared to pattern (1.b). This fact is illustrated in Table (3.10) which shows the relative change in age specific survival rates when α is changing from 0.5 to -0.5 and β is constant with value 0.6 and changing from 1.6 to 0.6.

Table (3.10)* % relative change in age specific survival rates from $\alpha = 0.5$ to $\alpha = -0.5$

x	$\beta = 0.6$ pattern (1.a)	β changing from 1.6 to 0.6 pattern (1.b)
5	2.653	4.110
10	1.898	3.233
15	2.958	5.791
20	3.883	8.354
25	3.708	9.027
30	3.601	9.774
35	3.794	11.[illegible]
40	4.17[illegible]	13.741
45	4.89[illegible]	17.7[illegible]7
50	5.955	24.084
55	7.234	32.749
60	9.311	47.980
65	11.365	67.417
70	14.623	105.548
75	16.878	157.648

*source: using values of α_x given in Brass (1971), table 5.

Examining the results of pattern (1.a) and (1.b), we note that - as expected - bigger deviation in the estimate appear in the latter.

We now turn our attention to a detailed investigation of the factors affecting the growth balance estimate. Preston showed that:

$$\frac{N_y}{P_y} = CDR\,\frac{D_y}{P_y} - \frac{\int_y^{\omega} N_x \, d\log \frac{N_x}{l_x}}{P_y} \qquad \text{(equation 2.4)}$$

where N_y, P_y, D_y, l_x and CDR are as defined.

For the growth balance estimate to be exact N_y, P_y and D_y should express the measures of the population experiencing the current mortality and fertility. In other words, if analytic expressions are available relating age composition to changing schedules of fertility and mortality just as the stable formula relates age composition to constant schedules, the growth balance method may be readily generalized to apply to cases when schedules of fertility and mortality has been changing rather then constant.

Let N_x^s denote the stable age distribution corresponding to the mortality and fertility schedules current in N_x, then:

$$\frac{N_y}{P_y} = CDR\,\frac{D_y}{P_y} - \frac{\int_y^{\omega} N_x \, d\log \frac{N_x}{l_x}\cdot\frac{N_x^s}{N_x^s}}{P_y} \qquad (3.1)$$

$$\frac{N_y}{P_y} = CDR\cdot\frac{D_y}{P_y} - \frac{\int_y^{\omega} N_x \, d\log \frac{N_x^s}{l_x}}{P_y} - \frac{\int_y^{\omega} N_x \, d\log \frac{N_x}{N_x^s}}{P_y} \qquad (3.2)$$

$$\frac{N_y}{P_y} = CDR\,\frac{D_y}{P_y} + r + z_y \qquad (3.3)$$

where $z_y = -\int_y^{\omega} N_x \, d\log \frac{N_x}{N_x^s} \,/\, P_y$

The deviation in the growth balance estimate results from ignoring this last term z_y. It is our purpose to discuss the magnitude of this term under different patterns of changing mortality.

Approximate analytic expressions for the age composition of a closed one sex population subject to a specific type of change as related to the stable age distribution with the current mortality and fertility ($\frac{N_x}{N_x^s}$) have been presented (Coale 1972). A summary of these expressions are given here:

'When mortality is declining after a long period of constant fertility and mortality (stability) and when the change in mortality is subject to the following postulates:

'(1) As mortality changes, a fixed age structure is assumed in the difference in age specific death rates from any moment to any other. Thus if $U(a,t_1)$ and $U(a,t_2)$ are age-specific mortality schedules at two moments when mortality is changing, $U(a,t_1) - U(a,t_2) = K\Delta U(a)$, where K is a constant and $\Delta U(a)$ is a non-altering characteristic age schedule of mortality change.

(2) The age pattern of change in mortality rates can be approximated by a steeply declining section from age 0 to about age 5, a section that can be considered level from age 5 to 45, and a section that rises linearly with age above 45.

(3) The time pattern of the mortality change initiated at $t = 0$ is one of linear change at each age. Thus $U(a,t) = U(a,0) - t\Delta U(a)$'

The age distribution of a population subject to this mortality decline relative to the stable age distribution is given by:

$$t < 3\frac{T}{4}$$

$$\frac{N_x(t)}{N_x^s(t)} = \frac{b_K}{b_s}\exp\left(-K(1-\frac{t}{T})x-1.27\frac{Kx}{T}\right) \qquad 5 < x < t$$

$$= \frac{b_K}{b_s}\exp\left(-K(1-\frac{x}{T})t-1.27\frac{Kx}{T}\right) \qquad t < x < 45 \qquad (3.4)$$

$$= \frac{b_K}{b_s}\exp\left(-K(1-\frac{x}{T})t-1.27\frac{Kx}{T}\right)s(x,t) \qquad 45 < x$$

$$t > \frac{3T}{4}$$

$$\frac{N_x(t)}{N_x^s(t)} = \frac{b_K}{b_s} \exp\left(-\frac{K}{2}x + \frac{K}{2T}x^2 - 1.27\frac{Kx}{T}\right) \qquad 0 < x < t - \frac{3T}{4}$$

$$= \frac{b_K}{b_s} \exp\left(-.09375KT - \frac{K}{2}t^2 - K\left(1 - \frac{t}{T}\right)x - 1.27\frac{Kx}{T} + \frac{K}{2}t\right) \qquad t - \frac{3T}{4} < x < t \tag{3.5}$$

$$= \frac{b_K}{b_s} \exp\left(-.09375KT - \frac{K}{2}t - \frac{K}{2T}t^2 + \frac{Kt}{T}x - 1.27\frac{K}{T}x\right) \qquad t < x < 45$$

$$= \frac{N_x(t)}{N_x^s(t)}\, s(x,t) \qquad 45 < x < \omega$$

where:

$$s(x,t) = \exp\left(-B1(x-45)^3/6\right) \qquad 45 < x < t + 45$$

$$= \exp\left(-B1\left[\frac{(x-45)^2 t}{2} - \frac{(x-45)t^2}{2} + \frac{t^3}{6}\right]\right) \qquad t + 45 < x < \omega$$

and:

T: mean length of generation in the stable population

t: the number of years that mortality has been declining

$\frac{N_x(t)}{N_x^s(t)}$: proportionate distribution at age x and time t relative to the current stable age distribution

b_K: birth rate in a population with a history of changing fertility at an annual rate K for t tyears

K: the annual proportionate increase in fertility to which the annual increase in survival at young ages is equivalent, or more precisely: $K = \int_0^5 z(x)dx$, $z(x)$ = the annual decline in age specific mortality – the level portion from 5 to 45.

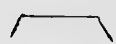
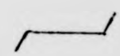
b_s: birth rate in the stable population

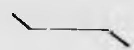
$B1$: the slope of the line approximating the rate at which the annual change in age specific mortality rates increases above 45.

Before using the previous expressions to evaluate z_y we need first to discuss the underlying postulates and their feasibility and estimates of values of K, Bl and T inherent in the pattern of mortality change used in the first approach.

Feasibility of the Underlying Postulates:

It was mentioned before that - within the available information - it is generally more likely that the age pattern of decline in mortality followed a reversed J or U. (when the pattern is reversed J, Bl may be set to zero). Also, in principle, Coale formulae do apply to cases when the age pattern of change resemble either of the following shapes:

 (K & Bl - ve) ,  (K - ve, Bl + ve),

 (K + ve, Bl - ve)

Thus the only restriction in postulate (2) is that the section between age 5 and 45 is considered level, this is a fair approximation to the actual pattern of change experienced in mortality data.

Postulates (1) and (3) may be challenged on the grounds that age patterns of change tend to be erratic over short periods, but once it is agreed that the cumulative effects of these short period changes may be closely approximated by the average changes over a longer interval the previous postulates may be readily accepted.

The data presented in Table (3.11) are presented in (Coale 1972) as an illustration of the feasibility of postulate (1).

Table (3.11)* Difference in age-specific mortality rates, 'west' female model life tables

age x	$\Delta U(x)$, e^o_o = 20 and 50	13.9 $\Delta U(x)$, e^o_o = 35 and 37.5
0.5	.351	.346
1.5	.106	.110
2.5	.054	.054
3.5	.038	.037
4.5	.030	.030
7.5	.011	.011
12.5	.008	.009
17.5	.011	.011
22.5	.013	.014
27.5	.015	.016
32.5	.017	.018
37.5	.018	.019
42.5	.019	.020
47.5	.019	.020
52.5	.024	.025
57.5	.030	.031
62.5	.044	.045
67.5	.055	.056
72.5	.075	.075
77.5	.096	.095

*source: reproduced from (Coale 1972), table 5.1

Estimates of the Parameters:

Estimate of K

$$e^K = e^{\int_0^5 z(a)da}$$

where z(a) = annual decline in age specific mortality - the level portion from 5 to 45.

$$z(a) \;=\; (U(a,t) - U(a,t+n) - U(25,t) + U(25,t+n))/n$$

where U(a,t) denotes the force of mortality around age a at time t, U(25,t) is an approximation for the level portion.

$$e^K \;=\; e^{\int_0^5 \{U(a,t) - U(a,t+n) + U(25,t+n) - U(25,t)\}/n}$$

$$e^K = \left[\frac{{}_5P_0(t+n)\,.\,{}_5P_{25}(t)}{{}_5P_0(t)\cdot{}_5P_{25}(t+n)}\right]^{\frac{1}{n}}$$

where ${}_nP_x$: age specific survival rates from x to x+n

$$K = \frac{1}{n}\left[\log\frac{l_5(t+n)\,.\,l_{25}(t+n)}{l_0(t+n)\cdot l_{30}(t+n)} - \log\frac{l_5(t)\,.\,l_{25}(t)}{l_0(t)\cdot l_{30}(t)}\right]$$

where $l_x(t)$: life table survivors at age x and time t.

Estimate of B1

$$B1 = (z(65) - z(50))/15$$

$$= \{U(65,t) - U(65,t+n) - U(50,t) + U(50,t+n)\}/15n$$

If we approximate U(x,t) by ${}_nm_x(t)$, where ${}_nm_x(t)$ denotes the age specific death rates for age group (x-x+n) at time t, then

$$B1 = \frac{({}_5m_{65}(t) - {}_5m_{50}(t)) - ({}_5m_{65}(t+n) - {}_5m_{50}(t+n))}{15n}$$

Estimate of T

Several approximate formulae are available for estimating T; they all depend on the current fertility and mortality schedules. Possible values of T range from 26 to 34 years. In all the following applications T will be assigned a constant value 30 simply because a small deviation in T will hardly affect the illustrations.

Table (3.12) shows value of Kt & B1.t inherent in changes in life expectancy by two and a half years increments from 30 to 50 in the west female model life tables as calculated by Coale (1972).

Table (3.13) shows values of K and B1 (calculated using the previous approximations) inherent in mortality change following Chile pattern from 1920 to 1959 and for Taiwan from 1920 to 1960 and Portugal from 1920 to 1960.

Table (3.12)* Values of Kt and B1.t inherent in the west female model life tables

change in e_o^o	Kt	B1.t . 10^3
30 to 32.5	.0424	.179
32.5 to 35.0	.0379	.161
35.0 to 37.5	.0345	.145
37.5 to 40.0	.0315	.133
40.0 to 42.5	.0289	.121
42.5 to 45.0	.0266	.111
45.0 to 47.5	.0246	.103
47.5 to 50.0	.0233	.095

*source: reproduced from Coale (1972), Table 5.2.

Table (3.13) Values of K & B1 inherent in Chile, Taiwan and Portugal's mortality change

country	K	B1
Chile (1920-1959)	.008	.00001
Taiwan (1920-1960)	.008	.00003
Portugal (1920-1960)	.008	.00001

Now let us reconsider the term z_y:

$$z_y = \int_y^w N_x \, d \log \frac{N_x}{N_x^s} / P_y$$

From formulae (3.4) and (3.5), it is clear that $\frac{N_x}{N_x^s}$ (and consequently $d \log \frac{N_x}{N_x^s}$) is not a continuous function over the whole range from y to w, but is continuous over several limited intervals within the whole range. The number of these intervals and their width depend on the length of mortality decline and the age considered.

For example, in the special case where $t < \frac{3T}{4}$ and $y < t$, $\frac{N_x}{N_x^s}$ is continuous within four smaller intervals given respectively by:

$$\begin{aligned}
d \log \frac{N_x}{N_x^s} &= -K(1-\frac{t}{T}) - 1.27\frac{K}{T} && 5 < x < t \\
&= \frac{Kt}{T} - 1.27\frac{K}{T} && t < x < 45 \\
&= \frac{Kt}{T} - 1.27\frac{K}{T} - \frac{B1}{2}(x-45)^2 && 45 < x < t+45 \\
&= \frac{Kt}{T} - 1.27\frac{K}{T} - B1(t(x-45)-\frac{t^2}{t}) && t+45 < x < \omega
\end{aligned}$$

$$\int_y^w N_x \, d \log \frac{N_x}{N_x^s} = \left[-K(1-\frac{t}{T}) - 1.27\frac{K}{T}\right] \int_y^t N_x dx + (\frac{Kt}{T} - 1.27\frac{K}{T}) \int_t^{45} N_x d_x$$

$$+ \int_{45}^{t+45} (\frac{Kt}{T} - 1.27\frac{K}{T} - \frac{B1}{2}(x-45)^2) N_x dx$$

$$+ \int_{t+45}^{w} \left(\frac{Kt}{T} - 1.27\frac{K}{T} - B1.t(x-45) - \frac{t^2}{2}\right) N_x dx$$

Let us deal with the general case when $\frac{N_x}{N_x^s}$ is continuous within m intervals defined by $y \to a_1$, $a_1 \to a_2$, ..., $a_{m-1} \to w$; then

$$z_y = \left\{ \int_y^{a_1} N_x \, d \log \frac{N_x}{N_x^s} + \int_{a_1}^{a_2} N_x \, d \log \frac{N_x}{N_x^s} + \cdots + \int_{a_{m-1}}^{w} N_x \, d \log \frac{N_x}{N_x^s} \right\} / P_y \qquad (3.6)$$

let $V_x = d \log \frac{N_x}{N_x^s} / dx$

$$z_y = \left\{ \int_y^{a_1} N_x V_x d_x + \int_{a_1}^{a_2} N_x V_x d_x + \cdots + \int_{a_{m-1}}^{w} N_x V_x d_x \right\} / P_y \qquad (3.7)$$

Since V_x is continuous within the small intervals, then using the mean value theorem, z_y may be approximated as:

$$z_y = \left\{ V_{x1} \int_y^{a_1} N_x d_x + V_{x2} \int_{a_1}^{a_2} N_x d_x + \cdots + V_{xm} \int_{a_{m-1}}^{w} N_x d_x \right\} / P_y \qquad (3.8)$$

where $y < x1 < a_1$

$a_1 < x2 < a_2$

$\vdots$

$a_{m-1} < xm < w$

$$z_y = \sum_{i=1}^{m} V_{xi} \, c_i \qquad (3.9)$$

where $c_1 = \int_y^{a_1} N_x d_x \Big/ \int_y^{w} N_x d_x$

$c_2 = \int_{a_1}^{a_2} N_x d_x \Big/ \int_y^{w} N_x d_x$

$\vdots$

$c_m = \int_{a_{m-1}}^{w} N_x d_x \Big/ \int_y^{w} N_x d_x$

$$\sum_{i=1}^{m} c_i = 1$$

In other words, z_y is the weighted mean for W_{xi}. It will be shown in the following cases that when the growth balance estimate is not affected by mortality decline, small values of W_{xi} are connected with big weights while big values of W_{xi} are connected with small weights; thus resulting in negligible values for z_y (especially corresponding to old ages).

Case (1)

Considering the values of Kt and B1.t - inherent in changes in life expectancy from 30 to 50 years in west female model life tables - presented in Table (3.12). Assuming that this gain in life expectancy was accomplished in 20 years - a plausible assumption with regard the speed of decline in mortality experienced in some developing countries - we get:

$K = .0124, \quad B1 = .0000524 \qquad T = 30 \qquad t = 20$, and:

$$W_{x1} = -K\left(1 - \frac{t}{T}\right) - 1.27\frac{K}{T} \qquad x < 20$$

$$W_{x2} = \frac{Kt}{T} - 1.27\frac{K}{T} \qquad 20 < x < 45$$

$$W_{x3} = \frac{Kt}{T} - 1.27\frac{K}{T} - \frac{B1}{2}(x3 - 45)^2 \qquad 45 < x < 65$$

$$W_{x4} = \frac{Kt}{T} - 1.27\frac{K}{T} - B1.t\,(x4 - 45) + B1\frac{t^2}{2} \qquad 65 < x < \omega$$

to estimate z_y (for $y < 65$), x4 is approximated as $\frac{65+\omega}{2}$ and ω set arbitrary to 85, x3 is approximated as $\frac{45+65}{2}$ for $y < 45$ and as $\frac{y+65}{2}$ for $45 < y < 65$. Actually, better approximations for z_y may be achieved, but since our aim is to illustrate roughly the magnitude of z_y the procedure is kept as simple as possible.

Using the age distribution of female, model west, corresponding to $e_o = 50$ and $r = 15$ (level 13) - the age distribution that should be used in this case is different - we get the following values of z_y as compared to $\frac{N_y}{P_y}$.

Table (3.14) The values of z_y and $\frac{N_y}{P_y}$ when mortality is declining following the west female model life tables pattern

age	z_y	N_y/P_y
5	.0014	.0282
10	.0022	.0289
15	.0032	.0305
20	.0044	.0324
25	.0040	.0345
30	.0033	.0370
35	.0024	.0401
40	.0015	.0441
45	.0003	.0491
50	-.0018	.0559
55	-.0050	.0651
60	-.0084	.0776

The negligible values of z_y illustrate the known fact that changes in mortality following the west model are believed to affect the stable distribution and consequently $\frac{N_y}{P_y}$ slightly.

<u>Case (2)</u>

Using the values of K and B1 presented in Table (3.13). We get for Chile:

$$t = 39$$

$$W_{x1} = -\frac{K}{2} - 1.27\frac{K}{T} + \frac{K}{T} \times 1 \qquad 5 < x < 16.5$$

$$W_{x2} = -K(1-\frac{t}{T}) - 1.27\frac{K}{T} \qquad 16.5 < x < 39$$

$$W_{x3} = \frac{Kt}{T} - 1.27\frac{K}{T} \qquad 39 < x < 45$$

$$W_{x4} = \frac{Kt}{T} - 1.27\frac{K}{T} - \frac{1}{2} Bl\ (x4 - 45)^2 \qquad 45 < x < 84$$

$$W_{x5} = \frac{Kt}{T} - 1.27\frac{K}{T} - Bl\ (x5 - 45)t + \frac{t^2}{2} Bl \qquad 84 < x < \omega$$

To simplify the calculations the previous first three expressions are assumed to correspond to ages: 5-15, 15-40 and 40-45. The last two expressions are modified to:

$$W_{x4} = \frac{Kt}{T} - 1.27\frac{K}{T} - \frac{1}{2} Bl\ (x4 - 45)^2 \qquad 45 < x < \omega$$

and W_{x5} is neglected. (The value of z_y, $y > 85$ does not exceed .0003.) x4 is approximated as: $\frac{45 + \omega}{2}$ for $y < 45$, ω is set arbitrary to 85 and as $\frac{y+\omega}{2}$ for $y > 45$.

Using the projected age distribution of Chile corresponding to 1959 we get the following values of z_y as compared to N_y/P_y.

Table (3.15) The values of z_y and $\frac{N_y}{P_y}$ when mortality is declining following Chile pattern from 1920 to 1959

age	z_y	N_y/P_y	z_y - .0032
5	.0032	.0377	0
10	.0045	.0383	.0013
15	.0056	.0390	.0024
20	.0053	.0401	.0021
25	.0052	.0420	.0020
30	.0055	.0446	.0023
35	.0065	.0471	.0033
40	.0078	.0498	.0046
45	.0071	.0540	.0039
50	.0090	.0612	.0068
55	.0087	.0721	.0055
60	.0074	.0865	.0042

From Table (3.15), we note that z_y or more precisely (z_y-.0032) - since a constant value of z_y over all ages does not affect the estimate of the slope - are negligible.

Case (3)

The previous two cases focused on situations when z_y is relatively small and thus the estimate is not affected. It may be of interest to show the magnitude of z_y when the estimate is affected.

Let us consider the first 10 years of mortality decline in Chile; following the same previous procedure we get: K = -.008, B1 = -.00013. The negative signs of both K and B1 and the magnitude of B1 illustrate the extraordinary age pattern of mortality change, during this short period, as the change is mainly due to decline in the mortality of middle age groups associated with an increase of mortality in both young and old age groups. This type of decline is not very common in actual situations - it may be a result of error in the data - and unlikely to persist for a long period.

Table (3.16) presents the annual change % in age specific death rate in the first ten years of mortality decline in Chile, Taiwan and Portugal.

Table (3.17) shows values of z_y as compared to N_y/P_y for Chile in 1920.

Comparing Table (3.17) and (3.15) we note that the values of z_y are considerably bigger in Table (3.17) (especially corresponding to old ages). This explains the deviation in the estimate of the crude death rate of Chile after 10 years.

Table (3.16) Annual change in age specific death rates in the first ten years

Country age	Chile	Taiwan	Portugal
0-	-.014%	-.475%	-.726%
1-	.139	-.245	-.236
5-	.028	-.082	-.056
10-	-.006	-.038	-.021
15-	-.030	-.060	-.022
20-	-.052	-.096	-.037
25-	-.066	-.129	-.025
30-	-.029	-.165	-.015
35-	-.039	-.183	-.019
40-	-.025	-.188	-.033
45-	-.012	-.175	-.032
50-	.028	-.195	-.049
55-	-.011	-.230	-.053
60-	.111	-.267	-.061
65-	.227	-.337	-.103
70-	.323	-.282	-.182
75-	.586	-.391	-.396
80-	.995	-.566	-.227

Table (3.17) The values of z_y and $\frac{N_y}{P_y}$ when mortality is declining following Chile pattern from 1909 to 1920

y	z_y	N_y/P_y
5	.002	.030
10	.002	.031
15	.003	.034
20	.004	.037
25	.005	.040
30	.007	.042
35	.009	.045
40	.012	.050
45	.017	.056
50	.024	.061
55	.033	.070
60	.033	.083

3.3.2 Possible Modifications on the Death Distribution Method to allow for Changing Mortality

The discussion up till now was only limited to justifying the deviation in the estimate due to declining mortality but no attempt was made to modify the method of estimation to allow for this mortality decline. Actually, if the pattern of decline is believed to be fairly approximated by the assumptions used to derive the formulae (3.4) and (3.5) and if the time since the initiation of the decline is known, the formulae are readily generalized.

For example, if mortality is declining for time t, where $t \geq 45 + \frac{3}{4}T$, then using equation (3.3) and (3.5) we get:

$$\frac{N_y}{P_y} = \cdot\, r + \mathrm{CDR}\,\frac{D_y}{P_y} - \int_y^{\omega} N_x \, d \log \frac{N_x}{N_x^s} \Big/ P_y$$

and

$$\frac{N_x}{N_x^s} = \frac{h_K}{h_s} \exp\left(-\frac{K}{2}x + \frac{K}{2T}x^2 - 1.27\frac{K}{T}x\right) \qquad 5 < x < 45$$

$$= \frac{h_K}{h_s} \exp\left(-\frac{K}{2}x + \frac{K}{2T}x^2 - 1.27\frac{K}{T}x - \frac{B1}{6}(x-45)^3\right) \qquad 45 < x < \omega$$

But,

$$\frac{N_y}{P_y} = r + CDR\frac{D_y}{P_y} - \int_y^{45} N_x \frac{d \log \frac{N_x}{N_x^s}}{P_y} - \int_{45}^{\omega} N_x \frac{d \log \frac{N_x}{N_x^s}}{P_y} \qquad 5 < y < 45$$

$$\frac{N_y}{P_y} = r + CDR\frac{D_y}{P_y} - \int_y^{\omega} N_x \frac{d \log \frac{N_x}{N_x^s}}{P_y} \qquad 45 < y < \omega$$

then,

$$\frac{N_y}{P_y} = \left(r + \frac{K}{2} + 1.27\frac{K}{T}\right) + CDR\frac{D_y}{P_y} - \frac{K}{T}\frac{\int_y^{\omega} x\, N_x dx}{P_y} + \frac{B1}{2}\frac{\int_{45}^{\omega} (x-45)^2 N_x dx}{P_y} \qquad 5 < y < 45 \qquad (3.10)$$

$$\frac{N_y}{P_y} = \left(r + \frac{K}{2} + 1.27\frac{K}{T}\right) + CDR\frac{D_y}{P_y} - \frac{K}{T}\frac{\int_y^{\omega} x\, N_x dx}{P_y} + \frac{B1}{2}\frac{\int_y^{\omega} (x-45)^2 N_x dx}{P_y} \qquad 45 < y < \omega$$

Thus instead of a linear equation of the form: $y = r + CDRx_1$, we have an equation of the form: $y = a_0 + CDRx_1 + a_2 x_2 + a_3 x_3$ where: $y = \frac{N_y}{P_y}$,

$$x_1 = \frac{D_y}{P_y},\ x_2 = \frac{\int_y^{\omega} x\, N_x dx}{P_y},\ x_3 = \frac{\int_{45}^{\omega} (x-45)^2 N_x dx}{P_y} \text{ for } y < 45 \text{ and}$$

$$x_3 = \frac{\int_y^{\omega} (x-45)^2 N_x dx}{P_y} \text{ for } y > 45.$$

Numerical methods for evaluating x_2 and x_3 may be easily suggested and the death rate estimated using multiple regression.

Theoretically, expression (3.10) is exact. But, in practise, we believe that the advantage gained by introducing this expression (or similar ones depending on the value of t) may be offset due to the extra complications involved and to the extra knowledge of the duration of mortality decline required and to the fact that the deficiencies associated with actual data (differential under-registration, age errors and deviations from the theoretical pattern of decline) are likely to affect this new expression. Thus, in view of the robustness of the growth balance estimate, the previous modification is not recommended for actual applications, unless very significant or atypical (with respect to the age pattern of decline) changes in mortality are suspected.

3.3.3 Justification for the Effect of Fertility Decline

Most developing countries experienced only a recent fertility change, thus the type of expression that is relevant to our discussion is the one concerned with short period changes in fertility.

When fertility is fixed in age structure but changing in level at a constant annual rate (K), then the proportionate age distribution in the years near t = 0 when the fertility decline begins at that point given in Coale (1972) as:

$$t < \frac{3T}{4}$$

$$N_x(t) = b(t)\ e^{-r(t)x-K(1-\frac{t}{T})^x}\ l_x \qquad x < t$$

$$N_x(t) = b(t)\ e^{-r(t)x-K(1-\frac{x}{T})^t}\ l_x \qquad x > t$$

where:

t: the number of years that fertility has been declining

T: mean length of generation in the stable population

$N_x(t)$: proportionate age distribution at age x and time t

b(t): birth rate at time t

r(t): intrinsic rate of growth corresponding to fertility and mortality of time t

l_x: life table survivors at age x when $l(o) = 1$.

Since:

$$\frac{N_y}{P_y} = CDR \frac{D_y}{P_y} - \frac{\int_y^{\omega} N_x \, d \log N_x / l_x}{P_y}$$

then,

$$\frac{N_y}{P_y} = CDR \frac{D_y}{P_y} - \left\{ \int_y^{t} N_x \left(-r(t) - K\left(1-\frac{t}{T}\right)\right) + \int_t^{\omega} \left(-r(t) + \frac{Kt}{T}\right) N_x \right\} / \int_y^{\omega} N_x d_x \qquad x < t$$

$$\frac{N_y}{P_y} = CDR \frac{D_y}{P_y} - \int_y^{\omega} \left(-r(t) + \frac{Kt}{T}\right) N_x \Big/ \int_y^{\omega} N_x d_x \qquad x > t$$

finally:

$$\frac{N_y}{P_y} = \left(r(t) - \frac{Kt}{T}\right) + CDR \frac{D_y}{P_y} + \frac{K \int_y^{t} N_x d_x}{\int_y^{\omega} N_x d_x} \qquad x < t$$

$$= \left(r(t) - \frac{Kt}{T}\right) + CDR \frac{D_y}{P_y} \qquad x > t \qquad (3.11)$$

Thus, when fertility is subject to this specific type of change, the growth balance method is exact for points corresponding to ages older than t, except of course that the intercept now denotes $\left(r(t) - \frac{Kt}{T}\right)$ which is approximately equivalent to the intrinsic rate of growth before fertility change begins (at $t = 0$).

Since t is small and points corresponding to young ages don't affect the estimate greatly, direct application of the growth balance formula on all ages yields good results.

3.3.4 Possible Modifications on the Growth Balance Method to allow for Fertility Changes

A case that may be theoretically interesting and result in a simple modification of Brass method is that when fertility has been declining for a long time. In this case the proportionate age distribution as presented by Coale (1972) is:

$$N_x(t) = b(t)\, e^{-\frac{K}{2} x + \frac{K}{2T} x^2 - r(t)\, x}\; l_x$$

thus:

$$\frac{N_y}{P_y} = (r(t) + \frac{K}{2}) + CDR \frac{D_y}{P_y} - \frac{K}{T} \frac{\int_y^{\omega} x\, N_x\, dx}{P_y} \qquad (3.12)$$

In practise, $\int_y^{w} x\, N_x dx$ may be approximated by Pl_y where:

$$Pl_y = \sum_{i=}^{\omega} \bar{x}_i\, A_i$$

$\bar{x}_i$: mean age corresponding to age group $(i - i+n)$

A_i: proportion in the age group $(i - i+n)$

For an illustration, a hypothetical age stable age distribution was projected for a long period. The initial growth rate was purposely assigned a high value of 0.05. Mortality was held constant corresponding to level 17 for female model west in Coale & Demeny (1966). Fertility was declining such that age specific fertility rate at (t+1) = age specific fertility rates at time (t) x $e^{-0.1}$. The actual death rate and the estimated death rate, using the growth balance method and formula (A) and the estimated death rate using expression (3.12) are presented in table (3.18) after 20 periods of fertility decline equivalent to 100 years.

In another application, similar to the previous illustration, but corresponding to high mortality (level 2, model west, females), a different method of estimation was attempted. Thus, instead of applying equation (3.12) directly we used:

Table (3.18) The actual and estimated death rate using the growth balance method and a modified method

period	actual death rate	estimated (growth balance method)	estimated (equation 3.12)
21	23.33	27.10	22.60
22	25.05	29.13	24.27
23	26.84	31.25	26.20
24	28.70	33.49	28.32

$$\frac{A_y}{n\ P^*_y} = (r(t) + \frac{K}{2}) + CDR\frac{D^*_y}{P^*_y} - \frac{K}{T}\frac{Pl^*_y}{P^*_y} \qquad (3.13)$$

where A_y, P^*_y and D^*_y are as identified in equation (2.4)

$$Pl^*_y = (Pl_y + Pl_{y+n})/2$$

This is of course the equivalent of formula (F) when using the modified equation. The actual death rate and the estimated using both the growth balance method (formula (A)) and formula (3.13) are given in Table (3.19) after 18 period of fertility decline equivalent to 80 years.

Table (3.19) The actual and estimated death rate using the death distribution method and a modified method

period	actual death rate	estimated (Brass)	estimated (equation 3.13)
18	39.87	51.65	41.95
19	39.59	50.66	39.42
20	39.46	49.74	39.58
21	39.49	49.34	37.21
22	39.68	49.40	37.80

Table (3.19) (continued)

period	actual death rate	estimated (Brass)	estimated (equation 3.13)
23	40.04	49.67	38.75
24	40.56	50.05	39.83
25	41.24	50.55	40.96
26	42.07	51.25	42.19
27	43.06	52.19	43.61
28	44.18	53.33	45.27

While the estimated rates using the modified expressions are a much better estimate for the actual death rate, it should be emphasized that the combinations of mortality and fertility, used in this illustration, are unrealistic and unfeasible in actual applications.

CHAPTER IV

THE EFFECT OF MIGRATION ON THE GROWTH BALANCE METHOD

4.1 INTRODUCTION

The effect of migration on the applicability of the growth balance method for mortality estimation is treated here. An adjustment procedure to allow for the effect of recent migration is presented and illustrated using actual data for Kuwait. A discussion of this procedure, especially in connection to its data requirements, and the general likely effect of migration is also introduced.

It should be stressed that we are not only interested in the flow of migrants during a certain period but also in the cumulative effect of migration over the past recent history of the country. Even if migration is not significant in terms of the total numbers, it may still affect the age structure due to its age selective nature.

In the following part, no attempt is made to differentiate between the effect of internal and international migration since the same principle applies to either case. Thus, the term foreign and home born does not necessarily apply to different countries but may denote different regions in the same country.

4.2 ADJUSTMENT PROCEDURE FOR THE EFFECT OF RECENT MIGRATION

Suppose we start with a population that follows a stable model. If this population is subjected to a migration movement for a certain period, then the resulting population at the end of this period is affected by the effective contribution of immigration or emigration to the age distribution.

Let G_x denote the population aged x at the end of the migration period, in case there was no in and out migrants. Let n_x denotes the resulting population aged x, affected by in and out migration.

For any age distribution, it was shown that:

$$\frac{N_y}{P_y} = \text{CDR}\ \frac{D_y}{P_y} - \frac{\int_y^{\omega} N_x . d \log \frac{N_x}{l_x}}{P_y} \qquad \text{(equation 2.4)}$$

rewriting the previous equation in terms of numbers, instead of proportions, we get:

$$\frac{n_y}{P_y} = \frac{f.\ d_y}{P_y} - \frac{\int_y^{\omega} n_y\ d \log \frac{n_x}{l_x}}{P_y} \qquad (4.1)$$

$$= \frac{f.\ d_y}{P_y} - \frac{\int_y^{\omega} n_y\ d \log \frac{G_x}{l_x} . \frac{n_x}{G_x}}{P_y} \qquad (4.2)$$

let $E_x = \frac{n_x}{G_x} = \frac{\text{total population aged x}}{\text{population aged x, in case of no in and out migrants}}$

Since G_x follows a stable form, then:

$$\frac{n_y}{P_y} = \frac{f\ d_y}{P_y} - \frac{\int_y^{\omega} n_x\ d \log (Be^{-rx} . E_x)}{P_y} \qquad (4.3)$$

where B and r denote the number of births and rate of growth of the population in case of no in and out migrants.

Differentiating and rearranging (4.3) we get:

$$\frac{n_y}{P_y} + \frac{\int_y^{\omega} n_x\ d \log E_x}{P_y} = r + \frac{f\ d_y}{P_y} \qquad (4.4)$$

and, finally:

$$\frac{n_y}{P_y} + \frac{\int_y^{\omega} n_x \frac{E'_x}{E_x}}{P_y} = r + \frac{f\, d_y}{P_y} \tag{4.5}$$

where E'_x is the first derivative of E_x.

Thus, to offset the effect of migration, an adjustment term is added to the formula. For numerical evaluation of this term the following is suggested:

$$\int_y^{\omega} n_x \frac{E'_x}{E_x}\, dx = \sum_{i=1}^{\frac{\omega-y}{m}} \frac{E'_{y+(i-\frac{1}{2})m}}{E_{y+(i-\frac{1}{2})m}} A_{y+(i-1)m}$$

where m: length of the age interval.

$A_{y+(i-1)m}$: age distribution corresponding to age group $(y + (i-1)m$ to $y + im)$.

$E_{y+(i-\frac{1}{2})m}$: the value of E corresponding to the age group $(y + (i-1)m$ to $y + im)$.

$E'_{y+(i-\frac{1}{2})m}$: the rate of change of E in the age group and may be approximated as $(E_{y+(i+\frac{1}{2})m} - E_{y+(i-1\frac{1}{2})m})/2m$.

To calculate E_x, we need to know the distribution by ages of the population in case of no in and out migration.

The population in the case of no in and out migration = natives calculated in the census + natives of the country studied living in other countries. (4.6)

Actually, the previous expression is a rough approximation since the natives calculated in the census are inflated by those who acquire nativity, and the natives living in other countries are deflated by those who acquire the nativity of other countries. These terms are quite difficult to estimate. A more exact expression, if the data permits, is given as:

The population in case of no in and out migration = natives calculated in the census - natives whose parents are foreign + natives of the country studied living in other countries + foreigners whose parents are natives of the country studied. (4.7)

4.3 ILLUSTRATION ON ACTUAL DATA

Migration in Kuwait plays a vital role. It is a country with rich oil resources and high standard of living which attracts immigrants. The age structure of those immigrants has typically a concentration around the labour force range. The data for Kuwait shows that male immigrants constitute 58% of the total male population in 1970 and 69% of ages 15 - 45.

Fertility of Kuwait natives seems to have remained at a relatively high level. Mortality has been reduced as a result of the development activities connected with the discovery of oil in Kuwait.

Kuwait is an ideal case for the application of the migration adjustment procedure; data on Kuwaiti and non-Kuwaiti population as well as deaths are available by age and given in table (4.1).

Table (4.1) Age and death distribution of Kuwaiti and non-Kuwaiti males, 1970 census.

age	population		deaths[1]	
	Kuwaiti	non Kuwaiti	Kuwaiti	non Kuwaiti
0-	34073	35109	444	411
5-	30607	25076	44	24
10-	23709	14633	21	11

age	population		deaths[1]	
	Kuwaiti	non Kuwaiti	Kuwaiti	non Kuwaiti
15–	16615	16237	14	12
20–	13638	28237	14	28
25–	11887	35473	21	54
30–	9924	30286	26	61
35–	8859	23883	28	55
40–	6417	15643	25	52
45–	5199	9398	27	48
50–	4606	5243	45	63
55–	2525	2352	37	33
60–	2584	1511	63	49
65–	1418	563	66	29
70–	1743	398	78	32
75–	696	154	57	11
80–	628	96	54	13
85+	386	67	74	16
Total	175514	244359	1138	1002

[1]Deaths: average of the three years (1969, 1970, 1971).

For a country like Kuwait, out migration may be assumed negligible, thus:

$$E_x = \frac{\text{population aged } x}{\text{natives aged } x \text{ calculated in the census} - \text{natives aged } x \text{ whose parents are foreign.}}$$

Suppose we approximate E_x as:

$$E_x = \frac{\text{population aged } x}{\text{Kuwaiti population aged } x.}$$

Of course, Kuwaiti population include all persons who changed their nationality and consequently the values calculated for E_x are not exact. Nevertheless, they may still serve our purpose.

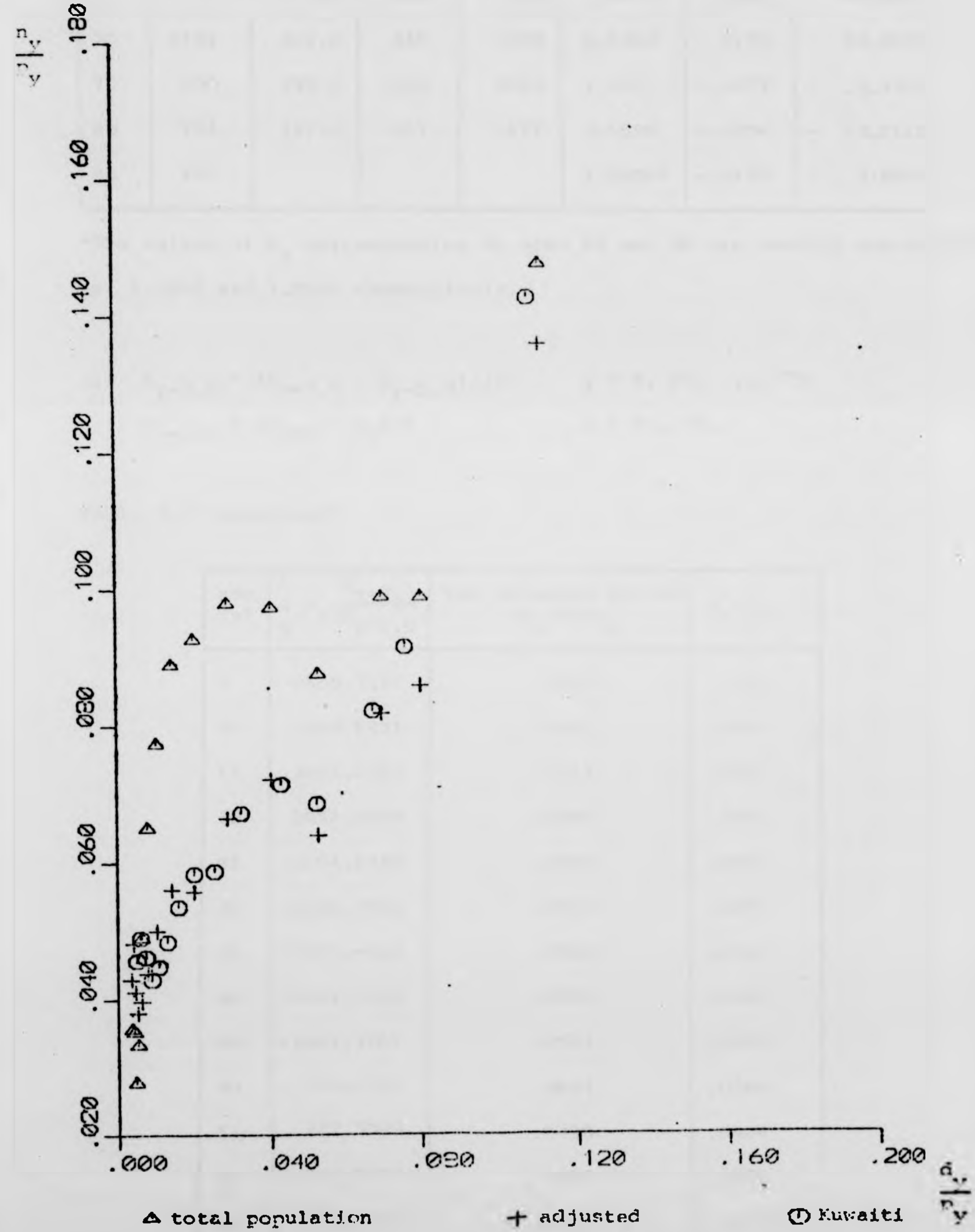
The details of calculating the adjustment term and the adjusted set of points are given in table (4.2). Graph (4.1) represents the sets of points $(\frac{n_y}{P_y}, \frac{d_y}{n_y})$ and the adjusted sets of points corresponding to total population. The sets of points corresponding to the data of Kuwaiti only are also shown in the same graph.

The improvement of the adjusted set and its similarity with the set corresponding to Kuwaiti only is quite noticeable.

Table (4.2) The details of calculating the adjusted sets of points to allow for migration.

age (y)	A_y	n_y	d_y	P_y	$E_{y+2.5}$	$E'_{y+2.5}$ [1]	$A_y \cdot \frac{E'_{y+2.5}}{E_{y+2.5}}$
0	69182				2.0304		
5	55683	12486.5	1285	350695	1.8192	-.0413	-1264.1314
10	38342	9402.5	1217	295012	1.6171	.0158	374.6234
15	32852	7119.4	1185	256670	1.9772	.1453	2414.2199
20	41875	7472.7	1159	223818	3.0704	.2006	2735.8406
25	47360	8923.5	1117	181943	3.9841	.0981	1166.1394
30	40210	8757.0	1042	134583	4.0517	-.0288	- 285.8178
35	32742	7295.2	955	94369	3.6959	-.0614	- 543.9429
40	22060	5480.2	872	61627	3.4377	-.0888	- 569.8368
45	14597	3665.7	795	39567	2.8076	-.1299	- 675.3634
50	9849	2444.6	720	24960	2.1382	-.0876	- 403.5040
55	4877	1472.6	612	15121	1.9314	-.0553	- 139.6386
60	4095	897.2	542	10244	1.5847	-.0534	- 137.9901

Graph (4.1) The points corresponding to total population of Kuwait, the adjusted points and the points corresponding to Kuwaiti only



	A_y	n_y	d_y	P_y	$E_{y+2.5}$[1]	$E^*_{y+2.5}$[1]	$A_y \cdot \frac{E^*_{y+2.5}}{E_{y+2.5}}$
65	1981	607.6	430	6149	1.3970	-.0356	- 50.4821
70	2141	412.2	335	4168	1.2283	-.0175	- 30.5035
75	850	299.1	225	2027	1.2212	-.0075	- 5.2202
80	724	157.4	157	1177	1.1528	-.0204	- 12.8119
85	453				1.0508*	-.0136	- 5.8629

*The values of E_x corresponding to ages 85 and 90 are roughly approximated as: 1.0848 and 1.0168 respectively.

(1) $E^*_{y+2.5} = (E_{y+7.5} - E_{y-2.5})/10$ y = 5, 10, ..., 75.

$E^*_{y+2.5} = (E_{y+5} - E_y)/5$ y = 80, 85.

Table 4.2 (continued)

age (y)	$\sum_y^{\omega} A_y \cdot \frac{E^*_{y+2.5}}{E_{y+2.5}}$	The adjusted points $(n_y+\Sigma)/P_y$	d_y/P_y
5	2565.7177	.0429	.0036
10	3829.8491	.0448	.0041
15	3455.2257	.0411	.0046
20	1041.0058	.0380	.0051
25	-1694.8348	.0397	.0061
30	-2860.9742	.0438	.0077
35	-2575.[illegible]564	.0500	.0101
40	-2031.2135	.0559	.0141
45	-1461.3767	.0557	.0200
50	- 786.0133	.0664	.0288
55	- 382.5093	.0720	.0404
60	- 242.8707	.0638	.0529
65	- 104.8806	.0817	.0699

age (y)	$\sum_{y}^{\omega} \Lambda_y \cdot \frac{E^*_{y+2.5}}{E_{y+2.5}}$	The adjusted points $(n_y+\Sigma)/p_y$	d_y/p_y
70	- 54.3985	.0858	.0803
75	- 23.8950	.1358	.1110
80	- 18.6748		

4.4 DISCUSSION OF THE ADJUSTMENT PROCEDURE

For a country where in migration plays a dominant role, the data needed to apply the adjustment procedure may generally be estimated either directly or through indirect calculations.

The same may not be true when out migration plays an important role, since the procedure requires the knowledge of native persons living in all other countries by age group. This presupposes detailed statistics on foreign born not only classified by age but also by country of origin. Actually, when the statistics are available one may reduce the analysis to practical proportions by considering only the principal country or countries receiving emigrants from the country under study.

The effect of urbanization may be taken into account by applying the balance growth method to the country as a whole and the adjusted procedure to the cities. The rural rates may then be estimated using residual procedures.

Another difficulty associated with this procedure is in estimating the native number of each age whose parents are foreign. We did not go further than the direct parents as it may be true that if migration has been going on for a long enough period with appropriate constancy the population is likely to stabilize.

In a country where in migration is dominant and recent, neglecting the natives of foreign born parents will deflate $\dot{E}_x$. This deflation decreases with age. Thus E_x is deflated and $\dot{E}_x$ is inflated, which results in a bigger adjustment term than is required for young ages.

Similarly, in a country where out migration is more dominant, E_x is inflated for young ages and $\dot{E}_x$ deflated, which results in a smaller adjustment than is required for young ages.

Actually, births to foreign born parents below certain young ages are not important in the application as they hardly affect the slope. In addition, the rate of change in E_x for the second generation is likely to be much smaller than the change for the first generation.

It should be pointed out that since the term involved in the equation is a relative rate of change $\left(\frac{\dot{E}_x}{E_x}\right)$, it may not be sensitive to errors which are not differential by age.

4.5 THE GENERAL LIKELY EFFECT OF MIGRATION

The effect of migration on the estimate of the death rate using Brass method depends on the type of the net movement, the magnitude of this movement, the age structure of migrants and the period since the movement started. An indication of the direction of this bias - according to the net movement - is given assuming that the age structure of migrants is mainly of labour force ages.

Using Graph (4.1) - corresponding to the application on Kuwait - we note that the unadjusted data have a bulge corresponding to middle ages. This bulge tends to increase the slope and results in higher estimate of the death rate. If this bulge is due to in migration - as suspected - one expects that the effect of out migration is to form a gap corresponding to middle ages, which in turn results in a lower estimate of the death rate.

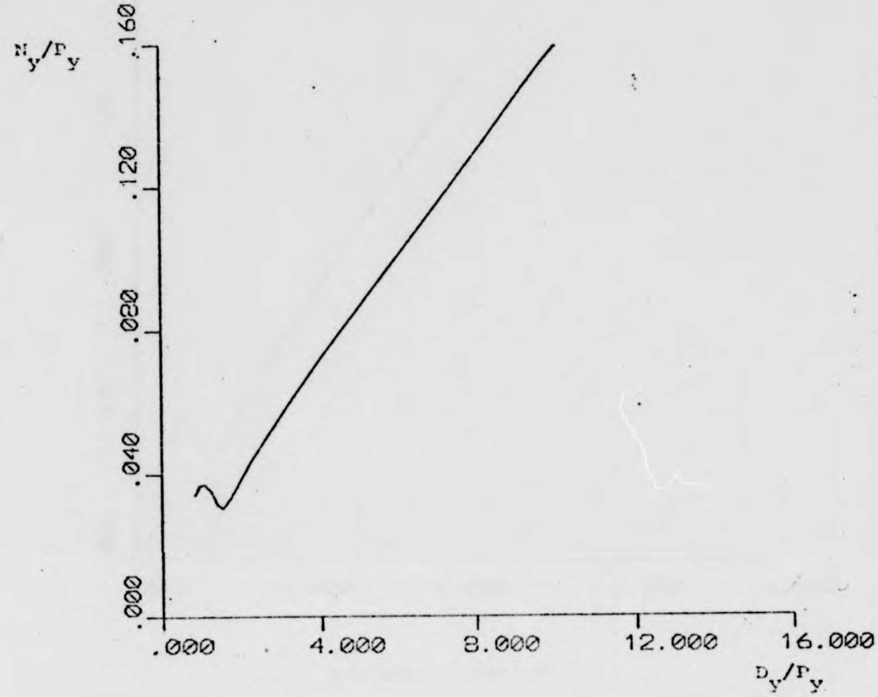
To confirm the previous observation, we projected a stable age distribution for 50 years, during which age specific mortality and fertility rates were held constant, with a continuous out-flow of migrants. For the sake of simplicity, a fixed proportion of the projected population in each age group was assumed to emigrate at the end of each projection interval, 5 year time interval in this case. These proportions are given in Table (4.3). Brass method of estimation was then applied to the projected distribution and the set of points $(\frac{N_y}{P_y}, \frac{D_y}{P_y})$ plotted.

In Table (4.4), the estimated and actual death rates are given at the end of each projection period. In Graph (4.2), the lines passing through the corresponding sets of points for each period are plotted. The estimated death rate is less than the actual rate and the plot form a gap corresponding to middle ages.

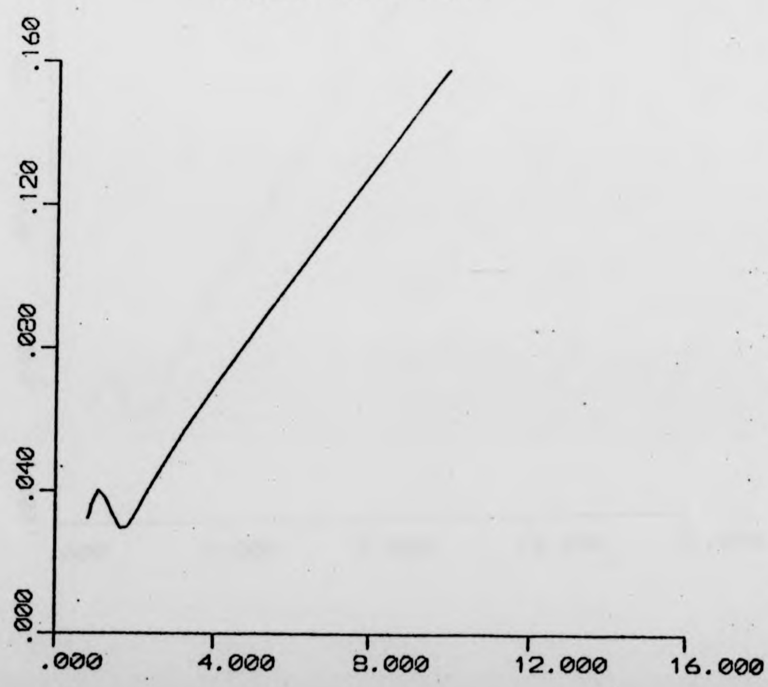
Table (4.3) The proportions assumed to emigrate at the end of each projection period

age	proportions
0-	.011
5-	.03
10-	.05
15-	.10
20-	.20
25-	.21
30-	.17
35-	.14
40-	.10

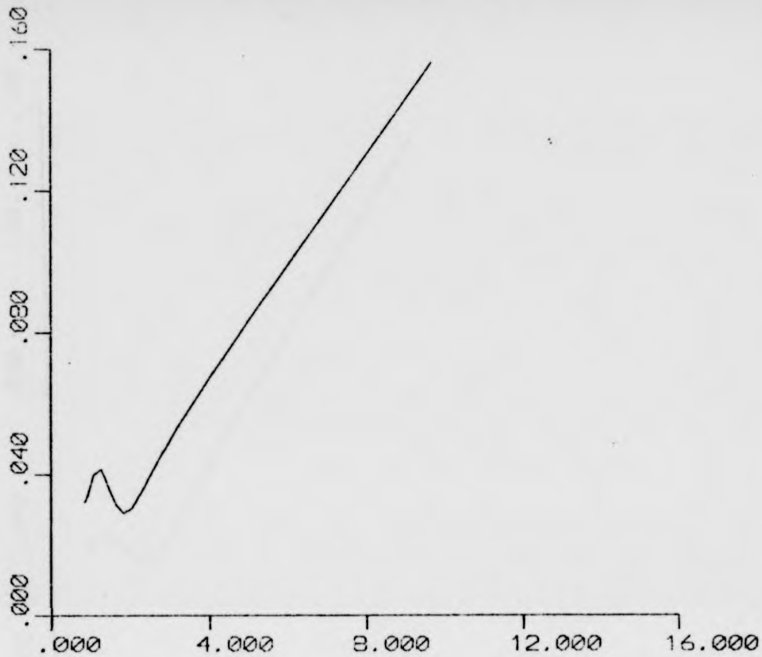
Graph (4.?) Effect of out migration.
Projection period (2)



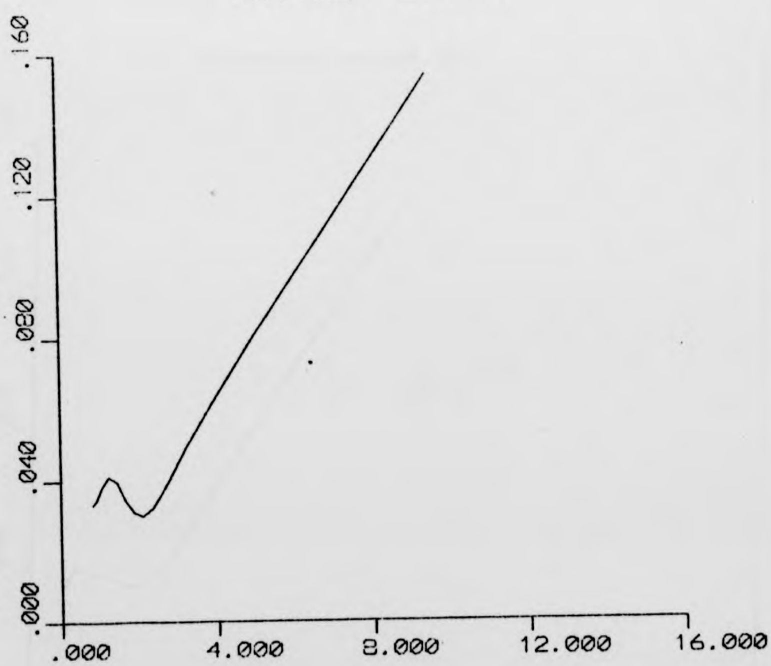
Projection period (3)



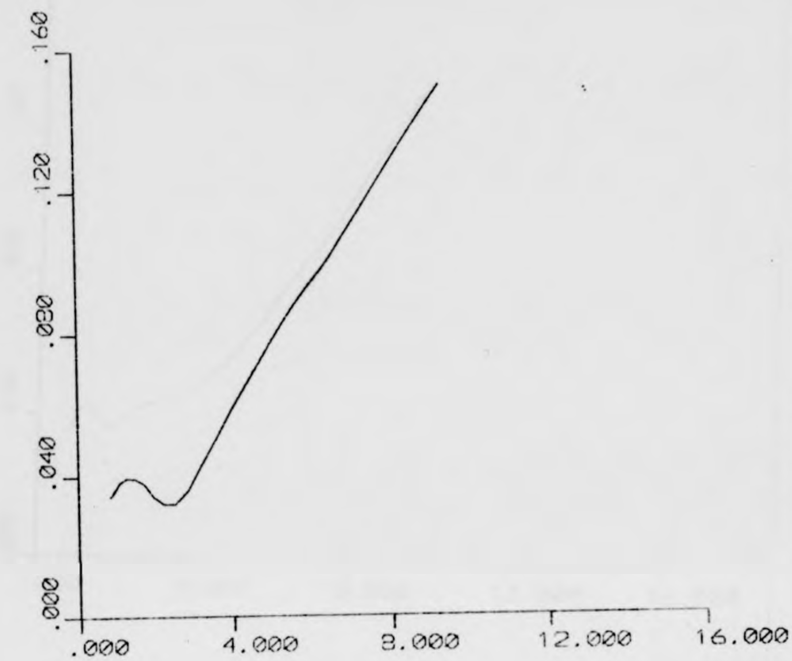
Projection period (4)



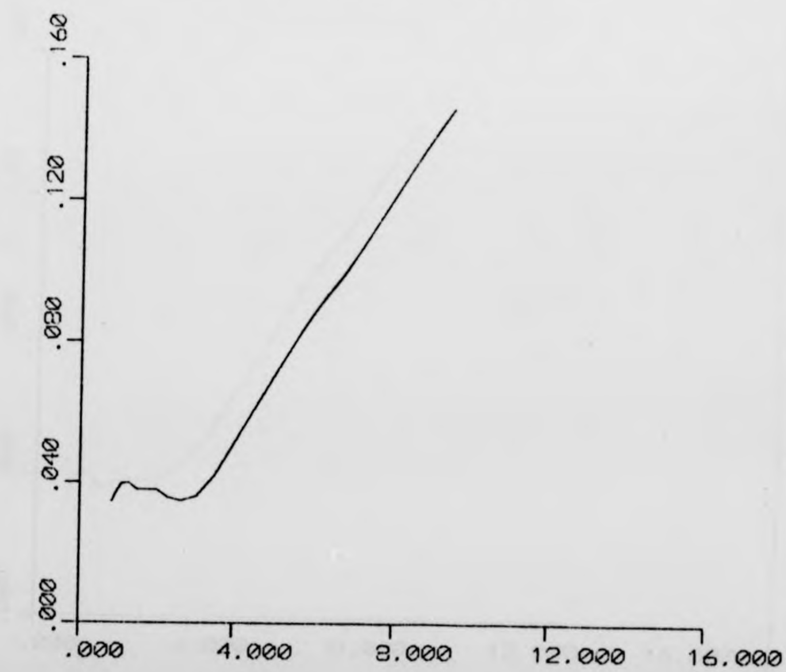
Projection period (5)



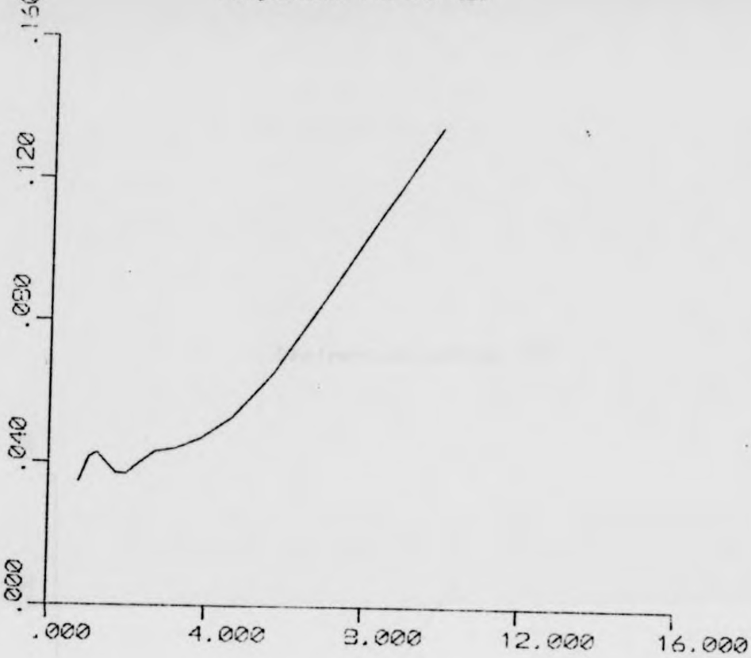
Projection period (6)



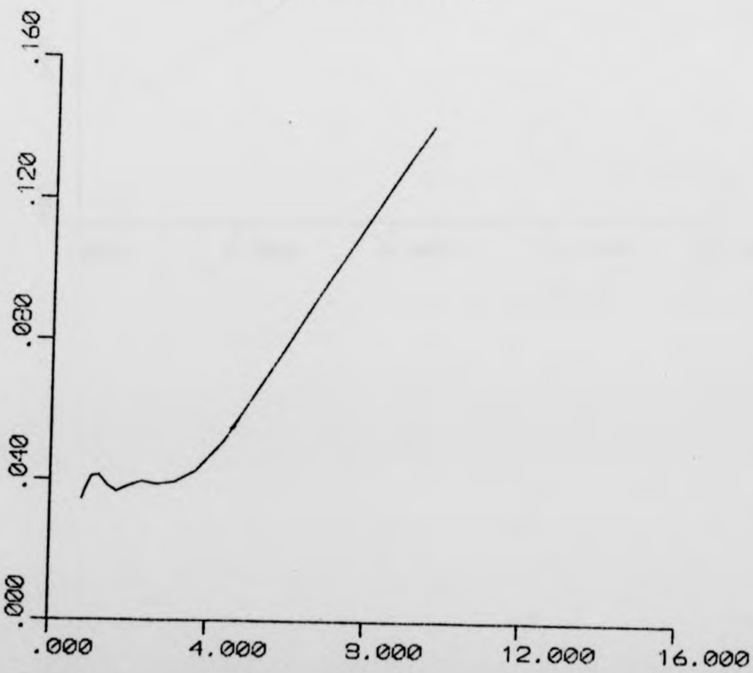
Projection period (7)



Projection period (8)



Projection period (9)



Projection period (10)

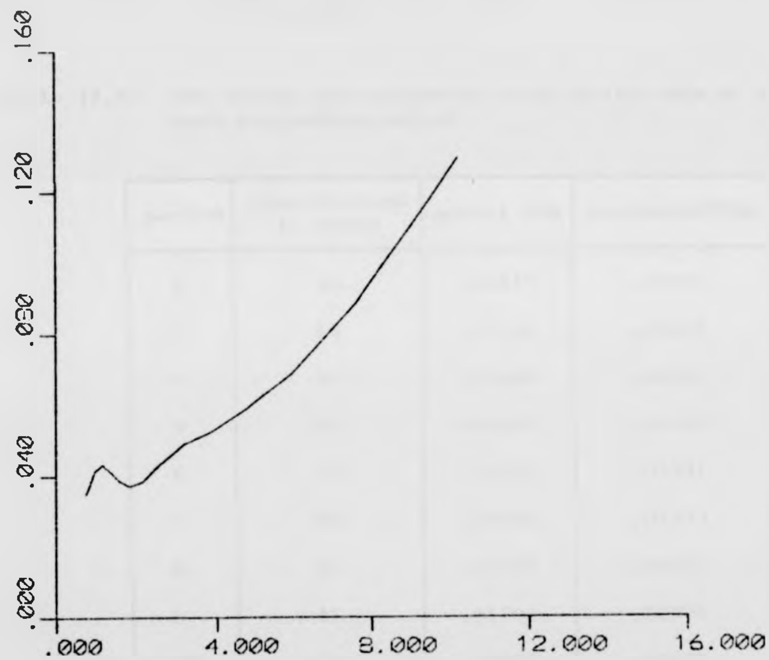


Table (4.3), continued.

age	proportions
45-	.08
50-	.05
55-	.03
60-	.02
65-	.01
70+	0.0

Table (4.4) The actual and estimated crude death rate at the end of each projection period

period	time elapsed in years	actual CDR	estimated*CDR
2	10	.01417	.01401
3	15	.01444	.01412
4	20	.01485	.01422
5	25	.01525	.01403
6	30	.01550	.01331
7	35	.01550	.01223
8	40	.01532	.01095
9	45	.01504	.00974
10	50	.01466	.00895

*The method of fit used is the least square.

For further illustration of the effect of out migration, the actual data for Puerto Rico (1960) given in Keyfitz & Flieger (1968) is used. The reason for choosing Puerto Rico is that the age composition of its population has been significantly affected by mass emigration to the United States; for example while 617,056 persons living in the United

States in 1960 were born in Puerto Rico, only 50,910 persons living in Puerto Rico were born in the United States.

The recorded crude birth rates for Puerto Rico during 1900 to 1940 show an increasing trend but this - as pointed out by Vazquez, J.L. (1968) - is only due to improvements in birth registration. This is illustrated in Table (4.5).

Table (4.5)* Reported and corrected birth rates for Puerto Rico 1900-1940

period	reported birth rate	corrected birth rate[1]
1900-1909	31.1	47.1
1910-1919	36.4	46.1
1920-1929	37.3	44.9
1930-1939	38.8	44.6

*source: reproduced from Vazquez, J.L. (1968), table 2.

[1] reported birth rates corrected for underregistration.

Very little change occurred in the age composition of the population of Puerto Rico during this period, so it may be concluded that the fertility rates remained more or less stationary during 1900-1940.

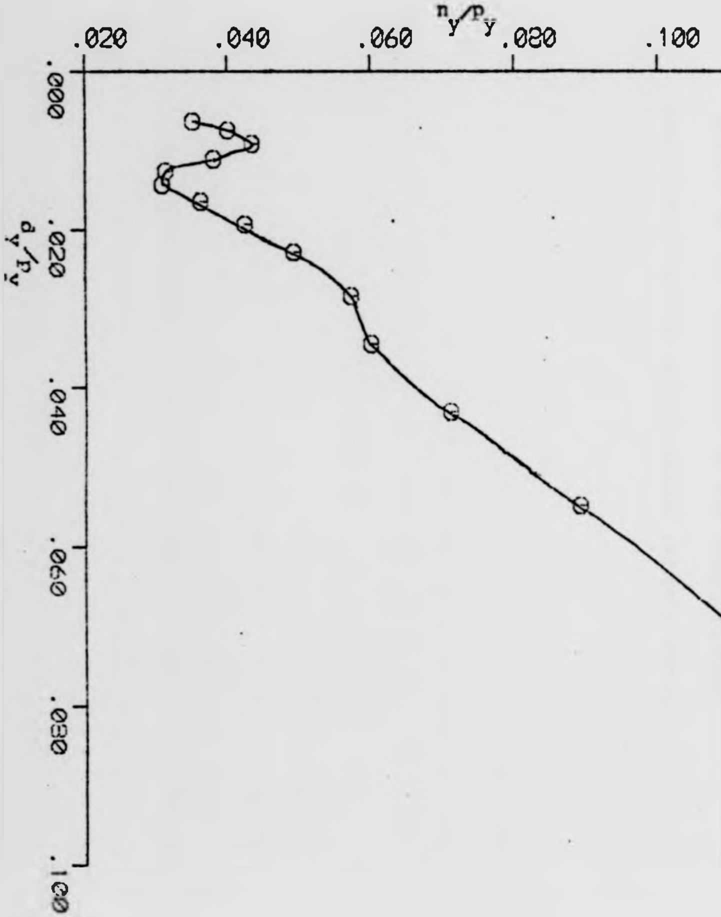
After 1940, more significant changes occurred in the birth rates as shown in Table (4.6). The birth rates declined more than 25% in 20 years. Of course, the age composition of the population of Puerto Rico has been strongly affected by emigration to the United States; the 1960 enumerated population was 30% less than the expected population in the absence of migration. When changes in age structure are taken into account, the age adjusted birth rates - the 1960 population is used as standard - still show the same picture of declining fertility - though to a lesser extent - as illustrated in Table (4.6).

Table (4.6)* Crude birth rates and age standardized birth rates Puerto Rico 1940-1960

year	crude birth rates	age standardized birth rates
1940	44.8	40.0
1950	40.1	37.0
1960	33.5	33.5

*source: reproduced, Vazquez, J.L. (1968), table 9.

Thus, Puerto Rico deviates from stability. Since the decline in fertility is recent and the emigration so strong, we expect that thought the sets of points $(\frac{n_y}{p_y}, \frac{d_y}{p_y})$ may deviate from linearity - the gap corresponding to middle ages will still be noticeable. Graph (4.3) illustrates this gap.



n_y/P_y
.020
.040
.060
.080
.100
d_y/P_y
.000
.020
.040
.060
.080
.100

GRAPH(4 . 3)

Puerto Rico, males 1960

CHAPTER (V)

EFFECT OF THE INEQUALITY OF THE PROPORTIONATE UNDER REGISTRATION ON THE GROWTH BALANCE METHOD

5.1 INTRODUCTION

The growth balance method of estimation requires that the proportionate under-registration of deaths is equal at all ages. This assumption is more likely to apply over the middle age range than for very young ages. Thus, this method is in practice used for estimating mortality of adult ages only.

It is our purpose in this chapter to extend the method to cover cases when there are two different proportionate underregistration. This is ideally suitable to allow for the different underregistration of young ages since as pointed out by Carrier (1958): 'a substantial proportion of infants die shortly after birth. For a variety of reasons and in a variety of ways this may lead to a proportion of infant deaths being treated differently from deaths at older ages, both as regards disposal of the remains and recording the event. Thus data which give adequate presentation to deaths at older ages, or at least equal deficiencies at all these ages, are liable to suffer from excessive deficiencies in infant deaths'.

In principle, of course, the extension of the method may apply to other cases, such as the differential underregistration of old age deaths. The proportionate underregistration of old ages is less or more than the general underregistration according to the significance and role of the older generation in different cultures.

In the following sections we will show that the difference in under-registration may be fully accounted for once the age groups suffering unequal under-report are located. Several numerical applications are illustrated. The effect on the graph due to the inequality of under-registration is also discussed; this may serve in locating the age groups suffering from different registration. A discussion of the advantages and disadvantages of the method is presented. Finally, illustrations of the magnitude of the error in the estimate - due to differential underregistration - according to several combinations of underregistration and shapes of age distribution are given.

5.2 A METHOD FOR ESTIMATING THE ACTUAL DEATH RATE WHEN THE PROPORTIONATE UNDER-REPORT IS NOT EQUAL

The general case when the first m age groups suffer from proportionate underregistration αu while age groups from m to M suffer from underregistration u is treated here.

In case α > 1, underregistration for young age groups 1 to m is higher than for age groups m to M. If α < 1 the opposite occurs.

- The first step is to calculate u:

using the reported number of deaths and population for ages over m and the relation:

$$\frac{n_y}{p_y} = r + \left(\frac{1}{1-u}\right) \frac{d^r_y}{p_y} \qquad y > m$$

or,

using the reported proportions of deaths and population for ages over m and the relations

$$\frac{N_y}{P_y} = r + CDR^{*} . \frac{D^r_y}{P_y} \qquad y > m$$

$$u = 1 - \frac{\text{total reported deaths}}{\text{total population } (CDR^{*})}$$

where,

n_y, p_y, r, N_y and P_y are as defined before. d^r_y and D^r_y denote the number and proportion of reported deaths over age y respectively.

- The second step is to estimate o using the following relations:

$$v_y = \frac{\frac{N_y}{P_y} - r}{CDR^*} P_y - D_y^r \qquad y < m$$

$$o = \frac{v_y}{u} + (D_y^r - D_m^r) \frac{1}{(D_y^r - D_m^r + v_y)} \qquad y < m$$

- Finally the actual death rate is equal to:

$$CDR = \left(\frac{\text{Reported deaths from 1 to m}}{1-ou} + \frac{\text{Reported deaths from m to M}}{1-u}\right)/\text{total population}$$

or,

$$CDR = CDR^*/K(u,o)$$

where,

$$K(u,\sigma) = 1 - \frac{u(\sigma-1)}{(1-u) + (1-\sigma u)\dfrac{D_m^r}{1-D_m^r}}$$

Note that when $o = 1$, there is no differential under registration. Then $K(u,o) = 1$ and $CDR = CDR^*$.

The proof of this method is given in detail in Appendix (A).

5.3 NUMERICAL APPLICATIONS

5.3.1 Application (1)

Starting with a stable distribution, model north, mortality level 11, r = 10.0 corresponding to actual death rate = 22.26 given in Coale & Demeny (1966). Subjecting the deaths corresponding to age groups from

0 to 20 to under-report 0.3, while the deaths corresponding to ages from 20 to 80+ are subjected to under-report 0.1 (o = 3, u = .1). Assuming the total population 100,000 and the total number of actual deaths 2,226, the actual and reported number of deaths and population is presented in Table (5.1).

The detailed calculations for estimating u are given in Table (5.2).

using least square fit, $\frac{1}{1-u} = \frac{\Sigma XY - \bar{X}\Sigma Y}{\Sigma X^2 - n\bar{X}^2}$

where $X = \frac{d_y^r}{p_y}$, $Y = \frac{n_y}{p_y}$.

$\frac{1}{1-u}$ = 1.120, then u = .107.

To estimate v_i, we need to calculate r and CDR`. r is the intercept of the straight line whether using proportion or numbers.

$r = \bar{Y} - 1.120, \bar{X} = .009 \simeq .01.$

Table (5.1) The actual and reported number of population and deaths in case of differential underregistration of deaths

age	Actual data		Reported data deaths
	population	deaths	
0-	2880	487.27	341.08
1-	9870	307.63	215.34
5-	10950	109.29	76.50
10-	10030	51.86	36.30
15-	9280	54.98	38.48
20-	8520	71.89	64.70
25-	7760	69.00	62.10
30-	7050	66.78	60.10

age	Actual data		Reported data deaths
	population	deaths	
35-	6370	68.78	61.90
40-	5710	74.34	66.91
45-	5060	79.69	71.72
50-	4400	88.14	79.33
55-	3740	97.72	87.94
60-	3040	111.52	100.37
65-	2320	123.98	111.58
70-	1590	130.44	117.39
75-	920	113.30	101.97
80+	510	119.53	107.58
Total	100000	2226.14	1801.29

Table (5.2) The details of calculating u

age y	number around age y (n_y)	pop. beyond age y (p_y)	reported deaths beyond age y (d^r_y)	X d^r_y/p_y	Y $\frac{n_y}{p_y}$
20	1780	56990	1093.377	.0191	.031
25	1628	48470	1028.708	.0212	.033
30	1481	40710	966.741	.0237	.036
35	1342	33660	906.647	.0269	.039
40	1208	27290	844.732	.0309	.044
45	1077	21580	777.849	.0360	.049
50	946	16520	706.137	.0427	.057
55	814	12120	626.805	.0517	.067
60	678	8380	538.862	.0643	.080
65	536	5340	438.508	.0821	.100
70	391	3020	326.931	.1082	.129
75	251	1430	209.536	.1465	.175

$$CDR^* = \frac{\text{Total reported deaths}}{\text{Total population}} \left(\frac{1}{1-u}\right)$$

$$= \frac{1801\ (1.12)}{100,000} = .020$$

The detailed calculations of v_i using ages less than 20 are given in Table (5.3).

Table (5.3) The detailed calculation of v_i

age	Y_i	P_i	D_i^r	$\frac{Y_i - r}{CDR^*} . P_i$ $r = .01,\ CDR^* = .02$	$v_i = \frac{Y_i - r}{CDR^*} . P_i - D_i^r$
5	.0265	87.25	69.077	71.981	2.904
10	.0274	76.3	64.837	66.381	1.544
15	.0291	66.27	62.827	63.287	.46

To calculate o, the following relation is used:

$$o = \left\{\frac{v_i}{u} + (D_i^r - D_m^r)\right\} \frac{1}{(v_i + D_i^r - D_m^r)}$$

$$D_m^r = D_{20}^r = 60.697$$

o (using v_1 and D_1^r corresponding to age 5) $= \frac{29.04 + 8.38}{2.904 + 8.38}$

$= 3.31$

o (using v_2 and D_2^r corresponding to age 10) $= 3.44$

o (using v_3 and D_3^r corresponding to age 15) $= 2.60$

The mean of the previous values is used as an estimate for o = 3.11

$$K(u,\sigma) = 1 - \frac{u(\sigma-1)}{(1-u) + (1-\sigma u)\dfrac{D_m^r}{1-D_m^r}} = .892$$

and finally,

$$\text{actual death rate} = \frac{CDR^*}{K(u,\sigma)} = 22.40$$

Thus instead of a reported death rate 18.01‰ this method results in an estimated death rate = 22.40‰ which is quite close to the actual death rate = 22.26‰

5.3.2 Numerical Application (2)

The principal application of this method is to allow for the highest under-registration of very young ages. This application illustrates this case. Starting with a stable age distribution, model west, males, level 13, R = 20.0‰ and CDR = 17.56 given in Coale & Demeny (1966). Subjecting age groups from 0-4 to under-report .3, while age 5-80+ are subjected to under-report .1 (u = .1, σ = 3). Table (5.4) illustrates the actual (stable) and reported age and death distribution.

Table (5.4)

age	Stable Dist.		Reported Dist.
	age dist.	death dist.	death dist.
0-	3.37	29.62	25.4139
1-	11.60	12.32	10.5705
5-	12.79	3.04	3.3535
10-	11.37	1.95	2.1511
15-	10.1	2.56	2.824
20-	8.89	3.21	3.541

age	Stable Dist.		Reported Dist.
	age dist.	death dist.	death dist.
25-	7.79	3.09	3.4087
30-	6.79	3.09	3.4087
35-	5.88	3.19	3.519
40-	5.04	3.41	3.7617
45-	4.27	3.60	3.9713
50-	3.55	4.02	4.4346
55-	2.86	4.35	4.7986
60-	2.20	4.82	5.3171
65-	1.59	4.96	5.4715
70-	1.03	4.75	5.2399
75-	.57	3.95	4.3574
80+	.31	4.04	4.4567

Assuming the total deaths are 1756 and total population 100,000. Then, the total reported deaths = 1432.6326 and the reported death rate = 14.326‰

Using age groups from 5 to 60, and least sqaure fit we get:

$\frac{1}{1-u} = 1.1004$ i.e. $u \simeq .1$.

$CDR^* = 15.76‰$ and $r = 20.55‰$

Since the life table survivors in single years between age 0 and 5 and the birth rate are supplied in Coale & Demeny (1966), we are in a position to estimate N_0 and N_1 more precisely than it is usually possible. The birth rate is equivalent to $N_0 = 3.756‰$. N_1 is estimated using the stable formula: N_1 = birth rate. $\exp(-r).l_1/l_0 = 3.168‰$

The detailed calculations for estimating v_i is given in Table (5.5).

Table (5.5)

age	N_y	$Y_i = \frac{N_y}{P_y}$	P_i	$\frac{Y_i - .020}{.01576}$ $P_i = D_i^c$	D_i^r	$v_i = D_i^c - D_i^r$
0	3.756	.03756	100	111.421	100	11.421
1	3.168	.03278	96.63	78.358	74.584	3.774

To calculate o:

$$o = \left(\frac{v_i}{u} + D_i^r - D_m^r\right) / \left(v_i + D_i^r - D_m^r\right)$$

$$o \text{ (using age 0)} = \frac{114.21 + 35.9844}{11.421 + 35.9844} = 3.16$$

$$o \text{ (using age 1)} = \frac{37.74 + 10.5705}{3.774 + 10.5705} = 3.36.$$

Thus, once N_y, r and u are estimated accurately, the method performs very well and a good estimate of o is available.

It is our purpose now to check the effect of deviations in the value of r and N_y on the method.

In the previous calculations the value of r was taken equal to .020 but the estimate of r was .02055. Using this estimate and repeating the calculations in Table (5.6) we get:

$v_0 = 7.93$, $v_1 = .402$

and, o (using v_0) = 2.62 and o (using v_1) = 1.32.

Thus a small change in r affects the value of σ.

In actual applications, it is unlikely the exact values of N_0 and N_1 are available. The age distribution for age group 0-1 is usually supplied. In case the age distribution for age group 1-5 is not available in single years the proportion of persons aged 1 needs to be estimated using more complicated techniques since it is known beforehand that the age distribution between 0-5 does not follow a linear decline. Suppose the proportions of persons between 0-1 and 1-2 are available, then an estimate of persons aged 1 may be taken as: $(N_{0-} + N_{1-})/2$ which is an overestimate of proportions aged 1. An illustration of the effect of overestimating N_1 on the estimate of σ in this application is given as follows:

Using the single years life table survivors $(l_1, l_2 ...)$ the proportion between age 1-2 = birth rate $e^{-r(1.5)}(l_1+l_2)/2 = 3.079$. v_1 may be calculated as:

age	N_y	P_i	$\frac{Y_i-0.020}{.01576}$ $P_i=D_i^c$	D_i^r	v_1
1	3.224	96.63	81.941	74.584	7.357

and finally σ = 4.69.

5.3.3 Numerical Application (3)

Brass (1976) applied the growth balance method to vital registration and census statistics for Iraq. It was noticed that the points at higher ages were quite close to linearity; those at younger ages were erratic and displayed a peculiar curvature upwards at the lower end of the graph. Brass suspected different underregistration of deaths at young ages (up to 30 years).

Ignoring the upturn of the lower points, the estimate of f was reached as 1.88 and used to inflate the reported deaths over the range for which the correction was taken as applicable.

To allow for differential underregistration at young ages, the same previous adjustment was extended to ages over 5. It was pointed out that since mortality over 5 was so low, little overall error was expected by this adjustment. To estimate the deaths corresponding to ages less than 5, the south set of Coale & Demeny model life table was used. Level 14 mortality was estimated to correspond to a population with the Iraq age distribution and the adjusted death rates over age 10. The adjusted crude death rate for the Iraq age distribution was then estimated as 15.5‰ Brass commented that this rate is somewhat lower than expected.

The adjustment procedure - to allow for the differential underregistration is applied using the same data for Iraq. Table (5.6) presents the original data for Iraq.

Table (5.6) Data for Iraq 1960-70, females

age group	number (thousands)	deaths (thousands)
0-4	766.7	2.13
5-9	603.0	.36
10-14	491.2	.34
15-19	343.4	.31
20-29	531.4	.74
30-39	459.2	.87
40-49	315.5	.95
50-59	227.6	1.02
60-69	155.8	1.90

age group	number (thousands)	deaths (thousands)
70-79	75.9	4.76
81 over	24.0	

* reproduced from Brass (1976), table 6.

Using the points corresponding to ages over 30:

$\frac{1}{1-u} = 1.883$ $\quad u = .4689$

$CDR^{*} = .0063$ $\quad r = .0262$

The detailed calculations for estimating o are presented in Table (5.7).

Table (5.7) The detailed calculations for estimating o

age (Y) (1)	D^r_y (2)	P_y (3)	$\frac{N_y}{P_y} - r\ /CDR^{*}$ (4)	$v_y = (4).P_y - D^r_y$	o
5	.8404	.8080	2.5873	1.2497	2.02
10	.8139	.6567	2.4603	.8017	2.00
15	.7884	.5340	2.0634	.3134	1.91
20	.7653	.4480	1.9047	.0880	1.70

Thus $o \simeq 2$.

The adjusted death rate, assuming the underregistration under age 30 is twice the underregistration over age 30 = 20‰

In view of the previous discussion and the near constancy of o, the adjusted death rate seems much more reasonable than the reported rate of 3.35‰

Note that, assuming the underregistration over age 5 is the same as the underregistration over age 30, the adjusted death rate = 14‰. This is quite close to the estimate provided by Brass.

5.4 EFFECT ON THE GRAPH DUE TO THE UNEQUALITY OF UNDERREGISTRATION

In Appendix (A) we proved that:

$$\frac{N_y}{P_y} = r + CDR^{\backprime}.\frac{D^r_y}{P_y} \qquad y > m$$

$$CDR^{\backprime} = CDR.K(u,o)$$

and

$$\frac{D^r_y}{P_y} = \frac{D_y}{P_y.K(u,o)} - v_y \qquad y < m$$

where

$$K(u,o) = 1 - \frac{u(o-1)\sum_{x=0}^{m} d_x}{(1-u)\sum_{x=0}^{M} d_x}$$

$$v_y = \frac{u(o-1)\sum_{x=y}^{m} d_x}{K(u,o)(1-u)\sum_{x=0}^{M} d_x}$$

If $\sigma > 1$ (higher under-registration of young ages),

then, $K(u,o) < 1$ and v_i = +ve.

Thus: $CDR^{\backprime} < CDR$ and

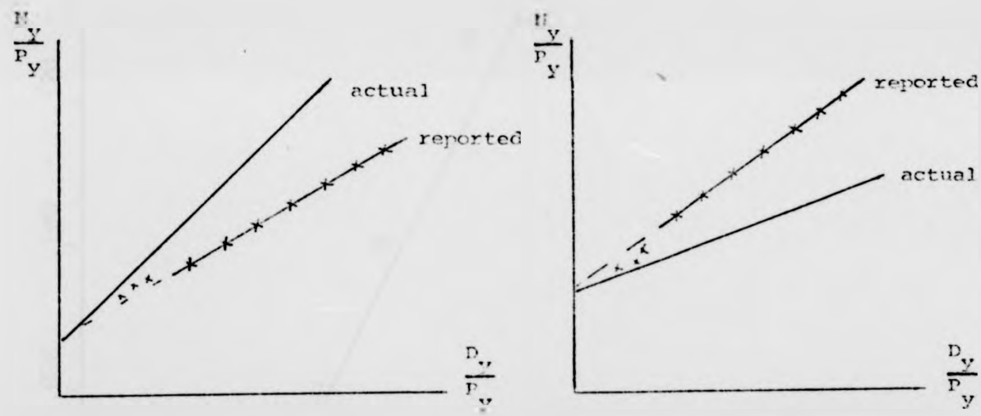
$\frac{D^r_y}{P_y}$ for $y < m$ is less than: $\dfrac{\frac{N_y}{P_y} - r}{CDR^{\backprime}} = \dfrac{D_y}{P_y.K(u,o)}$.

If $o < 1$, the opposite occurs.

Graph (5.1) illustrates the effect of this simple type of differential under-registration.

Misreporting of age, as will be shown in Chapter (6), tends to result in sets of points that deviate in both directions of the true line. This type of differential under-registration results in a line corresponding to old ages with a different slope than the true line, and a set of points corresponding to young ages deviating to only one side of this reported line.

Graph (5.1) Effect of differential under-registration



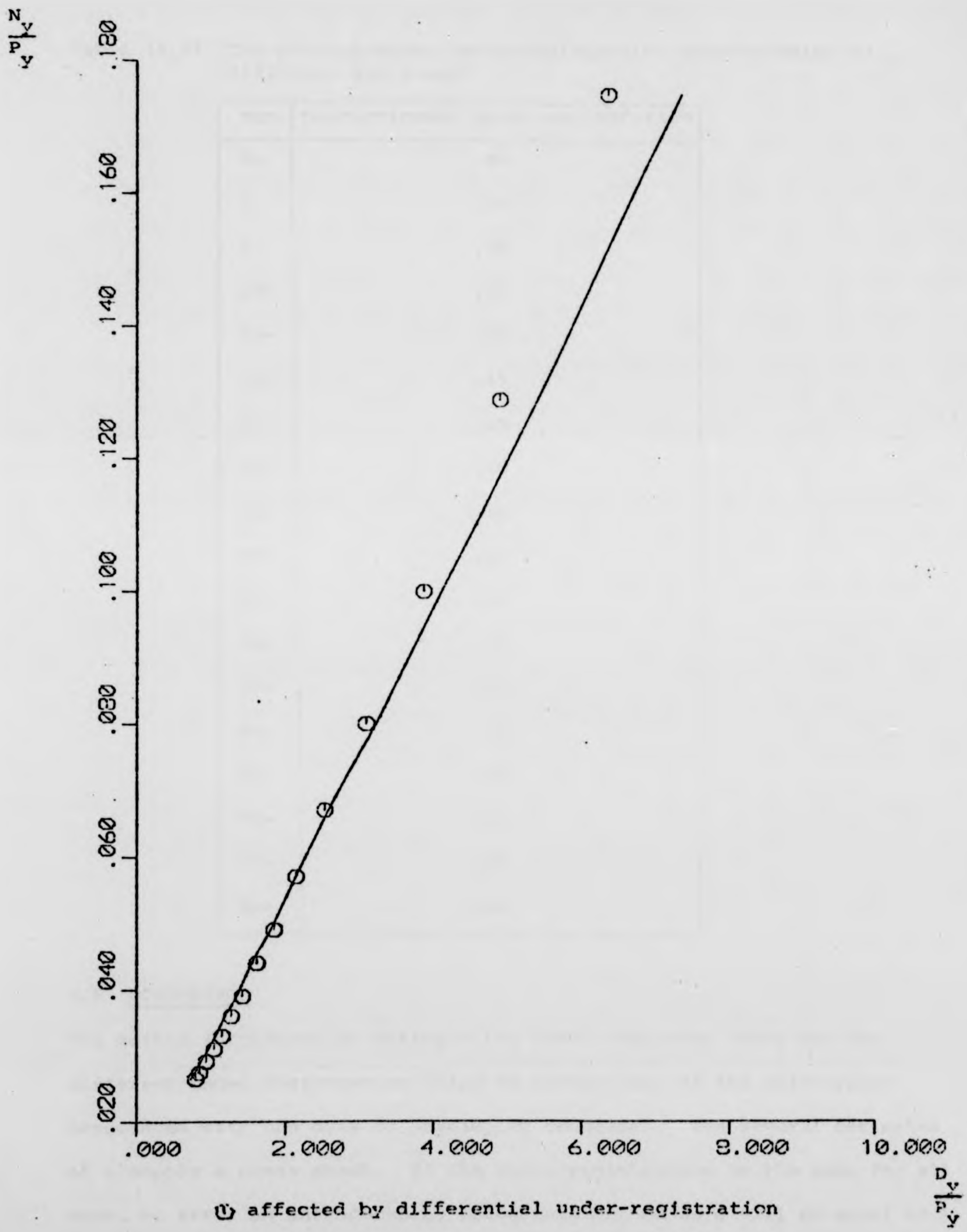
(a) higher under-registration of young ages

(b) higher under-registration of old ages

x: sets of points for $y < m$

It is of interest to illustrate the effect of a more complex type of under-registration on the graph of the sets of points $\frac{N_y}{P_y}$ and $\frac{D_y}{P_y}$. Let us consider the effect of under-registration that is high corresponding to young ages, declines till it reaches its minimum corresponding to middle ages and then increases again till it reaches the same level as young ages. More specifically, starting with the same age and death distribution utilized in application (5.3.1) and subjecting the death distribution to the proportionate under-registration given in Table (5.8), the resulting points

Graph (5.2) Effect of differential under registration
(more complex type of under-registration)



⊖ affected by differential under-registration

$\frac{N_v}{P_y}$ and $\frac{D_v^r}{P_y}$ are illustrated in graph (5.2) with the line drawn in case there was no differential under-registration.

Table (5.8) The proportionate under-registration corresponding to different age groups

age	proportionate under-registration
0-	.40
1-	.35
5-	.30
10-	.25
15-	.20
20-	.15
25-	.10
30-	.10
35-	.10
40-	.10
45-	.10
50-	.15
55-	.20
60-	.25
65-	.25
70-	.30
75-	.35
80+	.40

5.5 DISCUSSION

The method introduced to estimate the death rate when there are two different under-registration helps to correct one of the main errors associated with the data of developing countries. The several estimates of c supply a cross check. If the under-registration is the same for all ages, no error is introduced by using this method as c will be equal to 1 as a result of $v = 0$.

Also, there is no need to know exactly the ages suffering from differential under-registration. For example, if the researcher assumes a higher under-registration under age 5 while the data suffer from a higher under-registration under age 1 only. Theoretically, the value of o corresponding to age 1 should equal 1 as a result of $v_1 = 0$, the value of o corresponding to age 0 will be higher than 1 but o will be under-estimated in this case due to using the difference $(D^r_0 - D^r_5)$ instead of using $(D^r_0 - D^r_1)$ in calculating o. Using the data in numerical application (5.3.2) where actual CDR = 17.56‰ and CDR` = 15.76‰, suppose the researcher assumes that the different under-registration occurs in age group 0-1 (in fact it occurs for ages under 5), and that young under-registration is 3 times as the general under-registration, in this case K(u,o) = .932 and the estimated CDR = 16.244. The difference between 16.244 and 17.56 results from the failure to realize that the differential under-registration occurs under age 5 rather than under age 1.

The magnitude and sign of v_i indicates the degree and type of differential under-registration, in case v_i equal to zero age group i suffer from the same general under-registration, the bigger $|v_i|$ the more different the under-registration of age group i from the general under-registration. Also, if v_i is positive age group i suffers from higher under-registration while if v_i is negative age group i suffer from lower under-registration than the general under-registration.

From the previous applications, several disadvantages of the method are pointed out. First, the estimate of o is quite sensitive to the values of CDR`, r and $N_y$. A very small alteration in the values of CDR`, r or N_y may lead to a big difference in the estimate of o and consequently in the estimate of K(u,o). Considering the quality of data in developing countries it is quite unlikely that the estimate of CDR`, r and/or N_y are precise enough to yield a very accurate estimate of K(u,o) Thus, only an approximate estimate of the differential under-registration is obtained. It has been pointed out that (theoretically) the sets of points $\frac{N_y}{P_y}, \frac{D^r_y}{P_y}$ corresponding to young age groups with different under-registration deviate from the straight line and thus may help to identify these age groups. In some practical application, this identification may prove difficult.

For practical applications on actual data for developing countries the following steps are suggested to correct for higher under-registration of deaths of young ages:

- apply the growth balance method starting from age 5 on either the proportions to calculate CDR' or the numbers to calculate u. Find the missing parameter using the relation:

$$u = 1 - \frac{\text{total reported deaths}}{\text{CDR}' \cdot \text{Total population}}$$

- calculate N_0 and N_1, N_0 = birth rate while $N_1 = \frac{1}{2}$ (proportion aged (0-1) + proportion aged (1-2). If the birth rate is of doubtful accuracy we may only depend on N_1. Calculate v_0 and v_1; if $-2 \leq v_i \leq 2$ for $i = 0, 1$, then v_i is almost zero and we may conclude that there is no differential under-registration at young ages. The range -2 to 2 is allowed to cover for the deficiencies in the data resulting from age errors, deviation from stability ... etc.
- assume a differential under-registration under age 5 and calculate o using the relation:

$$o = \{\frac{v_i}{u} + (D^r_i - D^r_5)\}/(v_i + (D^r_i - D^r_5)). \quad i = 1.$$

Actually, two types of errors are probably present, first N_1 is likely to be overestimated resulting in an over-estimate of v_1 and consequently of o; on the other hand the higher under-registration is probably for ages younger than 5 and $(D^r_i - D^r_m)$ is overestimated resulting in an underestimate of o. Thus the two errors are different in directions and may offset each other.

5.6 MAGNITUDE OF THE ERROR DUE TO DIFFERENTIAL UNDER-REGISTRATION

The value of K(u,o) is not only affected by the magnitude of o and u but also by the slope of the death distribution. To get an idea of the magnitude of the effect of differential under-registration for young ages on the estimate of the crude death rate, values of K(u,o) corresponding to different values of o and u and different stable age distributions are given in Table (5.9).

The stable distributions are Coale & Demeny (1966) stable distributions, model west, r = 15, makes level 9 and 15 corresponding to CDR 25.48 and 15.43 and e_0 37.301 and 51.831.

Table (5.9) Values of K(u,o) corresponding to specified values of o and u level (9) differential under-registration from age 0-1

u \ o	1	2	3	4	5	6	7	8	9
.1	1	.963	.927	.890	.854	.817	.781	.745	.708
.2	1	.918	.836	.754					
.3	1	.859	.719						
.4	1	.781							
.5	1								
⋮	⋮								
.9	1								

differential under-registration from 0-5

u \ o	1	2	3	4	5	6	7	8	9
.1	1	.947	.895	.843	.790	.738	.686	.634	.581
.2	1	.882	.764	.647					
.3	1	.798	.596						
.4	1	.686							
.5	1								
⋮	⋮								
.9	1								

level (15) differential under-registration from age 0-1

u \ o	1	2	3	4	5	6	7	8	9
.1	1	.975	.951	.927	.903	.878	.854	.830	.806
.2	1	.945	.890	.836					
.3	1	.906	.812						
.4	1	.854							
.5	1								
⋮	⋮								
.9	1								

differential under-registration from age 0-5

u \ σ	1	2	3	4	5	6	7	8	9
.1	1	.966	.932	.898	.864	.831	.797	.763	.729
.2	1	.924	.848	.772					
.3	1	.869	.739						
.4	1	.797							
.5	1								
⋮	⋮								
.9	1								

CHAPTER VI

EFFECT OF AGE MISREPORT ON THE GROWTH BALANCE METHOD

6.1 INTRODUCTION

Age data of developing countries are greatly distorted. They suffer from traditional sources of error such as heaping, rounding and vagueness, also from errors relating to the specific population cultures. Responses regarding certain ages are affected by the social prestige accorded to that age or by laws and practises such as age for school attendance, voting, military service and marriage.

Age errors are not always easy to detect and may be difficult to measure. One source of error in the developing countries is simply ignorance of age; this makes the problem of detecting likely age errors even more difficult. By graphing the data and comparing with mathematically smoothed series or other accurate data or stable models, one may be able to identify certain patterns of age misreporting. The main danger in applying this procedure is in imposing an unrealistic model on the data and thus mistaking inherent features as errors.

Several methods for estimating demographic measures for developing countries depend on the relation between two age distributions. Age misreport distorts this relation and introduces a bias in the estimate. It is valuable if an indication of the magnitude of this bias can be presented.

Special attention is directed to the effect of age misreporting on the growth balance method of estimation. An advantage of the method is that it is only affected by the net transfer from one age group to the other so we may disregard the identity of individuals and allow for the offsetting effects of reporting into and out of a given age. On the other hand, a disadvantage may be due to the possibility of different errors associated with the statement of age for the deceased and the living which complicates the analysis of the bias introduced by age misreport.

Any procedure for calculating the range of this bias makes use of a model of age error; it is our purpose to introduce a general model for age reporting. The range of likely bias introduced in Brass estimate is shown using simulation procedure, which allows for a complicated and realistic model of age reporting. Finally, the important question regarding the effect of graduating the data before applying Brass method is dealt with.

6.2 A MODEL FOR AGE REPORTING

The two main reasons for age misreport in developing countries are ignorance of age and/or bias associated with this age. The persons aged x may be divided into two classes; the first includes everyone who knows his age correctly while the second includes those ignorant of their age. The first class may be subdivided to a_{1x} and a_{2x}; where a_{1x} includes those knowing their age and not biased in their reporting of this age, and a_{2x} those knowing their age and biased in their report. Similarly, the second class is divided to a_{3x} and a_{4x}; where a_{3x} includes those not knowing their age and not biased, and finally a_{4x} includes those not knowing their age but biased.

The model for age reporting may be given by:

$Y_x = x$ in group a_{1x},

$Y_x = x + BI_x$ in group a_{2x},

$Y_x = x + er_x$ in group a_{3x},

$Y_x = x + er_x + BI_{x+er_x}$ in group a_{4x}

where, Y_x: reported age when the true age is x.

BI_x: bias associated with age x.

er_x: random error associated with age x.

In other words, if a person knows his age and is not biased against this age or towards a neighbouring age he will state his age correctly. If he is biased the reported age depends on the kind of bias prevailing. If he

does not know his age but is not biased, he will attempt to state his age correctly, the deviation between the reported and actual age is simply a random error. Finally, if a person does not know his age but is biased against or towards a certain age - which is usually in the neighbourhood of his actual age - he will either avoid or report this age as his actual age.

Instead of dealing with exact age x we will consider single years age group x, where x denotes the age between x - ½ and x + ½.

For a full description of this model, the distribution of a_{jx}, er_x and BI_x has to be specified. A discussion of the general characteristics of these distributions follows.

6.2.1 The Distribution of the Different Groups

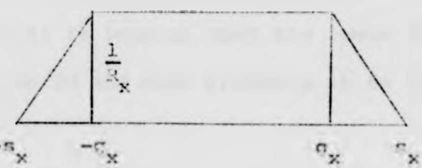
a_{1x} and a_{2x} include all those who know their age. The younger the age the closer the incident of birth and the more likely the age is known; it is expected that the probability of being in group a_{1x} and a_{2x} is a decreasing function of age. For older ages, a_{1x} and a_{2x} may be related to the educated percentages in age group x, or any other indicator of this education available.

Biases against or towards certain ages prevail in different communities according to their specific cultures and social customs. It is more likely that a smaller percentage of the educated are affected by these customs. In other words, the proportion in a_{2x} constitute a smaller percentage of group a_{1x} and a_{2x} than does the proportion in a_{4x} with respect to group a_{3x} and a_{4x}.

Apart from the previous guidelines, the assignment of values for the probabilities of being in the different groups is more or less arbitrary.

6.2.2 The Distribution of the Random Error

The general form of the distribution may be given by



$$f(er_x) = \frac{1}{R_x}\frac{(er_x + s_x)}{(s_x - g_x)} \qquad -s_x < er_x < -g_x$$

$$= \frac{1}{R_x} \qquad -g_x < er_x < g_x$$

$$= \frac{1}{R_x}\frac{(s_x - er_x)}{(s_x - g_x)} \qquad g_x < er_x < s_x$$

where $R_x = s_x + g_x$.

$$E(er_x) = 0 \quad \text{and} \quad Var(er_x) = \frac{g_x^2 + s_x^2}{6}.$$

The parameters of the distribution (g_x, s_x) may be set arbitrary; but they need to satisfy the following requirements to be realistic:

- though a person may not know his age, there is an upper limit for the s_x imposed by several factors, such as: appearance, social status, type of job... etc. For example, a person aged 40 is unlikely to state his age as 10. It is more likely that s_x, for x = 40 ranges between 5 and 15.
- The older the person the higher is the upper limit of his deviation (the higher the values of s_x),

$$s_x \geq s_z \qquad x \geq z.$$

Though $s_x \geq s_z \qquad x \geq z$

$$x - s_x \geq z - s_z$$

and $x + s_x \geq z + s_z$.

for example, if the lower limit for a person aged 40 is to state his age as 20 (s_{40} = 20); it is logical that the lower limit for a person aged 50 can not be less than 20 and most probably it is higher than 20.

6.2.3 Distribution of BI_x

Though there are several types of bias prevailing in developing countries, such as: digit preference, concentration of women in the middle of the reproduction period, overstatement of age for old people... etc.; these biases are basically the same. They show attraction to some ages and avoidance of others. One type of bias is illustrated here, others will be discussed later.

Consider two ages x and z such that x is an age where there is a bias against and z a bias towards. Persons aged x may either increase or decrease their age by 1 to a years, if this is done uniformally (other patterns of change may be assumed) a possible model may be:

$$f(BI_x) = \frac{1}{2(a-1)} \qquad -a < BI_x < -1$$

$$1 < BI_x < a$$

$$f(BI_x) = 0 \qquad \text{otherwise.}$$

Persons aged around z will report their age as z, thus:

$$P_r\{BI_w = z - w\} = 1 \qquad z - a < w < z + a$$

6.3 ESTIMATING THE MAGNITUDE OF ERROR IN THE GROWTH BALANCE ESTIMATE USING SIMULATION PROCEDURE

To estimate the magnitude of error in the growth balance estimate due to misreport of age, the general model presented in section 2 of this chapter is used. Two cases are considered, the first when both the age and death distribution are subjected to the same type of error; the second when the error in the death distribution is different from the error in the population age distribution.

For each case, we will discuss the values assigned to the parameters of the error distributions, the procedure used in simulating the reported age distribution, the results of several computer applications on different age distributions and the likely effect of age error on Brass estimate of the death rate, given the pattern of age error considered.

6.3.1 <u>The Same Kind of Error in the Population and Death Distribution</u>

6.3.1.1 <u>Values assigned to the parameters of the error distribution</u>

- The probabilities of being in different groups:

the values of the different probabilities are set arbitrary as follows:

$p(a_{1x} + a_{2x}) = 70\%$ $\quad x < 5$

$\quad = 50\%$ $\quad x > 5$

$p(a_{2x}) = 30\%\ p(a_{1x} + a_{2x})$

$p(a_{3x} + a_{4x}) = 30\%$ $\quad x < 5$

$\quad = 50\%$ $\quad x > 5$

$p(a_{4x}) = 40\%\ p(a_{3x} + a_{4x})$

Then:

$p(a_{1x}) = .49$

$p(a_{2x}) = .21$ $\quad$ for $x < 5$

$p(a_{3x}) = .18$

$p(a_{4x}) = .12$

$p(a_{1x}) = .35$

$p(a_{2x}) = .15$ for $x > 5$

$p(a_{3x}) = .30$

$p(a_{4x}) = .20$

Thus for ages over 5 it is assumed that 65% of the populations are influenced by some kind of error in reporting their ages.

- Bias error:

Types of bias studied under this model are twofold. The first, generally described as digit preference, shows itself as heaping on digits terminating with: 0, 5, 8, 2, 6 and 4; which of course imply shunning from ages terminating with 3, 7, 1 and 9. The second bias that characterizes most developing countries is a general movement on the age scale; we will consider the movement from age 11-19 to ages 20-29 (this movement is clear in female age distributions for African societies) and the movement from ages 51-59 to ages 60-69.

Digit preference:

If x is a preferred end digit, persons whose age ends with x states it correctly, unless they are affected by another error. Persons whose age ends with a digit different from x, states their age correctly or ending with another digit according to the following probabilities.

Movement out of age 1 and 9 are stronger than movement out of 3 and 7 as they are close to one of the most preferred end digits.

If f(x/y) denotes the probability of moving from an age ending in y to the closest age ending in x, the different probabilities may be given as:

f(0/1) = .55

f(1/1) = .20

f(2/1) = .25

f(2/3) = .25

f(3/3) = .35

f(4/3) = .25

f(5/3) = .15

f(6/7) = .25

f(7/7) = .35

f(8/7) = .25

f(5/7) = .30

f(8/9) = .25

f(9/9) = .20

f(0/9) = .55

General movement on the age scale

General movement on the age scale:

A person aged between 11-19 or 51-59 affected by this bias will move up the age scale from 1 to 10 years uniformally, thus:

$$p(BI_x = y) = \frac{1}{9} \qquad 1 < y < 10 \text{ and } 11 < x < 19.$$

also,

$$p(BI_x = y) = \frac{1}{9} \qquad 1 < y < 10 \text{ and } 51 < x < 59.$$

where $p(BI_x = y)$ denotes the probability a person aged x will add y years to his age.

Finally, a person aged 11-19 or 51-59 is subjected to either of the previous biases (digit preference, movement up the age scale) with equal probability. Thus, a random number decides first which type of bias a person is subjected to and another number reflects the value of this bias.

- Random error:

The same random error distribution introduced in section (2) is used here except for age 0-1, thus:

$$f(er_x) = \frac{1}{R_x}\,\frac{er_x + s_x)}{(s_x - g_x)} \qquad -s_x < er_x < -g_x$$

$$= \frac{1}{R_x} \qquad -g_x < er_x < g_x \qquad \text{for all } x > 1$$

$$= \frac{1}{R_x}\,\frac{(b_x - e_x)}{(b_x - a_x)} \qquad g_x < er_x < s_x$$

$$R_x = s_x + g_x$$

for $o < x < 1$

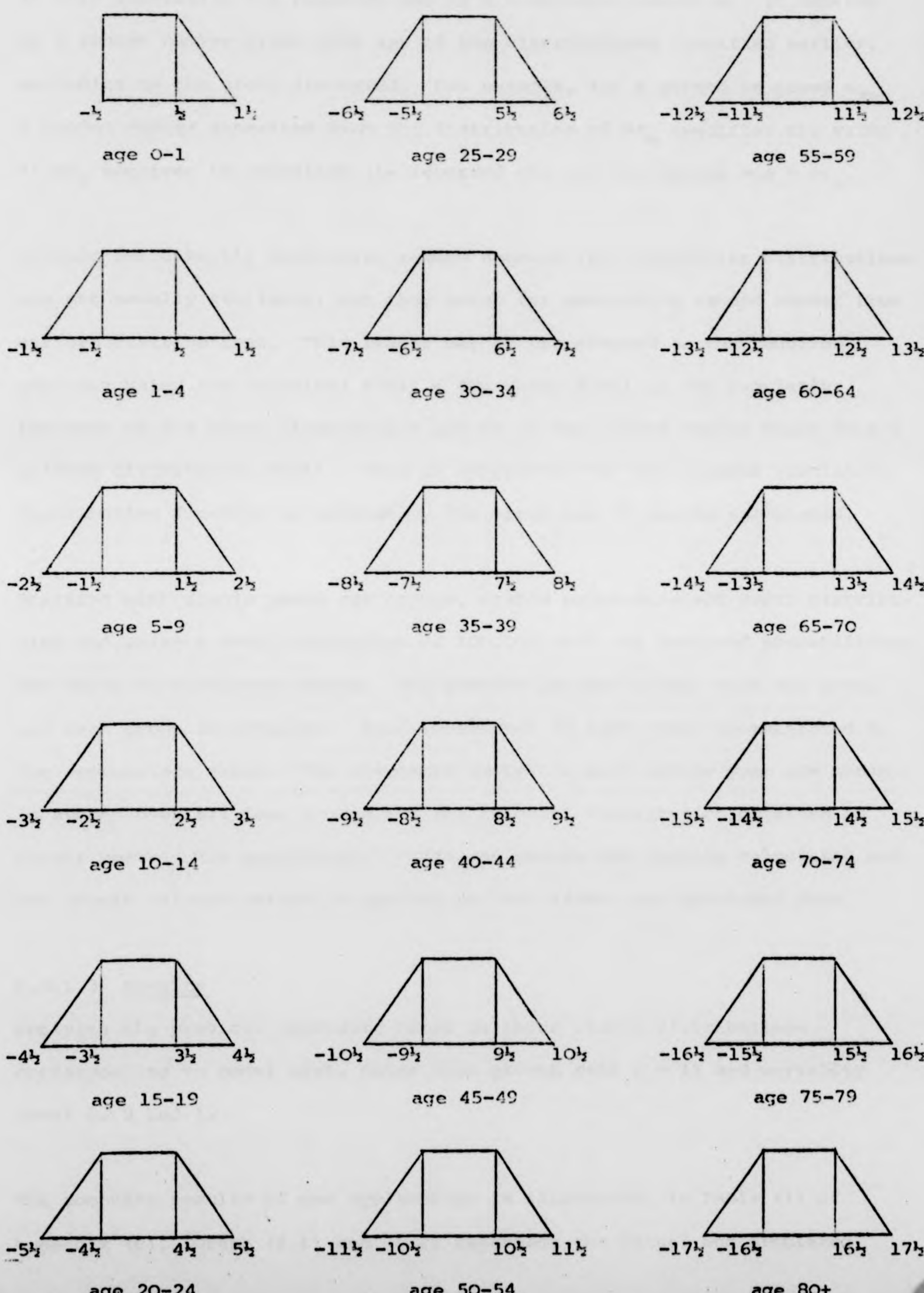
$$f(er_x) = \frac{1}{R_x} \qquad -g_x < er_x < g_x$$

$$= \frac{(s_x - er_x)}{R_x(s_x - g_x)} \qquad g_x < er_x < s_x$$

$$R_x = \frac{s_x + 3g_x}{2}$$

The values of the different parameters (g_x, s_x) are illustrated in the following graph.

Graph (6.D) Distribution of er_x for different x



6.3.1.2 <u>The procedure used in simulating the reported age distribution</u>

In this simulation the reported age is a stochastic variable. It depends on a random number drawn from any of the distributions specified earlier, according to the group discussed. For example, for a person in group a_{3x}, a random number generated from the distribution of er_x specifies the value of er_x required to calculate the reported age as: the actual age + er_x.

Methods for directly generating random numbers from particular distributions are not usually available, but they exist for generating random number from uniform distributions. This number may be transformed to the required sequence using the relation: F(er) = RN, where F(er) is the cumulative function of the error distribution and RN is the random number drawn from a uniform distribution (0,1). Once an expression for the inverse cumulative distribution function is available, the error may be easily calculated.

Starting with single years age groups, stable population and death distribution and using a total population of 100,000 with the assigned probabilities for being in different groups. The numbers in each single year age group and each group is obtained. Each individual in each group is subjected to the appropriate error. The simulated number in each single year age group is summed over all four groups and the reported distribution obtained in single years. The equivalent 5 years age groups are readily calculated and the growth balance method is applied to both stable and simulated data.

6.3.1.3 <u>Results</u>

Applying the previous procedure twice on three stable distributions corresponding to model west, males with growth rate r = 15 and mortality level 6, 9 and 12.

The computer results of one application is illustrated in Table (1) of Appendix (b). Graph (6.1) and (6.2) represent the actual and simulated age distributions in single years and 5 years age group respectively.

Graph (6.3) represents the average percentage female age distribution of 30 sets of census or survey data of various African countries and the stable model fitted to this average. This data are extracted from a study of the United Nations on age error in African data. (United Nations, 1975).

The similarity between the characteristics of age mis-statements in African countries and in the simulated data is apparent.

Finally, the actual death rate calculated by dividing the total deaths over the total population - and the estimated death rate - calculated by applying the growth balance method on the simulated age distributions - are given in Table (6.2).

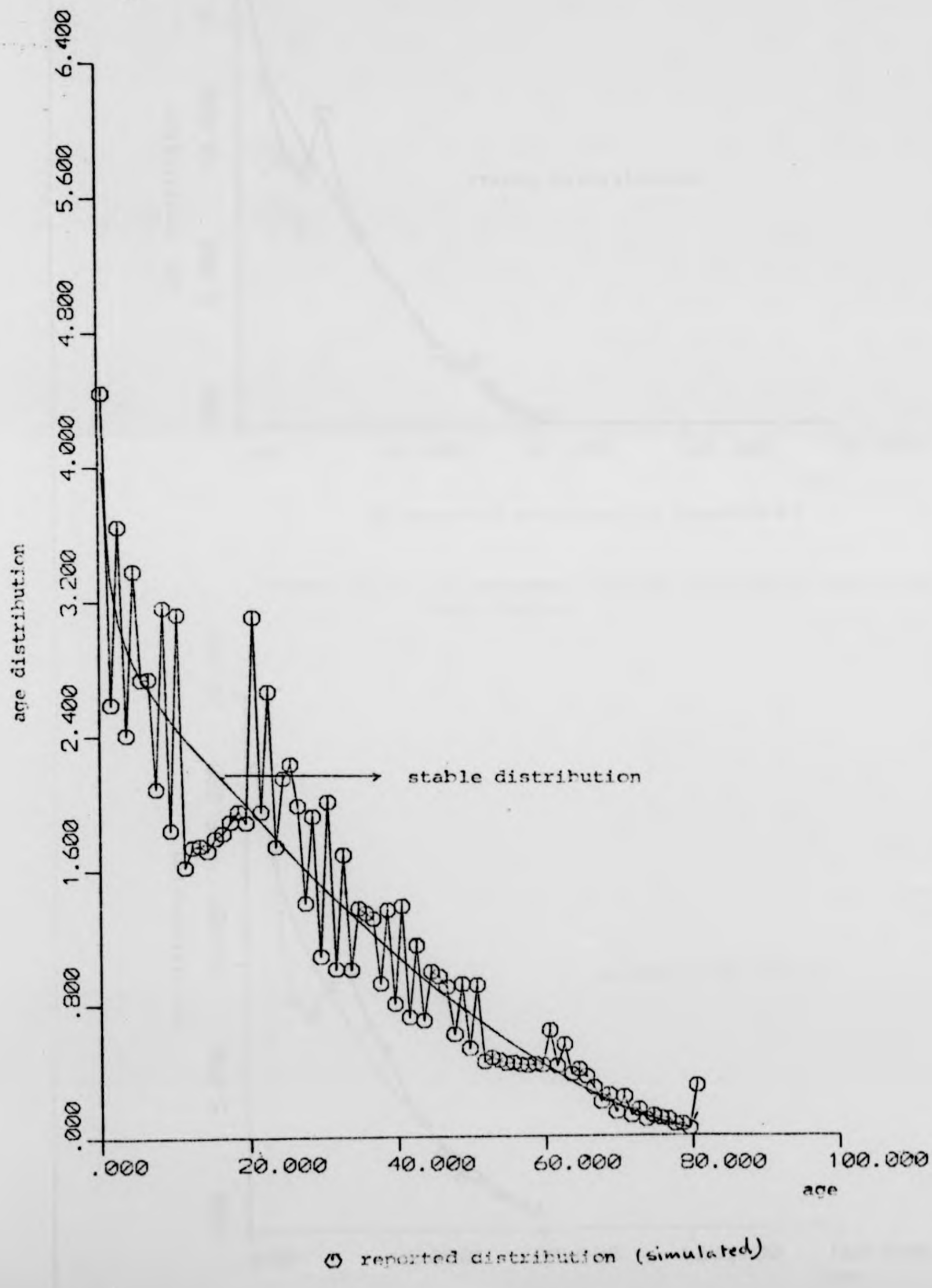
Table (6.2) Summary of the Results

actual CDR	Reported (simulated)			
	formula (A)		formula (B)	
	application (1)	application (2)	application (1)	application (2)
33.85	32.96	37.68	30.36	34.17
25.44	26.88	27.63	25.47	24.94
19.69	21.13	20.25	19.20	18.87

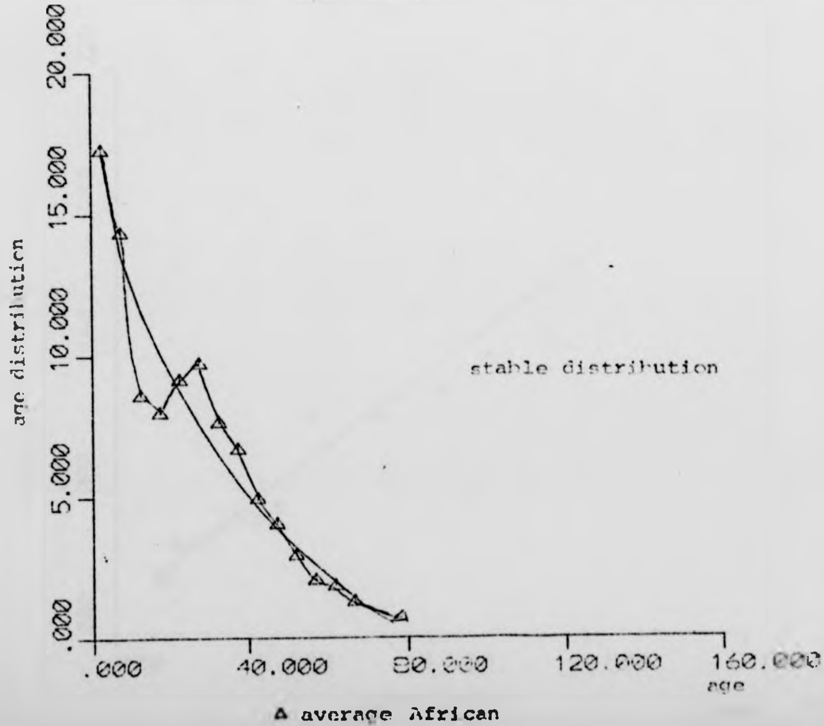
From the previous table, we note that the average deviation in the estimate of the death rate is within 1 to 2‰ and the maximum deviation is around 4‰.

Graph (6.4) represents the line passing through the points corresponding to actual death rate 33.85‰ and the sets of points formed using the simulated data of the first application and both formula (A) and (B).

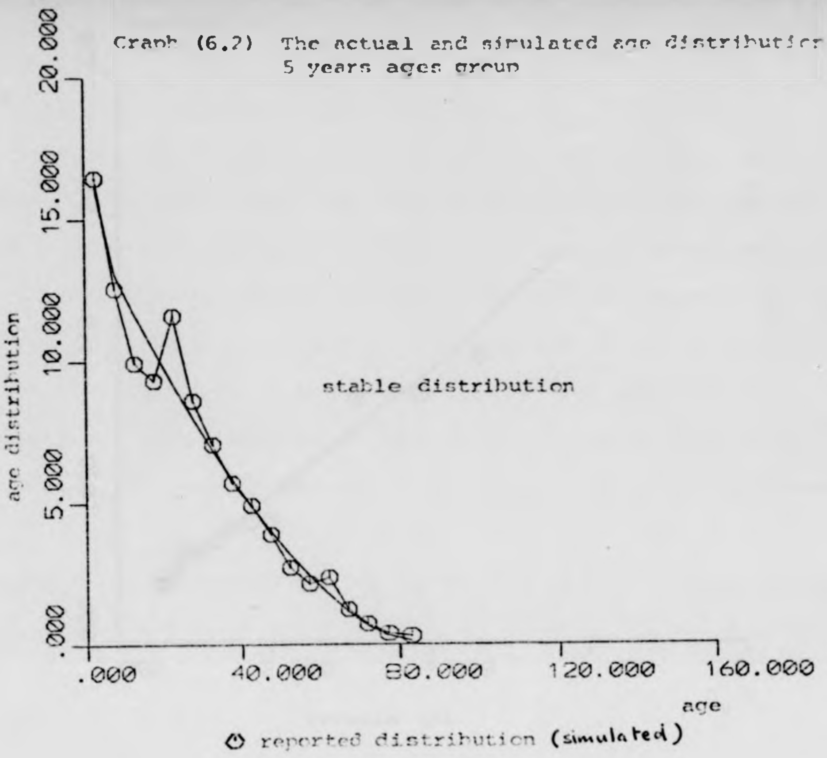
Graph (6.1) The actual and simulated age distribution in single years



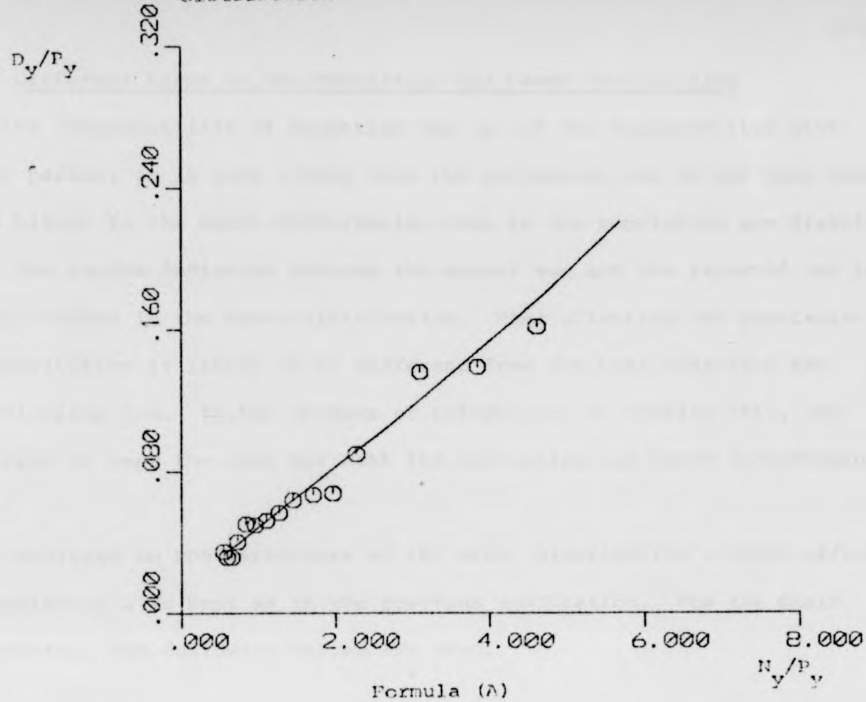
Graph (6.3) The averaged African and fitted stable age distribution

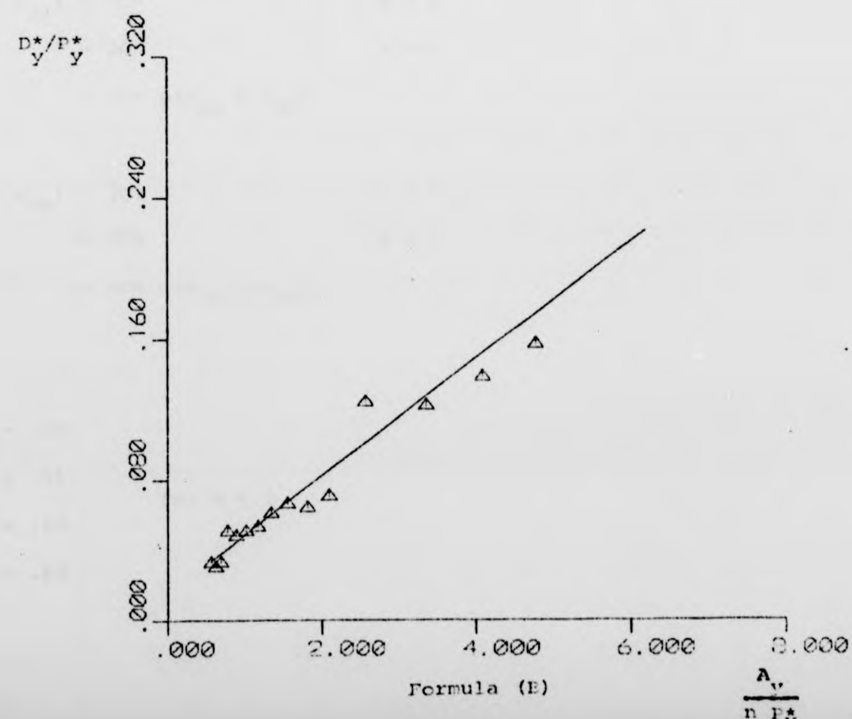


Graph (6.2) The actual and simulated age distribution in 5 years ages group



Graph (6.4) Effect of age misreport (same error in age and death distribution





6.3.2 Different Error in the Population and Death Distribution

Since the responsibility of reporting the age of the deceased lies with another person, it is more likely that the proportion who do not know their age is higher in the death distribution than in the population age distribution. The random deviation between the actual age and the reported age is probably higher in the death distribution. Bias affecting the population age distribution is likely to be different from the bias affecting the death distribution. In the absence of information to confirm this, the bias error is kept the same for both the population and death distribution.

Values assigned to the parameters of the error distribution - which affect the population - is kept as in the previous application. For the death distribution, the following values are used.

6.3.2.1 Values assigned to the parameters of the error distribution

- The probabilities of being in different groups:

$p(a_{1x} + a_{2x}) = 70\%$ $x < 5$

$= 30\%$ $x > 5$

$p(a_{2x}) = 30\%\ p(a_{1x} + a_{2x})$

$p(a_{3x} + a_{4x}) = 30\%$ $x < 5$

$= 70\%$ $x > 5$

$p(a_{4x}) = 40\%\ p(a_{3x} + a_{4x})$

then:

$p(a_{1x}) = .49$

$p(a_{2x}) = .21$ for $x < 5$

$p(a_{3x}) = .18$

$p(a_{4x}) = .12$

$p(a_{1x}) = .21$

$p(a_{2x}) = .09$ for $x > 5$

$p(a_{3x}) = .42$

$p(a_{4x}) = .28$

- Bias error

The same type of bias considered in the first case is considered here.

- Random error

The same distribution of random error applied in the first case is used here, except that the value of the parameter s_x is increased by 2 for all ages greater or equal 1. Thus, implying a bigger random error.

6.3.2.2 Results

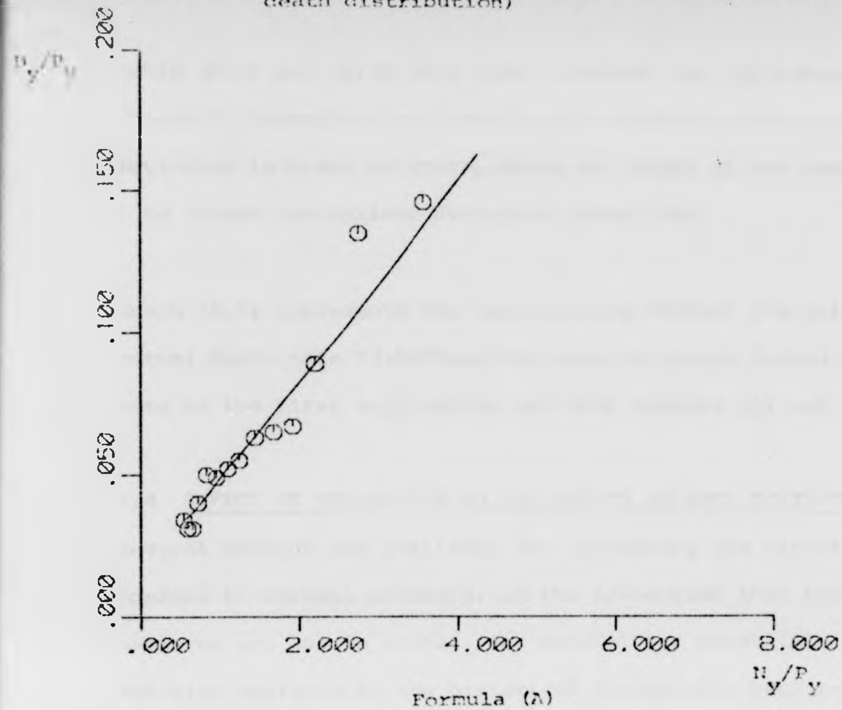
Applying the previous procedure twice on three stable distributions corresponding to model west, males with growth rate = 15 and mortality levels 6, 9 and 12. The computer results of one application is illustrated in Table (2) of Appendix (B).

The actual death rate - calculated by dividing the total deaths over the total population - and the estimated death rate - calculated by applying the growth balance method on the reported age distributions - are given in Table (6.4).

Table 6.4 Summary of the Results

actual CDR	Reported (simulated)			
	formula (A)		formula (B)	
	application (1)	application (2)	application (3)	application (4)
33.85	36.54	35.91	32.00	32.00
25.44	24.32	25.74	21.48	22.36
19.69	19.35	19.52	17.67	17.68

Graph (6.5) Effect of age misreport (different error in age and death distribution)



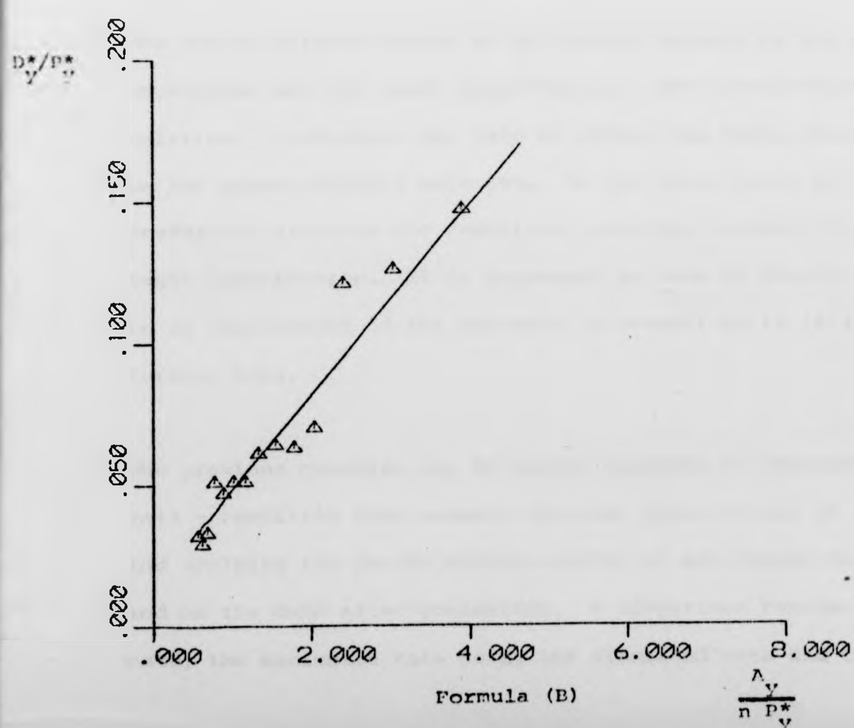


Table (6.2) and (6.4) show that - whether the population distribution and the death distribution share the same kind of error or not - the average deviation in Brass estimate, under our model of age reporting, is within 1 to 2% and the maximum deviation around 4%.

Graph (6.5) represents the line passing through the points corresponding to actual death rate 33.85‰ and the sets of points formed using the simulated data of the first application and both formula (A) and (B).

6.4 EFFECT OF GRADUATION ON THE GROWTH BALANCE ESTIMATE

Several methods are available for graduating age distributions to make them conform to certain patterns; on the assumption that the deviations from these patterns are due to error. Any attempt for graduation must be preceeded by a detailed analysis of the historical demographic background of the population under study, so that peculiarities of the data which have historical foundations are not treated as error.

The growth balance method of estimation depends on the relation between the population and the death distribution. Age misreporting affects this relation. Graduation may help to offset the bias - caused by age misreporting - in the growth balance estimate. On the other hand, it is possible that graduation distorts the underlying relation between the population and the death distribution. It is important to know if smoothing the data results in an improvement of the estimate in general or if it is a source of a further bias.

The previous question may be easily answered by considering the simulated data - resulting from several computer applications of the previous model - and applying the growth balance method of estimation on this data first and on the data after graduation. A comparison between the actual death rate, the estimated rate using the simulated data and the estimated using the graduated data provides the answer.

Since the data to be graduated are hypothetical, there is no historical factors requiring special attention and graduation becomes a simple exercise. Sophisticated techniques of graduation are more appropriate to application on data of high quality and thus two of the simpler methods of graduation are used here; mainly: Quadratic graduation and another technique devised by Brass. First, a brief account of these methods is given, then the results are introduced in Table (6.6).

- Quadratic graduation

Except at very young or old ages, it is assumed that the data is a quadratic function over a limited age range. The data required is the numbers in three consecutive age groups of length 10 each. Three sets of coefficients are provided, to permit splits of the youngest, middle and oldest of the three groups into five year groups. For each set, three coefficients are given. The number of persons in the first, middle and last of the age groups respectively are multiplied by these coefficients and the resulting products are finally added to construct the first five year age group, the second group is reached by substraction. The middle group is chosen as the one to split whenever this is possible.

Table (6.5)* Age splitting coefficients

coefficient to calculate the population from the younger side of an age group to the middle of the age group given three consecutive age groups of equal length

to calculate part of the								
youngest age group coefficients of			middle age group coefficients of			oldest age group coefficients of		
youngest age group	middle age group	oldest age group	youngest age group	middle age group	oldest age group	youngest age group	middle age group	oldest age group
.6875	-.25	.0625	.0625	.5	-.0625	-.0625	.25	.3125

*source: Carrier & Hobcraft (1971)

- Brass technique for graduation:

The proportions below various ages are calculated, the logits of these proportions are assumed to form a straight line when plotted against the logits of the proportions under the same ages of an appropriate reference population. Once a line is fitted to the data, the fitted points on this line corresponds to the logit of the graduated data. By reversing the logits, the proportion under any age is obtained and the graduated proportions in age groups may be reached by substraction. The reference population used is Brass standard life table (Brass, 1971).

Table (6.6) Effect of graduation on the estimate of the crude death rate

actual	Estimated*					
	Reported data		Graduated (Brass graduation)		Graduated (Quadratic)	
	formula (a)	formula (b)	formula (a)	formula (b)	formula (a)	formula (b)
33.85	35.60	31.53	38.25	37.66	39.97	21.35
	33.22	30.98	36.62	35.65	33.81	29.92
25.44	27.65	24.98	27.55	26.86	29.33	19.16
	26.67	24.66	29.79	29.30	26.82	25.29
19.69	21.16	19.24	22.98	22.70	22.46	15.35
	18.65	17.60	22.37	21.98	18.61	17.57

*the method of fit used is least square using 15 age groups.

From table (6.6) it is clear that graduation does not always improve the estimate and is likely to distort it considerably.

It may be argued that the use of Brass standard as our reference population is one of the reasons for this distortion, since the actual data correspond to a stable population affected by west mortality pattern. Table (6.7) presents the actual stable population and death distributions with the simulated distributions (affected by age error) and also, the graduated data using the stable distributions as the reference population. Though, in this case the graduation does improve the simulated data, the crude death rate calculated using the graduated data and Brass method is still distorted as illustrated in Table (6.8). Thus graduation is not recommended before applying Brass method for estimating mortality.

Table (6.7) The stable, simulated and graduated population and death distributions

age	stable		simulated		graduated using as reference stable distribution	
	population	deaths	population	deaths	population	deaths
0-	13.68	39.11	14.15	39.28	14.15	39.42
5-	11.87	2.87	11.2	2.82	11.61	2.81
10-	10.78	1.89	9.2	1.72	10.39	1.85
15-	9.8	2.45	8.62	2.6	9.41	2.4
20-	8.82	3.14	11.27	3.19	8.49	3.07
25-	7.89	3.10	8.31	3.14	7.65	3.03
30-	7.02	3.17	7.28	3.07	6.88	3.10
35-	6.21	3.33	5.96	3.48	6.14	3.26
40-	5.43	3.62	5.66	4.02	5.46	3.53
45-	4.68	3.85	4.39	3.44	4.78	3.77
50-	3.95	4.33	4.01	5.04	4.11	4.26
55-	3.24	4.70	3.07	4.49	3.43	4.62
60-	2.53	5.24	2.5	5.09	2.74	5.20
65-	1.85	5.41	1.78	4.59	2.05	5.39
70-	1.21	5.18	1.24	4.62	1.38	5.24
75-	.67	4.31	.71	3.89	.79	4.40
80-	.36	4.30	.65	5.46	.45	4.55

Table (6.8) Actual and estimated death rate using Brass method

actual	Estimated*			
		stable data	simulated	graduated
19.72	formula (A)	19.91	18.65	22.38
	formula (B)	19.21	17.60	22.02

*method of fit, least square using 15 age groups

CHAPTER VII

THE BEST METHOD OF FIT

7.1 INTRODUCTION

Several methods are available for fitting straight lines. These methods differ not only in their underlying assumptions but also in the effort and time they require. It is our purpose to discuss some of these methods and suggest which is likely to yield a superior estimate of the death rate in most cases.

The problem of fitting straight lines has been treated by many authors; the procedure suggested depends on the priori assumptions of the distribution law of error and whether one or both variables are subject to error and also on the criteria used for determining the best fit. Relating these assumptions to our special application is not attempted since our underlying distribution of error is not clear due to the many different combinations of factors likely to affect the relation between $\frac{N_y}{P_y}$ and $\frac{D_y}{P_y}$. Thus, though the theoretical basis for applying any of these methods may be questionable, the justification for suggesting them is simply the accuracy of the estimate they provide.

The methods discussed in this part are by no means exhaustive but they represent a selection of the more famous ones:

1. Wald method.
2. Bartlett method.
3. Least-square method.
4. Weighted least-square method.
5. Anscombe method.
6. Biweight regression method (Tukey).

After a brief presentation of these methods, numerical applications on several types of data are given and the best method suggested and discussed.

7.2 BRIEF PRESENTATION OF THE METHODS

7.2.1 Wald Method

Let us consider two sets of random variables: $x_1, x_2, x_3, \ldots x_N$; $y_1, y_2, y_3, \ldots y_N$ such that the relation between the true values is given by: Y = intercept + slope X.

Wald deals with the case when both variables X and Y are subject to errors. Under the assumptions that the errors in the X variables have the same distribution and are uncorrelated, the errors in the Y variables have the same distribution and are uncorrelated, and also, the errors in the X and Y are uncorrelated; a consistent estimate of the slope is given by dividing the data into two groups and drawing a line through the averages of these two groups, thus:

$$\hat{\text{slope}} = \frac{\dfrac{\sum_{i=1}^{m} y_i}{m} - \dfrac{\sum_{i=m+1}^{M} y_i}{M-m}}{\dfrac{\sum_{i=1}^{m} x_i}{m} - \dfrac{\sum_{i=m+1}^{M} x_i}{M-m}}$$

where M is the total number of observations and m is close to $\frac{N}{2}$.

$$\hat{\text{intercept}} = \bar{Y} - \hat{\text{slope}}\,\bar{X}$$

where $\bar{Y} = \dfrac{\sum_{i=1}^{M} y_i}{M}$ and $\bar{X} = \dfrac{\sum_{i=1}^{M} x_i}{M}$

7.2.2 Bartlett Method

Bartlett method is a modification of Wald's method with the same underlying assumptions. This modification appears in the use of three groups instead of two for estimating the slope. Thus, the number of observations M is divided into three groups - such that the equal number K2 in the two extreme groups is chosen as near $\frac{M}{3}$ as possible - and the estimate of the slope given by:

$$\widehat{slope} = \frac{\dfrac{\sum_{i=1}^{K2} y_i}{K2} - \dfrac{\sum_{i=M-K2+1}^{M} y_i}{K2}}{\dfrac{\sum_{i=1}^{K2} x_i}{K2} - \dfrac{\sum_{i=M-K2+1}^{M} x_i}{K2}}$$

$$\widehat{intercept} = \bar{Y} - \widehat{slope}\,\bar{X}.$$

7.2.3 Least-Square Method

Let us consider the case when: y_i = intercept + slope $x_i + \varepsilon_i$. The least square method minimizes the sum of squares of deviations of the observed y from the estimated y $(\sum_i (y_i - \text{intercept} - \text{slope } x_i)^2)$. The method is appropriate when the deviation law has a symmetrical form and when it is assumed that the scatter of the observations about the regression curve is the same at all points.

The assumption of normality of the error - usually associated with this method - is only required when confidence limits and tests of significance are used.

$$\widehat{slope} = \frac{\sum_{i=1}^{M} (y_i - \bar{Y})(x - \bar{X})}{\Sigma (x - \bar{X})^2}$$

$$\widehat{intercept} = \bar{Y} - \widehat{slope}\,\bar{X}$$

$$\bar{Y} = \frac{\Sigma y_i}{M}, \quad \bar{X} = \frac{\Sigma x_i}{M}.$$

7.2.4 Weighted Least-Square Method

For cases when the scatter of error is different for different points; in other words, when the variances of the y_i satisfy the relationship:

$$\sigma_i^2 = \sigma^2/V_i \qquad i = 1, 2, \ldots n$$

The estimate of the slope that minimizes the weighted sum of squares is calculated as:

$$\hat{slope} = \frac{\Sigma V_i (x_i - \bar{X})(y_i - \bar{Y})}{\Sigma V_i (x_i - \bar{X})^2}$$

$$\hat{intercept} = \bar{Y} - \hat{slope}\,\bar{X}$$

where $\bar{X}$ and $\bar{Y}$ denote the weighted means, i.e.:

$$\bar{Y} = \frac{\Sigma V_i y_i}{\Sigma V_i} \text{ and } \bar{X} = \frac{\Sigma V_i x_i}{\Sigma V_i}$$

$$V_i \propto \frac{1}{\sigma_i}.$$

7.2.5 Anscombe Method

This method has the same principle as the weighted least square. It deals with the situation when the distribution of error is more skew than symmetrical (with the same dispersion). It allows also for the possibility that some errors in the observations may be due to a mistake, which results in some points that ought to be neglected (since their variability are different from the underlying variability of the phenomena). Thus, this method modifies the least square procedure to allow for a long tailed distribution of error of good observations and for the possible occurrence of bad observations.

The objective function minimized by Anscombe is:

$$\Sigma_1 (y_i - u_i^*)^2 + \Sigma_2 K_1 (2|y_i - u_i^*| - K_1) + \Sigma_3 K_1 (2K_2 - K_1).$$

where: $\hat{u}_i$ = intercept + slope x_i, K_1 and K_2 are chosen numbers, Σ_1 denotes summation over the values of i such that $|y_i - \hat{u}_i| \leq K_1$, Σ_2 denotes summation over the values of i such that $K_1 \leq |y_i - \hat{u}_i| \leq K_2$, Σ_3 denotes summation over the remaining values such that $|y_i - \hat{u}_i| > K_2$.

In other words, we minimize the weighted sum of squares:

$\Sigma\, V_i\, (y_i - \hat{u}_i)^2$, where the weights satisfy:

$$V_i = 1 \qquad \text{if } |y_i - \hat{u}_i| \leq K_1$$

$$V_i = K_1/|y_i - \hat{u}_i| \qquad \text{if } K_1 < |y_i - \hat{u}_i| < K_2$$

$$V_i = 0 \qquad \text{if } |y_i - \hat{u}_i| > K_2$$

where $\hat{u}_i$ = $\widehat{\text{intercept}}$ + $\widehat{\text{slope}}$ x_i.

Values of K_1 may be chosen in the neighbourhood of twice the standard deviation of the error distribution and K_2 around 3 or 4 times as large as K_1.

Generally, this procedure requires a number of iterations, unless we are able to assign the observations correctly to the summation at the outset. In application, we took the initial values for V_i for the first third of observations equal to 1, for the second third equal to K_1 and for the last third equal to zero. We iterated until there is little change in the estimates of the intercept and slope recomputing the weights each time.

7.2.6 <u>Biweight Regression Method (Tukey)</u>

Both weighted least squares and Anscombe methods require an estimate of weights supplied by the researcher; on the other hand, Tukey procedure uses weights dependent on the residual in the previous iteration. Thus, if $\varepsilon_i = \dfrac{y_i - \hat{y}_i\ (K)}{h\ s_K^{*}}$; where h is a numerical constant, s_K^{*} a measure of spread of the residuals left by the K^{th} fit and $\hat{y}_i$ (K) is the fitted value for y_i at the K^{th} step, then: $V_i = (1 - \varepsilon_i^2)^2$, a good choice for h is 6 and for s_K^{*} is median $|y_i - \hat{y}_i\ (K)|$. In application, we took some initial values for $V_i - V_i = 1$ - we iterated until there is little change in the estimates of the intercept and slope, recomputing the weights each time.

7.3 NUMERICAL APPLICATIONS ON SEVERAL TYPES OF DATA

The first type of data considered is stable data. These data satisfy all the requirements for applying Brass method, mainly stability and no error introduced through age misreport or differential under-registration. The only source of error that appear is due to our procedure for estimating N_Y when using formula (A) and of estimating P_y^* and D_y^* when using formula (B).

The estimate of the dispersion of the error presents no problem in this case since the actual growth rate (r) and death rate (CDR) are available. Thus, the dispersion is proportionate to $\sqrt{\Sigma(y_i - r - CDR\ x_i)^2}$. In the weighted least square it was assured that the dispersion is equal in the first, second and last third of the observations. For the other methods it was assumed equal all over the age span.

Table (7.1) shows the results of several applications on age and death distributions given in Coale & Demeny (1966), model west, males, corresponding to various mortality levels and growth rates.

From Table (7.1) it is clear that all the methods perform well when using formula (B); while only, Wald, Bartlett, Least square methods using 10 observations and weighted least square method perform well when using formula (A).

Table (7.1) The actual death rate corresponding to different stable population given in Coale & Demeny (1966) and the estimated death rate using the growth balance method and several methods of fit applied on both formula (A) and formula (B)

ACTUAL DEATH RATE .059770

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.064491	.059165
WALD	10	.058490	.059376
BARTLETT	15	.065631	.059140
BARTLETT	10	.057890	.059384
LEAST SQUARE	15	.067627	.059051
LEAST SQUARE	10	.058489	.059376
WEIGHTED L.S.	15	.060255	.059300
ANSCOMBE	15	.067627 (1)	.059051 (2)
TUKEY	15	.069177 (3)	.059342 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 4
(2)*NUMBER OF STEPS TILL CONVERGENCE= 2
(3)*NUMBER OF STEPS TILL CONVERGENCE=333
(4)*NUMBER OF STEPS TILL CONVERGENCE= 2

ACTUAL DEATH RATE .047500

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.050228	.046808
WALD	10	.046607	.047163
BARTLETT	15	.050539	.046864
BARTLETT	10	.046188	.047174
LEAST SQUARE	15	.052135	.046324
LEAST SQUARE	10	.046616	.047184
WEIGHTED L.S.	15	.047841	.047059
ANSCOMBE	15	.052135 (1)	.046324 (2)
TUKEY	15	.050842 (3)	.046756 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 4
(2)*NUMBER OF STEPS TILL CONVERGENCE= 2
(3)*NUMBER OF STEPS TILL CONVERGENCE= 24
(4)*NUMBER OF STEPS TILL CONVERGENCE= 1

Table (7.1) (continued)

ACTUAL DEATH RATE .037680

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.038941	.037033
WALD	10	.036910	.037328
BARTLETT	15	.039114	.036997
BARTLETT	10	.036646	.037331
LEAST SQUARE	15	.039966	.036377
LEAST SQUARE	10	.036898	.037345
WEIGHTED L.S.	15	.037605	.037225
ANSCOMBE	15	.039966 (1)	.036377 (2)
TUKEY	15	.040865 (3)	.036588 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 4
(2)*NUMBER OF STEPS TILL CONVERGENCE= 2
(3)*NUMBER OF STEPS TILL CONVERGENCE= 31
(4)*NO CONVERGENCE TILL 400 STEPS

ACTUAL DEATH RATE .030700

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.031301	.030078
WALD	10	.030144	.030441
BARTLETT	15	.031398	.030038
BARTLETT	10	.029974	.030455
LEAST SQUARE	15	.031887	.029841
LEAST SQUARE	10	.030130	.030456
WEIGHTED L.S.	15	.030566	.030388
ANSCOMBE	15	.031887 (1)	.029841 (2)
TUKEY	15	.031527 (3)	.029224 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 4
(2)*NUMBER OF STEPS TILL CONVERGENCE= 2
(3)*NUMBER OF STEPS TILL CONVERGENCE= 3
(4)*NO CONVERGENCE TILL 400 STEPS

Table (7.1) (continued)

ACTUAL DEATH RATE .018050

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.018023	.017752
WALD	10	.017844	.017959
BARTLETT	15	.018037	.017729
BARTLETT	10	.017796	.017954
LEAST SQUARE	15	.018155	.017652
LEAST SQUARE	10	.017834	.017956
WEIGHTED L.S.	15	.017995	.017921
ANSCOMBE	15	.017937 (1)	.017652 (2)
TUKEY	15	.018161 (3)	.017385 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 2
(2)*NUMBER OF STEPS TILL CONVERGENCE= 2
(3)*NUMBER OF STEPS TILL CONVERGENCE= 4
(4)*NO CONVERGENCE TILL 400 STEPS

ACTUAL DEATH RATE .015430

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.015383	.015216
WALD	10	.015268	.015345
BARTLETT	15	.015394	.015201
BARTLETT	10	.015237	.015350
LEAST SQUARE	15	.015492	.015153
LEAST SQUARE	10	.015256	.015346
WEIGHTED L.S.	15	.015359	.015322
ANSCOMBE	15	.015301 (1)	.015153 (2)
TUKEY	15	.015447 (3)	.014937 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 1
(2)*NUMBER OF STEPS TILL CONVERGENCE= 2
(3)*NUMBER OF STEPS TILL CONVERGENCE= 3
(4)*NO CONVERGENCE TILL 400 STEPS

Table (7.1) (continued)

ACTUAL DEATH RATE .013290

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.013187	.013100
WALD	10	.013171	.013223
BARTLETT	15	.013189	.013087
BARTLETT	10	.013152	.013222
LEAST SQUARE	15	.013234	.013047
LEAST SQUARE	10	.013156	.013216
WEIGHTED L.S.	15	.013189	.013192
ANSCOMBE	15	.013122 (1)	.013047 (2)
TUKEY	15	.013359 (3)	.012992 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 1
(2)*NUMBER OF STEPS TILL CONVERGENCE= 2
(3)*NO CONVERGENCE TILL 400 STEPS
(4)*NUMBER OF STEPS TILL CONVERGENCE= 1

ACTUAL DEATH RATE .011430

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.011318	.011266
WALD	10	.011322	.011359
BARTLETT	15	.011318	.011255
BARTLETT	10	.011316	.011360
LEAST SQUARE	15	.011342	.011206
LEAST SQUARE	10	.011313	.011359
WEIGHTED L.S.	15	.011320	.011339
ANSCOMBE	15	.011342 (1)	.011206 (2)
TUKEY	15	.011389 (3)	.011067 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 1
(2)*NUMBER OF STEPS TILL CONVERGENCE= 2
(3)*NUMBER OF STEPS TILL CONVERGENCE= 1
(4)*NO CONVERGENCE TILL 400 STEPS

The justification for this is very simple. In formula (B) the error introduced by our procedure for estimating P^*_y and D^*_y is very minor; thus the observations lie on a straight line and any method of fit should perform well in this ideal case. In formula (A) the error introduced by our procedure for estimating N_y is big only corresponding to old ages and especially when the death rate is high (the decline in the age distribution is not linear). Any method neglecting observations corresponding to old ages perform well.

Weighted least square method is the only method that was able to remedy the error introduced by our procedure and result in a plausible estimate for the death rate (note that the dispersion of error increases with age). Both Anscombe and Tukey methods in spite of their complicated structure did not perform well.

From the previous remarks our suggestions is to reject both Anscombe and Tukey methods and choose one of the following methods: Wald (using 10 observations), Bartlett (using 10 observations), Least square (using 10 observations) or weighted least square.

A recommendation for using either Wald or Bartlett methods with only 10 observations is expected, since they are the simplest in application. But, it should be pointed out that the previous applications are under ideal circumstances when no deviation from the assumption exist. Actual data are affected by age misreport, migration and a change in mortality and fertility. For example, even if it is expected that certain error may affect old ages it may be true that changes in fertility or migration or age misreport have altered the age distribution for young ages and data corresponding to old ages are more representative of the magnitude of the death rate. Thus before a final recommendation for a certain method, more applications on representative data should be attempted.

The model of age error introduced in Chapter (6) provides us with a vast source of information. First, stable data before subjecting it to age error, then simulated data affected by age error, also graduated data using Brass and quadratic graduation. To confirm the previous conclusion for rejecting both Anscombe and Tukey methods and to help choose the best method of fit, all the previous methods are applied on several sets of data (each set comprises: stable, reported and graduated data) and the results are given in Table (7.2).

Table (7.2) confirm our conclusion for the inadequacy of Anscombe and Tukey procedure in applying the growth balance method; for example, the estimates of the graduated data (Brass technique) in the first set using Anscombe and Tukey are given by: 36.7‰, 35.7‰, 36.7‰ and 41.7‰ instead of 33.85.

The method that performs well in most cases is the weighted least square method; for example, considering the graduated data of the first set and comparing the estimates of the death rates using Wald method (10 observations), Bartlett method (10 observations) and the weighted least square we get:

actual death rate = 33.85‰

method	estimated death rate using either formula (A) or (B) in several runs			
Wald	29.72	24.58	37.12	32.95
Bartlett	30.39	24.76	38.04	37.89
W. Least square	32.18	34.80	33.77	31.42

Thus, as a general recommendation, sophisticated techniques of fitting the straight line are not accepted and the weighted least square method is suggested. It should be emphasized that one method cannot be expected

Table (7.2) The actual and estimated death rate, using the growth balance method and several methods of fit applied on both formula (A) and formula (B), corresponding to different sets of data.

ACTUAL DEATH RATE .033850

STABLE DATA

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.035132	.033620
WALD	10	.033347	.033691
BARTLETT	15	.035287	.033611
BARTLETT	10	.033140	.033701
LEAST SQUARE	15	.036085	.033541
LEAST SQUARE	10	.033344	.033692
WEIGHTED L.S.	15	.033010	.033681
ANSCOMBE	15	.036085 (1)	.033541 (2)
TUKEY	15	.035381 (3)	.032418 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 3
(2)*NUMBER OF STEPS TILL CONVERGENCE= 2
(3)*NUMBER OF STEPS TILL CONVERGENCE= 9
(4)*NO CONERGENCE TILL 400 STEPS

REPORTED DATA(SIMULATED)

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.032877	.031602
WALD	10	.033274	.033318
BARTLETT	15	.033100	.031317
BARTLETT	10	.036341	.037840
LEAST SQUARE	15	.033224	.030983
LEAST SQUARE	10	.033475	.034214
WEIGHTED L.S.	15	.032936	.031274
ANSCOMBE	15	.033268 (1)	.030983 (2)
TUKEY	15	.033631 (3)	.030865 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 3
(2)*NUMBER OF STEPS TILL CONVERGENCE= 2
(3)*NUMBER OF STEPS TILL CONVERGENCE= 9
(4)*NUMBER OF STEPS TILL CONVERGENCE= 7

Table (7.2) (continued)

GRADUATED DATA(BRASS TECHNIQUE)

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.033159	.032807
WALD	10	.029725	.024583
BARTLETT	15	.033524	.034074
BARTLETT	10	.030394	.024768
LEAST SQUARE	15	.036629	.035651
LEAST SQUARE	10	.030127	.021352
WEIGHTED L.S.	15	.032185	.034809
ANSCOMBE	15	.036705 (1)	.035730 (2)
TUKEY	15	.036756 (3)	.041740 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 4
(2)*NUMBER OF STEPS TILL CONVERGENCE= 2
(3)*NO CONERGENCE TILL 400 STEPS
(4)*NO CONERGENCE TILL 400 STEPS

GRADUATED DATA(QUADRATIC FORMULA)

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.033410	.031214
WALD	10	.037124	.032952
BARTLETT	15	.034039	.031006
BARTLETT	10	.038045	.037395
LEAST SQUARE	15	.033810	.029029
LEAST SQUARE	10	.035185	.034365
WEIGHTED L.S.	15	.033775	.031421
ANSCOMBE	15	.033875 (1)	.029029 (2)
TUKEY	15	.031880 (3)	.029740 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 2
(2)*NUMBER OF STEPS TILL CONVERGENCE= 3
(3)*NO CONERGENCE TILL 400 STEPS
(4)*NO CONERGENCE TILL 400 STEPS

Table (7.2) (continued)

ACTUAL DEATH RATE .025441

STABLE DATA

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.025875	.025191
WALD	10	.025119	.025341
BARTLETT	15	.025941	.025179
BARTLETT	10	.025003	.025328
LEAST SQUARE	15	.026310	.025135
LEAST SQUARE	10	.025103	.025324
WEIGHTED L.S.	15	.025362	.025295
ANSCOMBE	15	.026310 (1)	.025135 (2)
TUKEY	15	.026075 (3)	.025054 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 3
(2)*NUMBER OF STEPS TILL CONVERGENCE= 2
(3)*NUMBER OF STEPS TILL CONVERGENCE= 16
(4)*NUMBER OF STEPS TILL CONVERGENCE= 3

REPORTED DATA(SIMULATED)

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.025749	.024388
WALD	10	.025181	.025311
BARTLETT	15	.026110	.024779
BARTLETT	10	.027477	.023739
LEAST SQUARE	15	.026679	.024668
LEAST SQUARE	10	.025307	.026196
WEIGHTED L.S.	15	.025860	.024728
ANSCOMBE	15	.026679 (1)	.024663 (2)
TUKEY	15	.025717 (3)	.024113 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 2
(2)*NUMBER OF STEPS TILL CONVERGENCE= 2
(
(4)*NUMBER OF STEPS TILL CONVERGENCE= 5

Table (7.2) (continued)

GRADUATED DATA(BRASS TECHNIQUE)

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.027411	.027438
WALD	10	.024809	.021345
BARTLETT	15	.027566	.028289
BARTLETT	10	.025175	.021363
LEAST SQUARE	15	.029798	.029365
LEAST SQUARE	10	.025267	.019029
WEIGHTED L.S.	15	.025991 (1)	.028221 (2)
ANSCOMBE	15	.029798 (3)	.029300 (4)
TUKEY	15	.029370	.029177

(1)*NUMBER OF STEPS TILL CONVERGENCE= 3
(2)*NUMBER OF STEPS TILL CONVERGENCE= 2
(3)*NUMBER OF STEPS TILL CONVERGENCE=150
(4)*NUMBER OF STEPS TILL CONVERGENCE= 17

GRADUATED DATA(QUADRATIC FORMULA)

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.025864	.025110
WALD	10	.027617	.024300
BARTLETT	15	.026477	.025189
BARTLETT	10	.028396	.028622
LEAST SQUARE	15	.026828	.025290
LEAST SQUARE	10	.026387	.025927
WEIGHTED L.S.	15	.026575 (1)	.025202 (2)
ANSCOMBE	15	.026828 (3)	.025347 (4)
TUKEY	15	.026731	.024342

(1)*NUMBER OF STEPS TILL CONVERGENCE= 2
(2)*NUMBER OF STEPS TILL CONVERGENCE= 3
(3)*NUMBER OF STEPS TILL CONVERGENCE= 5
(4)*NUMBER OF STEPS TILL CONVERGENCE= 18

Table (7.2) (continued)

ACTUAL DEATH RATE .019691

STABLE DATA

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.019821	.019509
WALD	10	.019480	.019627
BARTLETT	15	.019350	.019407
BARTLETT	10	.019420	.019628
LEAST SQUARE	15	.020026	.019463
LEAST SQUARE	10	.019465	.019626
WEIGHTED L.S.	15	.019593	.019603
ANSCOMBE	15	.020026 (1)	.019463 (2)
TUKEY	15	.020195 (3)	.019222 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 3
(2)*NUMBER OF STEPS TILL CONVERGENCE= 2
(3)*NUMBER OF STEPS TILL CONVERGENCE= 4
(4)*NO CONERGENCE TILL 400 STEPS

REPORTED DATA(SIMULATED)

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.018826	.018325
WALD	10	.019767	.019815
BARTLETT	15	.018912	.018007
BARTLETT	10	.021602	.022505
LEAST SQUARE	15	.018653	.017606
LEAST SQUARE	10	.019710	.020376
WEIGHTED L.S.	15	.018429	.018127
ANSCOMBE	15	.018653 (1)	.017606 (2)
TUKEY	15	.018459 (3)	.017340 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 3
(2)*NUMBER OF STEPS TILL CONVERGENCE= 3
(3)*NUMBER OF STEPS TILL CONVERGENCE= 10
(4)*NUMBER OF STEPS TILL CONVERGENCE= 8

Table (7.2) (continued)

GRADUATED DATA(BRASS TECHNIQUE)

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.021267	.021273
WALD	10	.020440	.017970
BARTLETT	15	.021143	.021663
BARTLETT	10	.020490	.017900
LEAST SQUARE	15	.022375	.021982
LEAST SQUARE	10	.020752	.015963
WEIGHTED L.S.	15	.021061	.021690
ANSCOMBE	15	.022381 (1)	.021986 (2)
TUKEY	15	.024962 (3)	-.034309 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 2
(2)*NUMBER OF STEPS TILL CONVERGENCE= 3
(3)*NO CONERGENCE TILL 400 STEPS
(4)*NO CONERGENCE TILL 400 STEPS

GRADUATED DATA(QUADRATIC FORMULA)

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.018311	.018245
WALD	10	.021409	.019429
BARTLETT	15	.019078	.018124
BARTLETT	10	.022085	.022365
LEAST SQUARE	15	.018615	.017576
LEAST SQUARE	10	.020459	.020220
WEIGHTED L.S.	15	.018559	.018454
ANSCOMBE	15	.018615 (1)	.017576 (2)
TUKEY	15	.018468 (3)	.017653 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 3
(2)*NUMBER OF STEPS TILL CONVERGENCE= 3
(3)*NUMBER OF STEPS TILL CONVERGENCE= 2
(4)*NUMBER OF STEPS TILL CONVERGENCE= 12

Table (7.2) (continued)

ACTUAL DEATH RATE .010509

STABLE DATA

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.010495	.010465
WALD	10	.010529	.010548
BARTLETT	15	.010491	.010455
BARTLETT	10	.010516	.010551
LEAST SQUARE	15	.010519	.010425
LEAST SQUARE	10	.010509	.010545
WEIGHTED L.S.	15	.010507	.010529
ANSCOMBE	15	.010519 (1)	.010425 (2)
TUKEY	15	.010568 (3)	.010386 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 1
(2)*NUMBER OF STEPS TILL CONVERGENCE= 2
(3)*NUMBER OF STEPS TILL CONVERGENCE= 1
(4)*NUMBER OF STEPS TILL CONVERGENCE= 1

REPORTED DATA(SIMULATED)

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.010436	.010263
WALD	10	.010995	.010934
BARTLETT	15	.010518	.010200
BARTLETT	10	.011900	.012242
LEAST SQUARE	15	.010519	.010045
LEAST SQUARE	10	.010851	.011184
WEIGHTED L.S.	15	.010458	.010118
ANSCOMBE	15	.010534 (1)	.010045 (2)
TUKEY	15	.017662 (3)	.009985 (4)

(1)*NUMBER OF STEPS TILL CONVERGENCE= 2
(2)*NUMBER OF STEPS TILL CONVERGENCE= 3
(3)*NUMBER OF STEPS TILL CONVERGENCE= 2
(4)*NUMBER OF STEPS TILL CONVERGENCE= 2

Table (7.2) (continued)

GRADUATED DATA(BRASS TECHNIQUE)

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.012167	.012234
WALD	10	.012804	.011667
BARTLETT	15	.011932	.012263
BARTLETT	10	.012688	.011564
LEAST SQUARE	15	.012186	.012001
LEAST SQUARE	10	.012000	.010312
WEIGHTED L.S.	15	.012277	.012114
ANSCOMBE	15	(1) .012186	(2) .012001
TUKEY	15	(3) .012112	(4) .012147

(1)*NUMBER OF STEPS TILL CONVERGENCE= 2
(2)*NUMBER OF STEPS TILL CONVERGENCE= 3
(3)*NUMBER OF STEPS TILL CONVERGENCE= 3
(4)*NUMBER OF STEPS TILL CONVERGENCE=356

GRADUATED DATA(QUADRATIC FORMULA)

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WALD	15	.010386	.010271
WALD	10	.011625	.010772
BARTLETT	15	.010549	.010269
BARTLETT	10	.011565	.012255
LEAST SQUARE	15	.010367	.010074
LEAST SQUARE	10	.011032	.011064
WEIGHTED L.S.	15	.010259	.010112
ANSCOMBE	15	(1) .010367	(2) .010074
TUKEY	15	(3) .010366	(4) .010207

(1)*NUMBER OF STEPS TILL CONVERGENCE= 2
(2)*NUMBER OF STEPS TILL CONVERGENCE= 3
(3)*NUMBER OF STEPS TILL CONVERGENCE= 1
(4)*NUMBER OF STEPS TILL CONVERGENCE= 10

to perform well in all actual applications. Plotting the sets of reported points with the line formed using any method of fit, or even with the different lines formed using different methods, is indispensible before the fitted slope is accepted as an estimate of the adjusted death rate.

7.4 Discussion

In all the previous applications the actual growth and death rate were known. Thus, the dispersion of error (and consequently the weights) were easily calculated. It is our purpose now to show that the weighted least square method performs equally well when the dispersion is not known and the weights are estimated from the data.

Two estimates are proposed. The first is to divide the points into 3 equal groups; the sum of square of the differences between the y points and the mean of the y points of each group is used for estimating the weights, thus:

$$V_i : V_j : V_K = \frac{1}{\sum_{s=1}^{\frac{M}{3}} \left(Y_s - \frac{\sum_{r=1}^{\frac{M}{3}} Y_r}{\frac{M}{3}} \right)^2} : \frac{1}{\sum_{s=\frac{M}{3}+1}^{\frac{2M}{3}} \left(Y_s - \frac{\sum_{r=\frac{M}{3}+1}^{\frac{2M}{3}} Y_r}{\frac{M}{3}} \right)^2} :$$

$$\frac{1}{\sum_{s=\frac{2M}{3}+1}^{M} \left(Y_s - \frac{\sum_{r=\frac{2M}{3}+1}^{M} Y_r}{\frac{M}{3}} \right)^2}$$

$i = 1, 2, \ldots, \frac{M}{3}$, $j = \frac{M}{3} + 1, \ldots, \frac{2M}{3}$ and $K = \frac{2M}{3} + 1, \ldots, M.$

The second uses the sum of squares of first differences between the consecutive y points in each groups as an indication of the weights, thus:

$$V_i : V_j : V_K = \frac{\frac{M}{3} - 1}{\sum_{s=1}^{\frac{M}{3}-1} (y_s - y_{s+1})^2} : \frac{\frac{M}{3}}{\sum_{s=\frac{M}{3}}^{\frac{2M}{3}-1} (y_s - y_{s+1})^2} : \frac{\frac{M}{3}}{\sum_{s=\frac{2M}{3}}^{M-1} (y_s - y_{s+1})^2}$$

$i = 1, 2, \ldots, \frac{M}{3}$, $j = \frac{M}{3} + 1, \ldots, \frac{2M}{3}$ and $K = \frac{2M}{3} + 1, \ldots, M$.

Table (7.3) shows the actual and estimated death rates for stable and simulated data using the weighted least square method where the weights are calculated using the mean of the y points.

Table (7.4) shows the actual and estimated death rates for stable and simulated data using the weighted least square method where the weights are calculated using the first differences.

From both Table (7.3) and (7.4) we conclude that the weighted least square is still recommended even when the weights are estimated from the data. It should be pointed out that, in some applications, more complex methods for estimating the weights may be required; these methods will probably involve some iteration.

Table (7.3) The actual and estimated death rates for different stable and simulated data using the weighted least square fit where the weights are calculated using the mean of $\frac{N}{P_y}$

Stable data

ACTUAL DEATH RATE .059770

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WEIGHTED L.S.	15	.060357	.059231

ACTUAL DEATH RATE .047580

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WEIGHTED L.S.	15	.047793	.047007

ACTUAL DEATH RATE .037680

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WEIGHTED L.S.	15	.037623	.037122

ACTUAL DEATH RATE .030730

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WEIGHTED L.S.	15	.030543	.030183

ACTUAL DEATH RATE .018050

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WEIGHTED L.S.	15	.017895	.017817

ACTUAL DEATH RATE .015430

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WEIGHTED L.S.	15	.015313	.015261

ACTUAL DEATH RATE .013290

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WEIGHTED L.S.	15	.013167	.013141

ACTUAL DEATH RATE .011430

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WEIGHTED L.S.	15	.011318	.011292

Table (7.3) (continued)

Simulated data

ACTUAL DEATH RATE .033850

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WEIGHTED L.S.	15	.033395	.032012

ACTUAL DEATH RATE .025441

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WEIGHTED L.S.	15	.026036	.025197

ACTUAL DEATH RATE .019697

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WEIGHTED L.S.	15	.019165	.018460

ACTUAL DEATH RATE .010599

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WEIGHTED L.S.	15	.010708	.010422

Table (7.4) The actual and estimated death rates for different stable and simulated data using the weighted least square fit where the weights are calculated using the first differences

Stable data

ACTUAL DEATH RATE .059770

METHOD OF FIT	N	FORMULA (A)	FORMULA (B)
WEIGHTED L.S.	15	.060661	.059228

ACTUAL DEATH RATE .047500

METHOD OF FIT	N	FORMULA (A)	FORMULA (B)
WEIGHTED L.S.	15	.047961	.047004

ACTUAL DEATH RATE .037680

METHOD OF FIT	N	FORMULA (A)	FORMULA (B)
WEIGHTED L.S.	15	.037003	.037118

Table (7.4) (continued)

Simulated data

ACTUAL DEATH RATE .033850

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WEIGHTED L.S.	15	.032204	.028488

ACTUAL DEATH RATE .025441

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WEIGHTED L.S.	15	.024046	.022324

ACTUAL DEATH RATE .019691

METHOD OF FIT	N	FORMULA(A)	FORMULA(B)
WEIGHTED L.S.	15	.019123	.017497

CHAPTER VIII

APPLICATION

8.1 INTRODUCTION

In the previous chapters, the effects of departures from the underlying assumptions of the growth balance method were studied. Also, certain practical considerations about the formulae and the method of fit used were discussed. Each deviation from the basic assumption was treated separately and adjusting procedures were suggested accordingly.

In actual data, several deviations occur, and the effect of the interaction of these factors is quite important in the analysis. It is the purpose of this chapter to illustrate the previous conclusions and apply the growth balance method under more realistic circumstances.

Two types of data are considered; the first hypothetical data affected by mortality decline, differential under-registration, age error and a migration movement. The second using actual data of a developing country.

8.2 APPLICATION ON HYPOTHETICAL DATA

Mortality decline, differential under-registration of deaths and age errors are common features in the data of developing countries. In the third chapter, a representative and flexible pattern of mortality decline was reached using the model logit system with parameters α and β; where β was fixed and α decreased as to achieve a specified increase in e_o. This pattern of mortality decline is used with the model of age error presented in Chapter Six and an assumption of higher under-registration for deaths at young ages.

The migration factor is slightly more complicated. For some countries, international migration may be assumed negligible, for others in or out migration is considerable. Data on the proportions of migrants from the total population and their age and sex composition are not generally available for developing countries and are expected to show a great deal of diversity. Three cases are considered with the data affected by mortality decline, age error and differential under-registration; the first when no migration occur, the second when out migration is dominant and finally when in migration is dominant.

8.2.1 Detailed Procedure and Data Used

Starting with a single year stable population, corresponding to an intrinsic growth rate = .010 and an underlying pattern of mortality based on the logit system and the standard model life table (Brass, 1971) with parameters $\alpha = .1$ and $\beta = .7$. The initial parameters of this population are as follows: CDR = .025, CBR = .035 and $e_o = 39$. The fertility schedule used is based on model fertility, pattern 6 of the United Nations.

This population is projected in single year periods, with constant fertility and declining mortality and a type of migration movement. Mortality is assumed to decline through a decrease in α equivalent to an approximate yearly increase in $e_o = .5$. When out migration is dominant, it is assumed to constitute a fixed proportion of the population at the end of each year; these proportions are estimated using the data in Table (4.3). When in migration is dominant, it is assumed to constitute a fixed number at the end of each year; this number is calculated using the same proportions as out migrants and the initial population before mortality change.

The projected single years population age distribution - after 10 and 20 years of declining mortality and migration - is subjected to age error. The corresponding death distribution is subjected first to an under-registration of deaths aged 0-5 by 50% and under-registration of deaths over 5 by 10%.

8.2.2 Results

The results after 10 years of projection are presented in the following section. The corresponding results after 20 years of projection are presented in Appendix (C). The conclusions drawn are the same for both sets of results.

A. Effect of deviations from stability

In Table (8.1), the actual death rate (total actual deaths/total population) - for the projected data after 10 years of mortality and migration - are presented with the death rate estimated using the growth balance procedure and three different methods of fit.

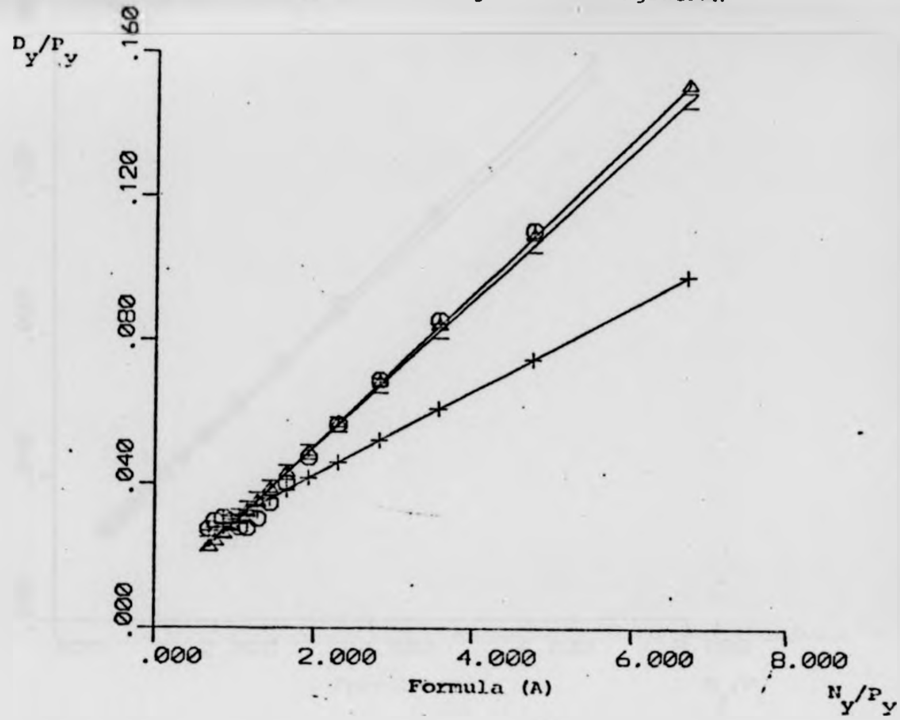
Graph (8.1), (8.2) and (8.3) represents the set of projected points for both formula (A) and (B), $\frac{N_y}{P_y}$, $\frac{D_y}{P_y}$ and $\frac{A_y}{h.P^*_y}$, $\frac{D^*_y}{P^*_y}$, and the lines drawn using the three methods of fit.

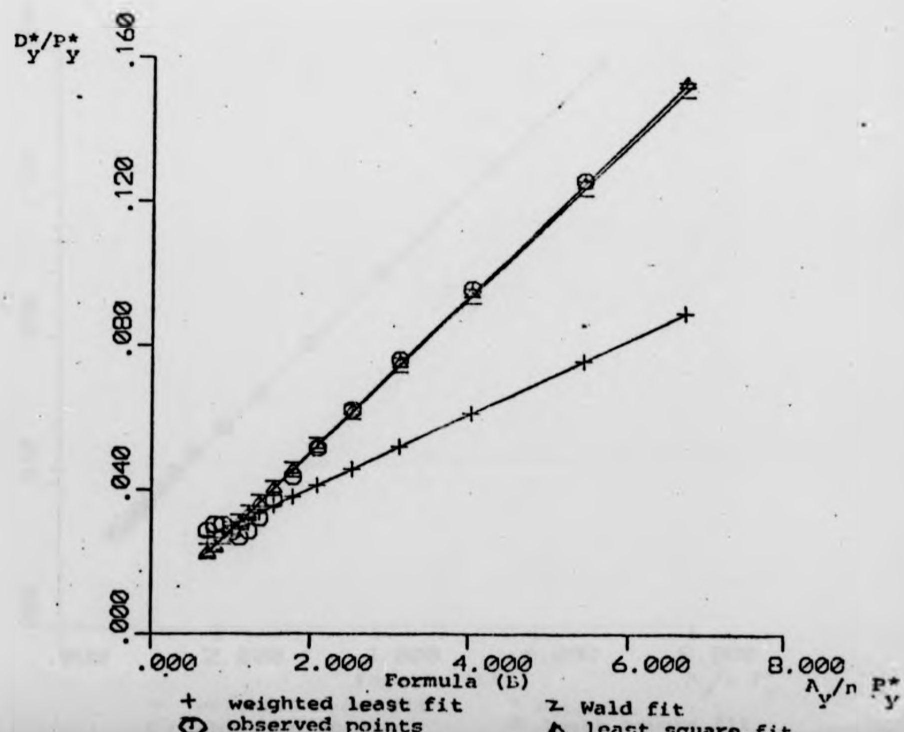
Table (8.1) The actual and estimated death rate after 10 years of mortality decline and migration

	Out migration		No migration		In migration	
Actual CDR	.021		.021		.021	
ESTIMATED CDR						
Formula	(A)	(B)	(A)	(B)	(A)	(B)
method of fit						
Least square	.021	.022	.021	.021	.021	.020
Weighted least square	.012	.010	.020	.021	.026	.026
Wald	.021	.021	.021	.021	.022	.021

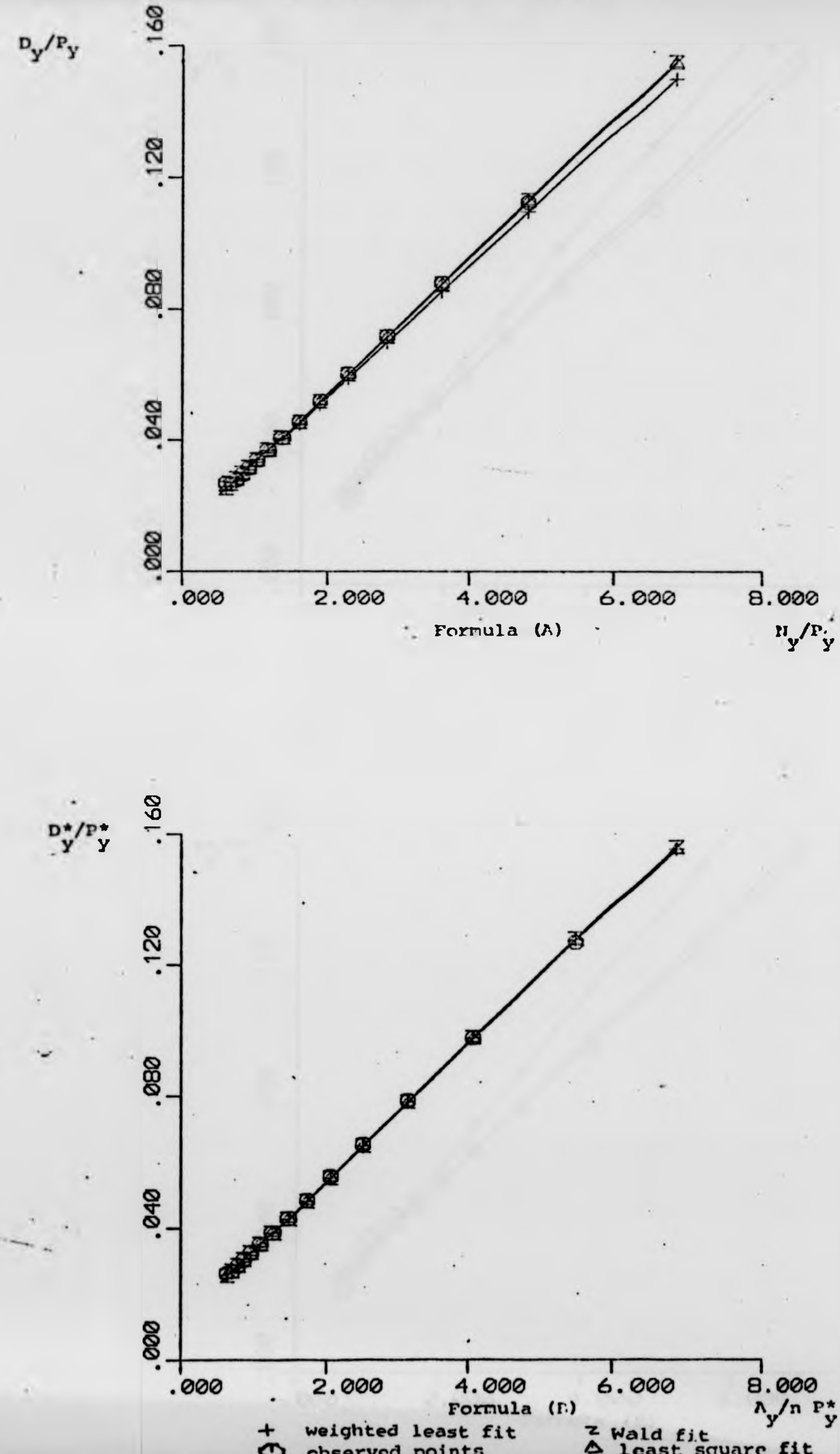
From Table (8.1), it is clear that the pattern of mortality decline considered hardly affect the estimate of the death rate. Out and in migration considerably affect the estimate of the crude death rate when the weighted least square method is used.

Graph (8.1) Changing mortality and out migration

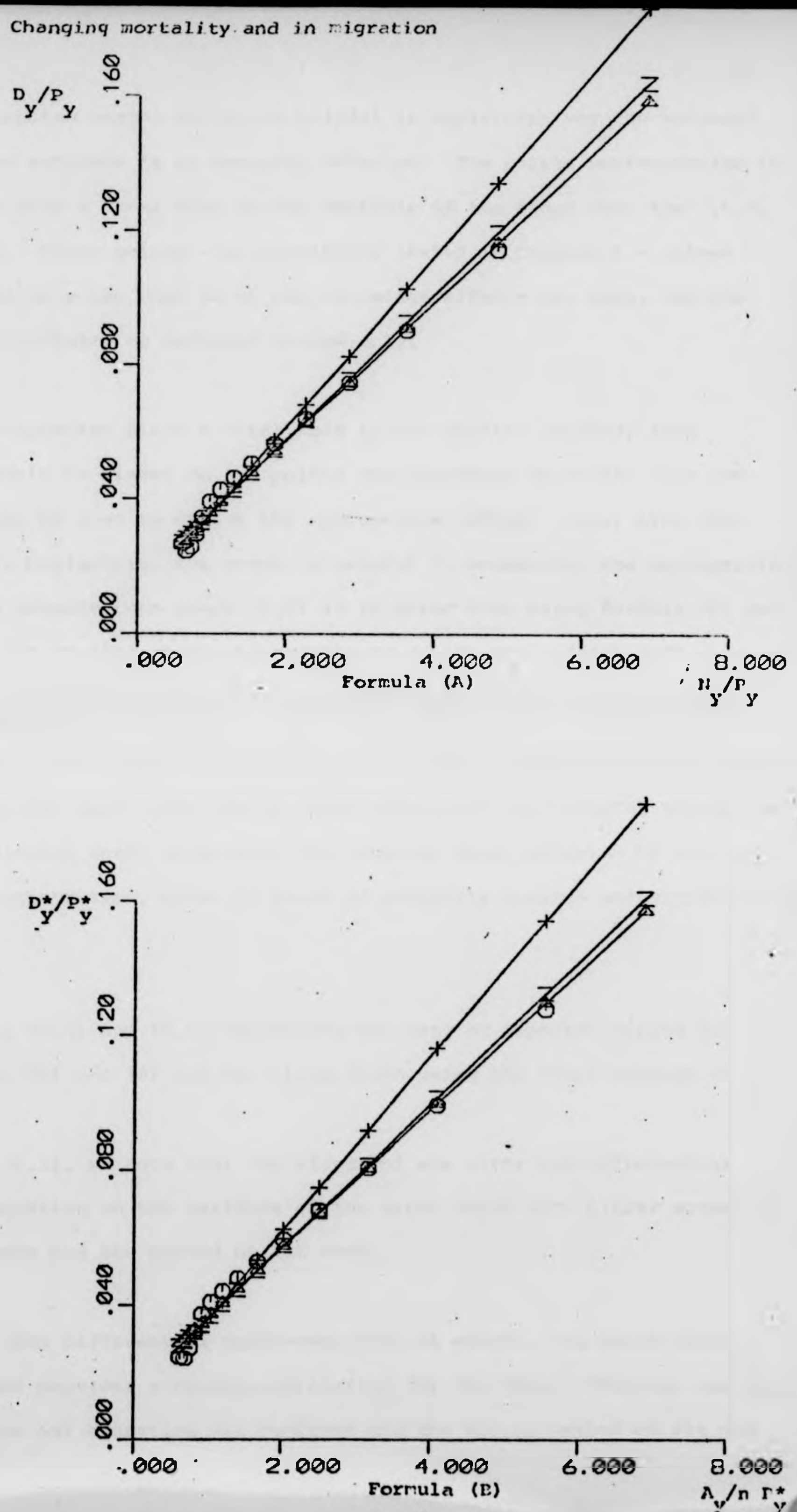




Graph (8.2) Changing Mortality and no Migration



Graph (8.3) Changing mortality and in migration



The corresponding graphs are quite helpful in explaining why the weighted least square estimate is so strongly affected. The points corresponding to middle ages play a vital role in the estimate of the slope when the W.L.S. fit is used. These points - as previously stated in Chapter 4 - either form a bulge or a gap when in or out migration affects the data, and the estimate is inflated or deflated accordingly.

Thus, when migration plays a vital role in the country studied, less emphasis should be placed on the points corresponding to middle ages and the graph may be used to choose the appropriate method. Also, even when migration is negligible, the graph is helpful in suggesting the appropriate method; for example from graph (8.2) it is clear that using formula (A) and the W.L.S. fit results in an under-estimate of the crude death rate.

B. Effect of deviations from stability, differential under-registration and misreport

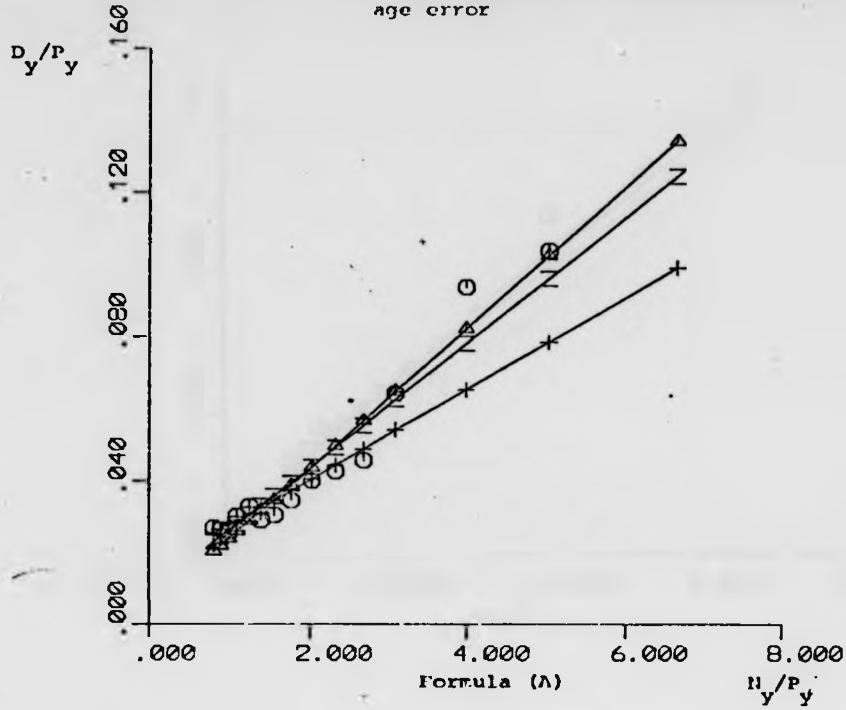
In Table (8.2), the actual death rate (total actual deaths/total population) and the reported death rate (total under-registered deaths/total population) and the estimated death rate using the reported data, affected by age error and under-registration, after 10 years of mortality decline and migration are presented.

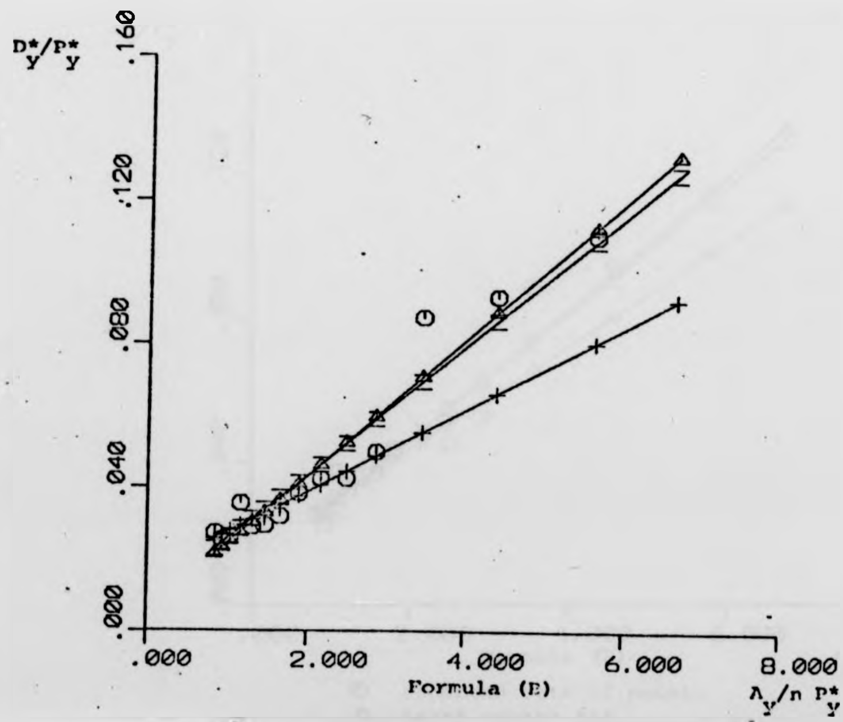
Graph (8.4), (8.5) and (8.6) represents the sets of reported points for both formula (A) and (B) and the lines drawn using the three methods of fit.

From Table (8.2), we note that the effect of age error and differential under-registration on the estimate of the crude death rate differ according to the formula and the method of fit used.

Generally, when differential under-registration exists, the death distribution method provides a minimum correction for the data. The only exception occurred when out migration was dominant and the W.L.S. method of fit was

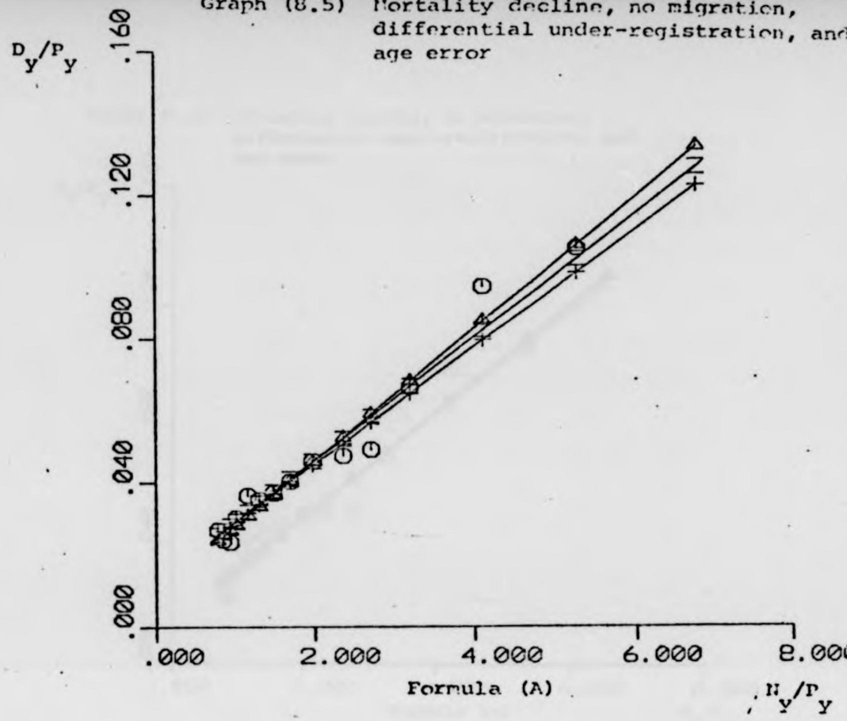
Graph (8.4) Mortality decline, out migration, differential under-registration, and age error

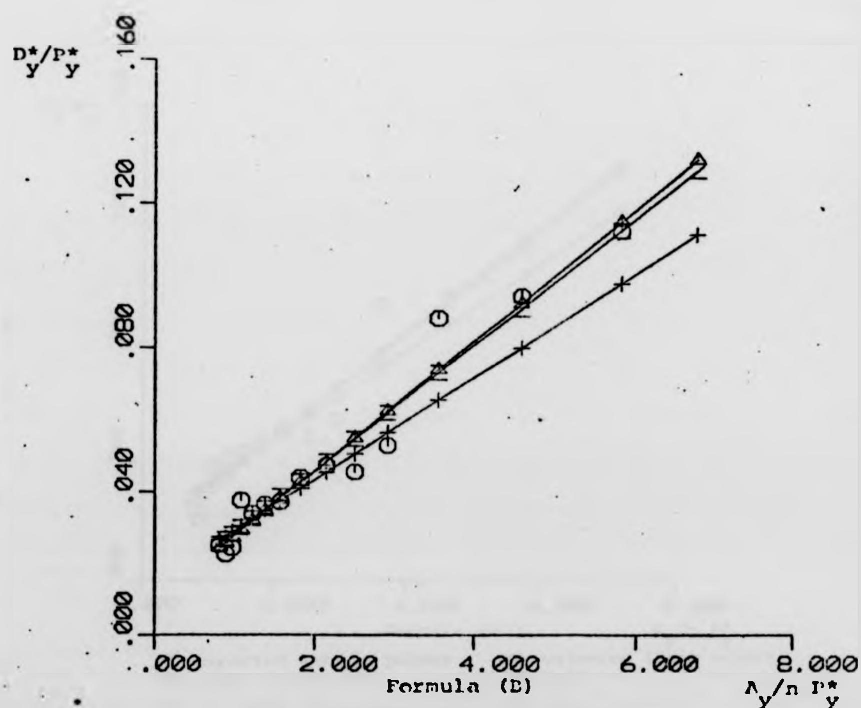




⊙ reported sets of points
△ least square fit
+ weighted least square
Z Wald

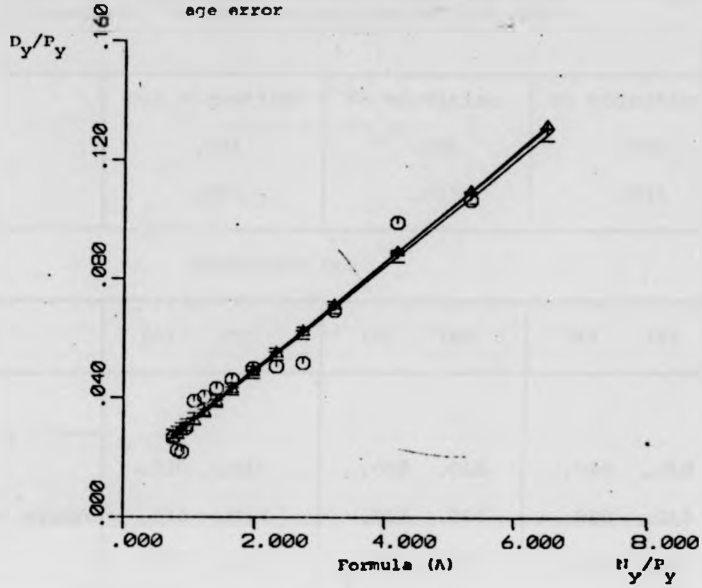
Graph (8.5) Mortality decline, no migration, differential under-registration, and age error





⊙ reported sets of points
△ least square fit
+ weighted least square
z Wald

Graph (8.6) Mortality decline, in migration, differential under-registration, and age error



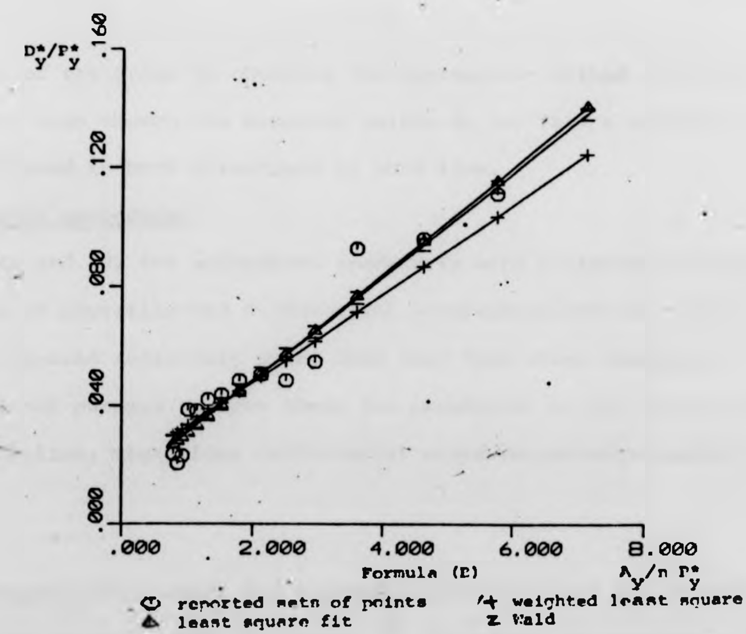


Table (8.2) The actual, reported and estimated death rate, after 10 years of mortality decline and migration, for data affected by differential under-registration and age error

	Out migration		No migration		In migration	
Actual CDR	.021		.021		.021	
Reported CDR	.015		.015		.015	
ESTIMATED CDR						
Formula	(A)	(B)	(A)	(B)	(A)	(B)
method of fit						
Least square	.020	.019	.018	.018	.018	.018
Weighted least square	.013	.013	.016	.015	.018	.015
Wald	.017	.018	.017	.017	.018	.017

used. The reason for this exception was discussed in the previous section.

The importance of the graph in choosing the appropriate method of fit is still apparent, even though the reported points do not form a straight line but are distributed in both directions of this line.

C. The adjustment procedures

In Chapters (4) and (5) two adjustment procedures were presented to allow for the effect of migration and differential under-registration. Each procedure was applied separately using data free from other sources of errors. It is our purpose to test these two procedures on the data affected by mortality decline, migration, differential under-registration and age error.

Table (8.3) presents the actual and reported death rates and the estimated death rate using the adjustment procedure for the effect of migration. The data on the proportions of the total population to the population in case of no in and out migration (P_x) are calculated using the reported population by age groups assuming in and out migration and the corresponding numbers in case of no migration.

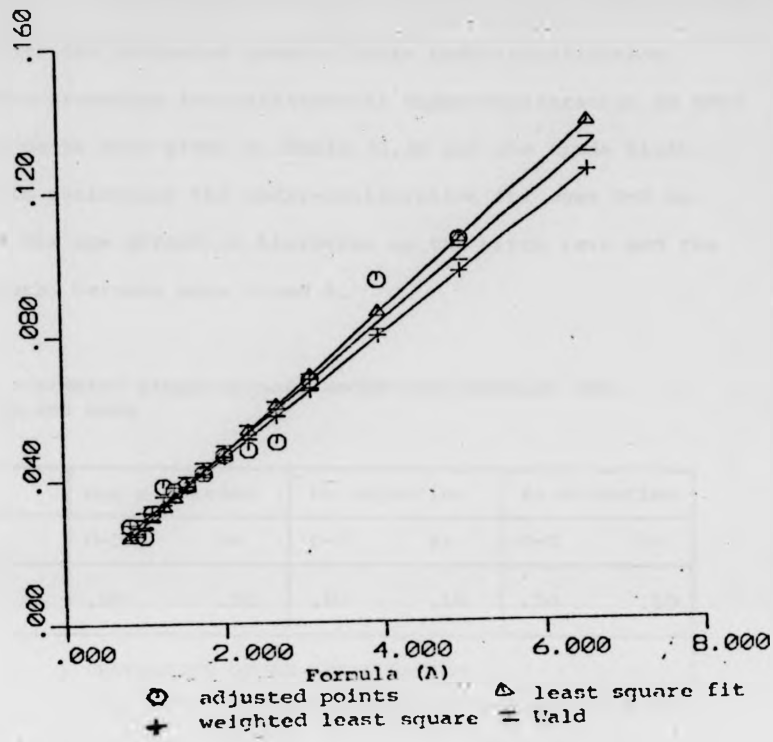
Graph (8.7) and (8.8) represent the adjusted sets of points for the effect of out and in migration and the lines drawn using the three methods of fit.

Table (8.3) The actual, reported and adjusted death rate for the effect of migration, after 10 years of mortality decline and migration, for data affected by differential under-registration and age error

	Out migration	In migration
Actual CDR	.021	.021
Reported CDR	.015	.015
ADJUSTED CDR		
Formula	(A)	(B)
method of fit		
Least square	.020	.018
Weighted Least Square	.017	.016
Wald	.018	.017

From Table (8.3), we conclude that the adjustment procedure for the effect of migration is acceptable and not very sensitive to age errors and deviations from assumptions. Also, the improvements in the adjusted sets of points given in Graph (8.7), (8.8) as compared to (8.4) and (8.6) is noticeable.

Graph (8.7) Adjustment for out migration



Graph (8.8) Adjustment for in migration

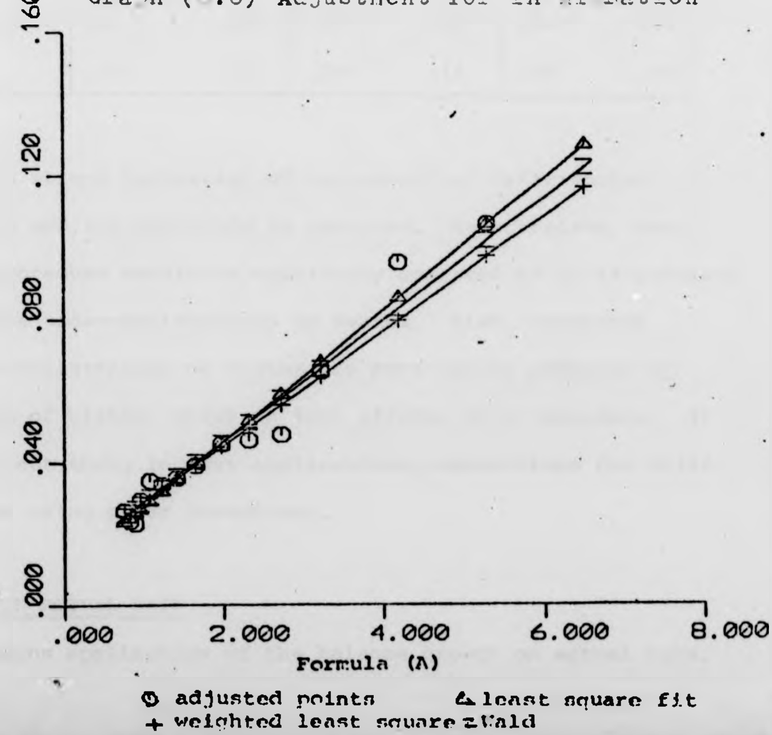


Table (8.4) presents the estimated proportionate under-registration when the adjustment procedure for differential under-registration is used with the adjusted death rate given in Table (8.3) and the crude birth rate. Note that in estimating the under-registration for ages 0-5 no allowance is made for the effect of migration on the birth rate and the proportions of deaths between ages 0 and 5.

Table (8.4) The estimated proportionate under-registration for different ages

	Out migration		No migration		In migration	
Ages	0-5	5+	0-5	5+	0-5	5+
Actual under-registration	.50	.10	.50	.10	.50	.10
ESTIMATED UNDER-REGISTRATION						
method of fit						
Least square	.52	.19	.63	.17	.68	.19
Weighted least square	.37	.08	.57	.07	.63	.07
Wald	.39	.13	.58	.11	.63	.12

From Table (8.4) a strong indication of the nature of differential under-registration and its magnitude is provided. Nevertheless, the results of this procedure should be cautiously accepted as it is possible to overestimate the under-registration of deaths. Also, countries affected by under-registration of deaths are more likely affected by under-registration of births, which in turn affects this procedure. It should be pointed out that, in many applications, corrections for child deaths may be made using other procedures.

8.3 APPLICATION ON ACTUAL DATA

This section contains application of the balance growth on actual data.

The data considered are for Guinea 1954-55. The analysis is divided into three parts; the first discusses the nature and characteristics of the data and the general techniques used, the second is a detailed study of the data, and finally the third part contains several cross checks and assessment of the results.

8.3.1 Data and General Techniques Used

The data analysed are from a sample inquiry covering Guinea (1954-55). This is the first large scale inquiry conducted in African territories formerly administered by France, and is one of the largest ones held.

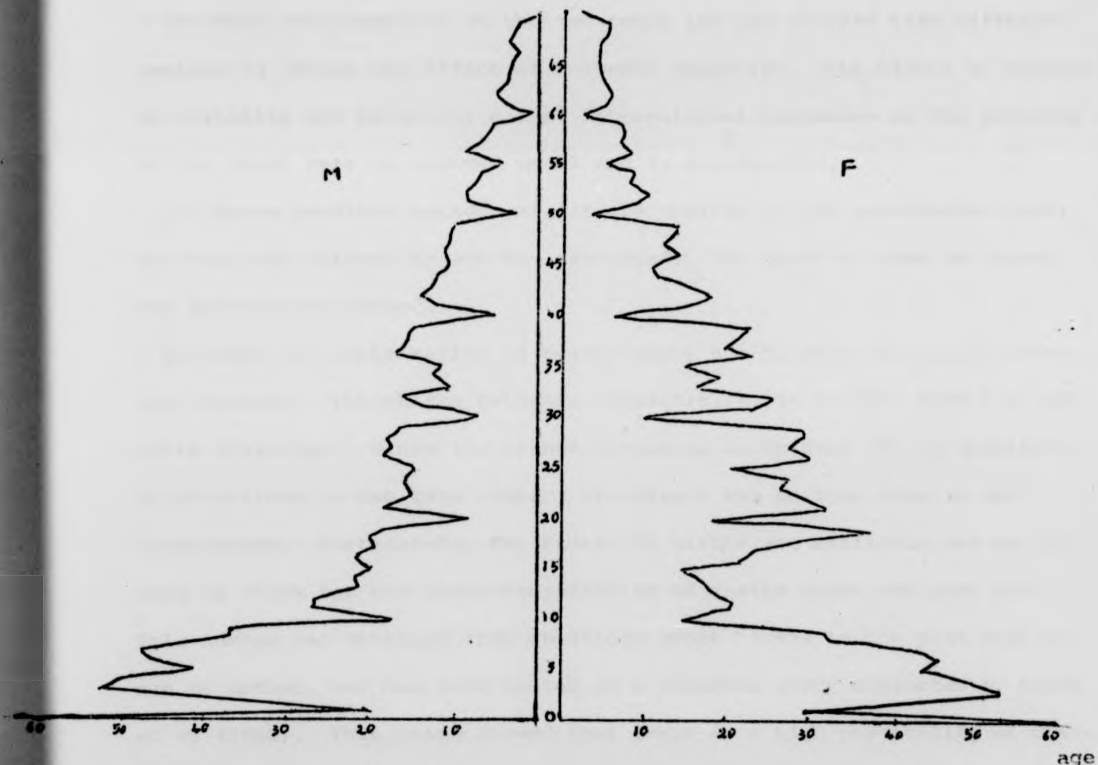
The available data for the estimation of mortality are of two kinds; current which are obtained from questions about deaths in the last year by age of deceased and retrospective consisting of reports by mothers, divided by age group, on the total number of children born to them and those still alive at the time of the survey. Both current and retrospective data are given separately for males and females and also for the four different regions of the country: Guinea Maritime, Fouta Djallon, upper Guinea and Forest.

It should be pointed out that the use of 'deaths in past year' reflects an important application of the growth balance method.

Though the data collection was through enumerators who attempted to check that the questions are understood and the answers reasonable; the data still suffer from several deficiencies associated with age. The age distribution by single years is reproduced in Graph (8.9). It shows an uncommon shunning from ages ending with 0 and 5; this is explained by the emphasis in training the enumerators against the general tendency of the population to round their ages with numbers ending with 0 and 5 which probably created a counter reaction to these digits. Another feature of the age distribution is the marked deficit in those aged 1 and 2 for both sexes. A final remark is the apparent deficiency of females aged 10 to 15 as compared to the neighbouring age groups.

Graph (8.9) The age distribution by single years

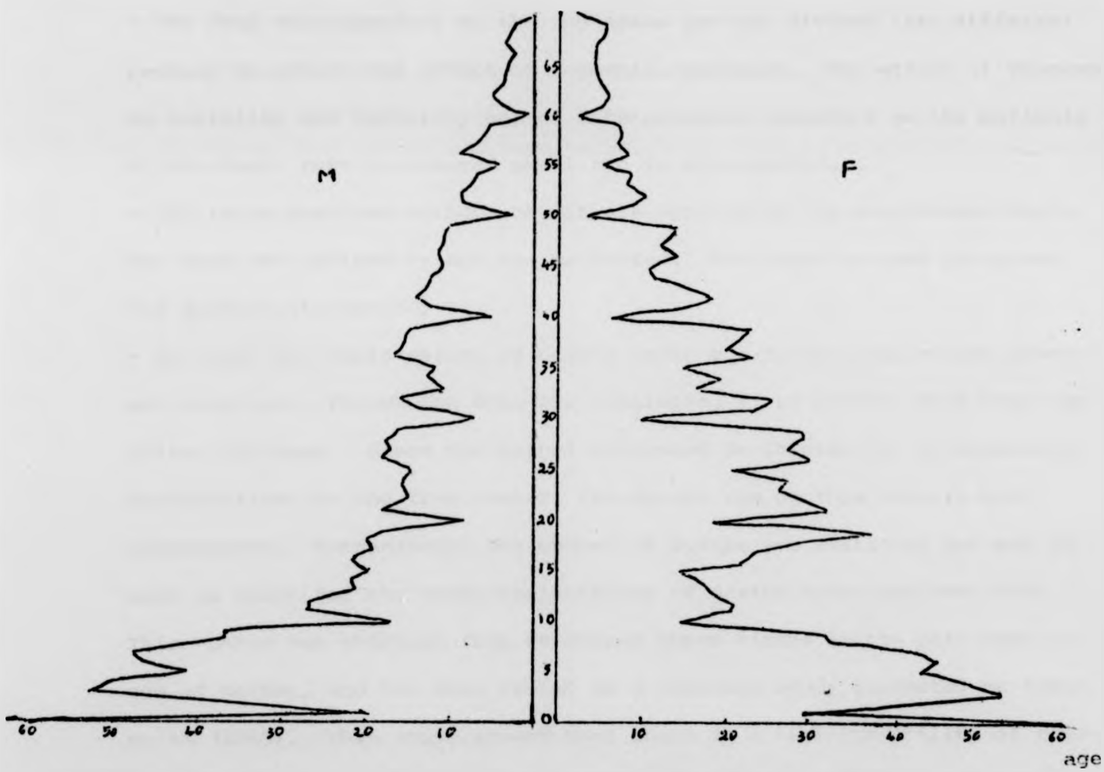
~~Graph (8.9) The age distribution by single years~~



the age distribution is the marked deficit in those aged 1 and 2 for both sexes. A final remark is the apparent deficiency of females aged 10 to 15 as compared to the neighbouring age groups.

Graph (8.9) The age distribution by single years

~~Graph (8.9) The age distribution by single years~~



This type of age error affecting the age distribution is similar to the bias already considered in the model discussed in Chapter (6). Actually instead of heaping on certain digits we have shunning from them but there is no reason for the overall effect to be different.

The data used in the following analysis are given in details in Tables (1), (2) and (3) of Appendix (d).

The procedure of analysis applied is described as follows:

- The data corresponding to the two sexes are not divided into different regions to offset the effect of internal migration. The effect of changes in mortality and fertility and of international migration on the estimate of the death rate is assumed small and is disregarded.
- The three previous methods of fit are applied to the ungraduated data; the data not defined by age are neglected. The graph is used to choose the appropriate method.
- To study the registration of deaths under age 5, data for single years are required. Though the data are available, it is evident that they are quite distorted. Since the method suggested in Chapter (5) is sensitive to deviations in the true number, its direct use on this data is not recommended. Fortunately, the number of births are available and may be used to check for the under-registration of deaths under one year old. This number was obtained from questions about births in the past year by age of mother, and has been tested in a separate study conducted by Brass et al (1968). This study showed that there is a high possibility of over-reporting of live births in the preceding year and suggested adjustment factors for the over-reporting of births in the four regions as follows:

Guinea Maritime	Fouta Djallon	Upper Guinea	Forest
.77	.82	.92	.73

These factors are accepted and the adjusted births are used to check for the under-registration under age 1.

- Once the under-registration for ages 0-1 and ages over 5 are accepted, it remains to check the under-registration for ages 1-4. Since we already rejected the data for single years, another procedure should be used. The relation between the deaths in infancy and at one to four years is not sufficiently strong to enable us to estimate one from the other. Nevertheless, if we agree that the registration from 1-4 years old is in the range of the under-registration within 0-1 and 5+ - which is not unlikely - the problem may be simpler. The relation between ${}_4p_1$ to ${}_1p_0$ - denoting the probability of surviving from age 1 to 4 and 0 to 1 respectively - under the assumption that the registration from 1-4 is either the same as ages over 5 or as ages from 0-1 is compared to the relation in the general standard life table introduced by Brass. (Brass, 1971). The value of ${}_4p_1$ which conform more closely with this standard is accepted as correct. The relations of probabilities of dying to the specific rates used are: ${}_1q_0 = \dfrac{{}_1m_0}{1 + 0.7\,{}_1m_0}$ and ${}_4q_0 = \dfrac{4\;{}_4m_1}{1 + 2.7\,{}_4m_1}$ where q_x and m_x denote probabilities of dying and specific death rates respectively.

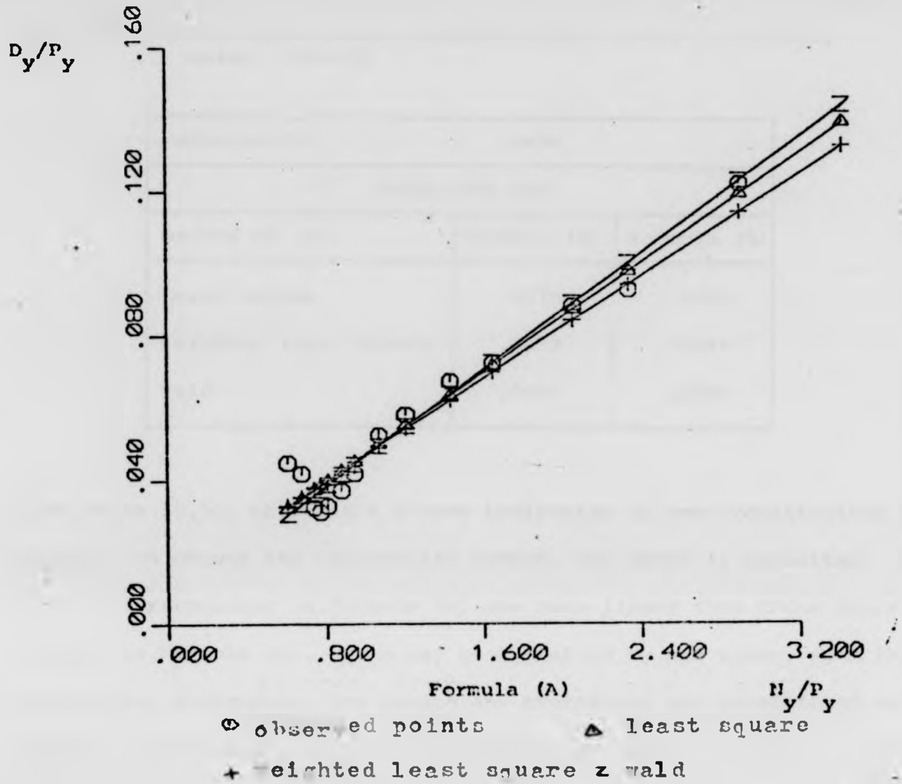
- Since the procedure adopted involve an arbitrary element, the corrections suggested are tested before being accepted. The results reached using each sex separately are compared to the results reached if the method is applied to the data of both sexes. Also, the estimates reached are checked against other estimates calculated by different analysis of the data.

8.3.2 Detailed Studies of the Data

a) males

In Table (8.5), the reported death rate and the estimated death rate using the reported data for males are presented. Graph (8.10) represents the sets of reported points for both formula (A) and (B) and the lines drawn using the three methods of fit.

Graph (8.1) Guinea, males



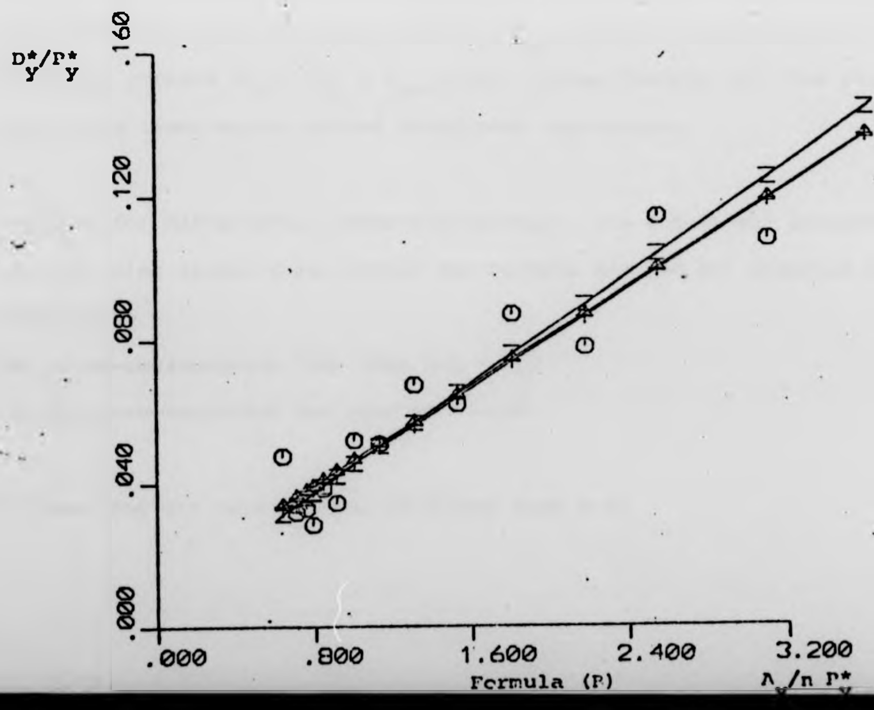


Table (8.5) The reported and estimated death rates for Guinea, males, 1954-55

reported CDR	.0456	
ESTIMATED CDR		
method of fit	Formula (A)	Formula (B)
Least square	.0378	.0345
Weighted least square	.0353	.0348
Wald	.0401	.0379

From Table (8.5), there is a strong indication of over-registration of deaths. To choose the appropriate method, the graph is consulted. The points corresponding to Formula (A) are more linear than those corresponding to Formula (B). This may be explained by age error, such that consecutive quinquenial age groups are overstated and understated respectively. For example, the male proportions in ages: 10-, 15-, 20-, 25-, 30- and 35-39 are reported as: .068, .094, .091, .099, .067 and .072. Formula (B) is more sensitive to this type of age error since it uses the proportions in each age group directly (A_y), while Formula (A) uses an averaging process ($N_y = (A_y + A_{y+n})/2n$). Using Formula (A), the fitted line using least square method seems more appropriate.

To check for differential under-registration, the adjustment procedure is applied using least square method and Formula (A) and the adjusted crude birth rate;

The under-registration from ages 0-1 = 0.0

The under-registration for ages 5+ = -.20

To check for the registration of deaths from 1-4:

the relation $\frac{{}_4D_1}{{}_1D_0}$ under several assumptions.

General Standard	Assumption (1) The registration from 1-4 the same as 5+ (over-registration 20%)	Assumption (2) The registration from 1-4 the same as 0-1 (No error)
1.086	1.091	1.052

Thus, we accept there is no error in stating the deaths at ages 0-1 while the deaths over age 1 are over-registered by 20%.

b) females

In Table (8.6), the reported death rate and the estimated death rate using the reported data for females are presented. Graph (8.11) represents the sets of reported points for both formula (A) and (B) and the lines drawn using the three methods of fit.

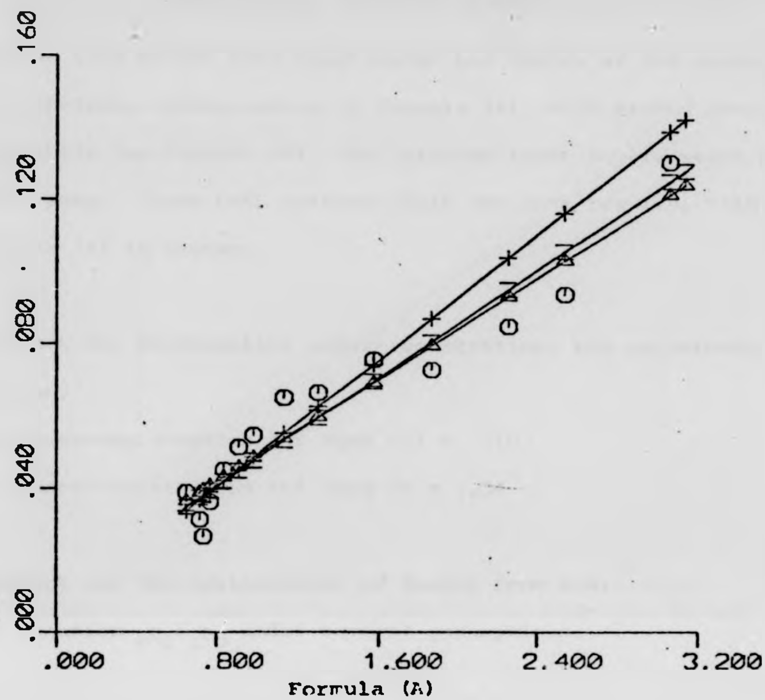
Table (8.6) The reported and estimated death rate for Guinea, females 1954-55

reported CDR	.0389	
ESTIMATED CDR		
method of fit	Formula (A)	Formula (B)
Least square	.0353	.0357
Weighted least square	.0437	.0376
Wald	.0374	.0353

All the previous methods, except weighted least square using Formula (A), suggest over-registration of deaths.

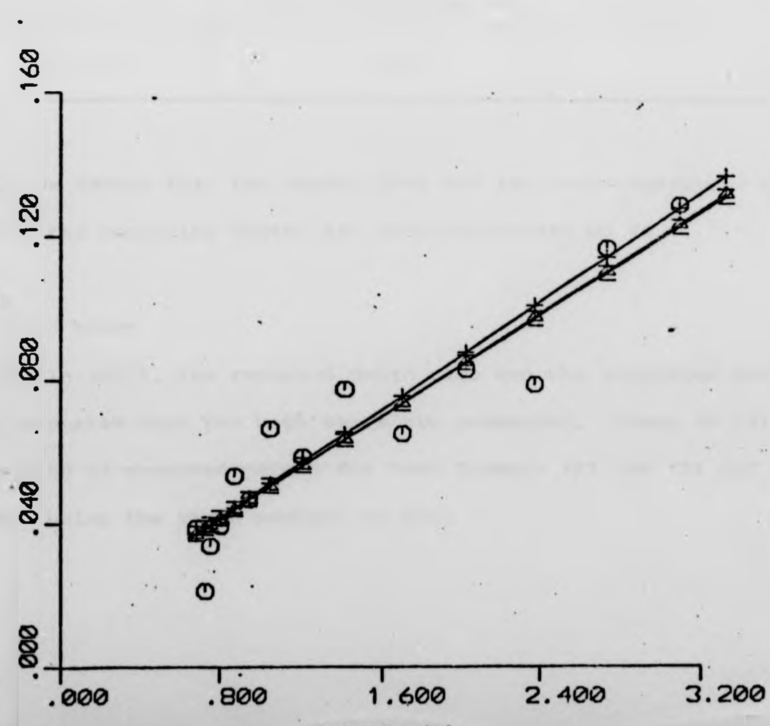
The deviation of the reported points from a straight line is stronger for

Graph (8.11) Guinea, females



⊖ observed points   △ least square

\+ weighted least square z Wald



females than males; this complicates the choice of the appropriate method. Nevertheless, corresponding to Formula (A), Wald method provides an average fit; while for Formula (B), the weighted least square seems appropriate for older ages. Since both methods yield the same results, Wald method using Formula (A) is chosen.

To check for differential under-registration, the adjustment procedure is applied,

The under-registration for ages 0-1 = -.10

The under-registration for ages 5+ = -.04

To check for the registration of deaths from 1-4:

the relation ${}_4p_1/{}_1p_0$ under several assumptions.

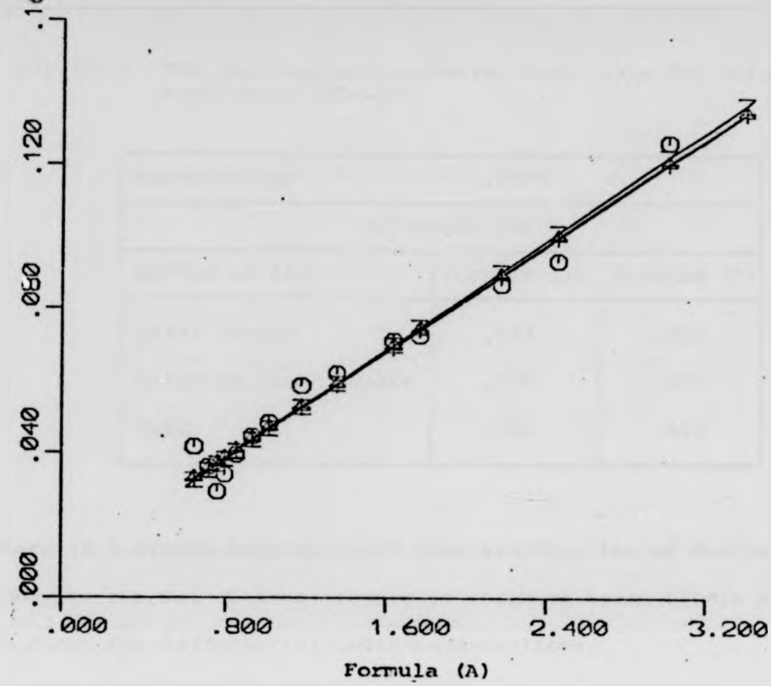
General Standard	Assumption (1) The registration from 1-4 the same as 5+ (over-registration 4%)	Assumption (2) The registration from 1-4 the same as 0-1 (over-registration 10%)
1.045	1.023	1.033

Thus we accept that the deaths from 0-4 are over-registered by 10% while the remaining deaths are over-registered by 4%.

c) both sexes

In Table (8.7), the reported death rate and the estimated death rate using the reported data for both sexes are presented. Graph (8.12) represents the sets of reported points for both formula (A) and (B) and the lines drawn using the three methods of fit.

Graph (8.12) Both Sexes



observed points   least square

weighted least square   Wald

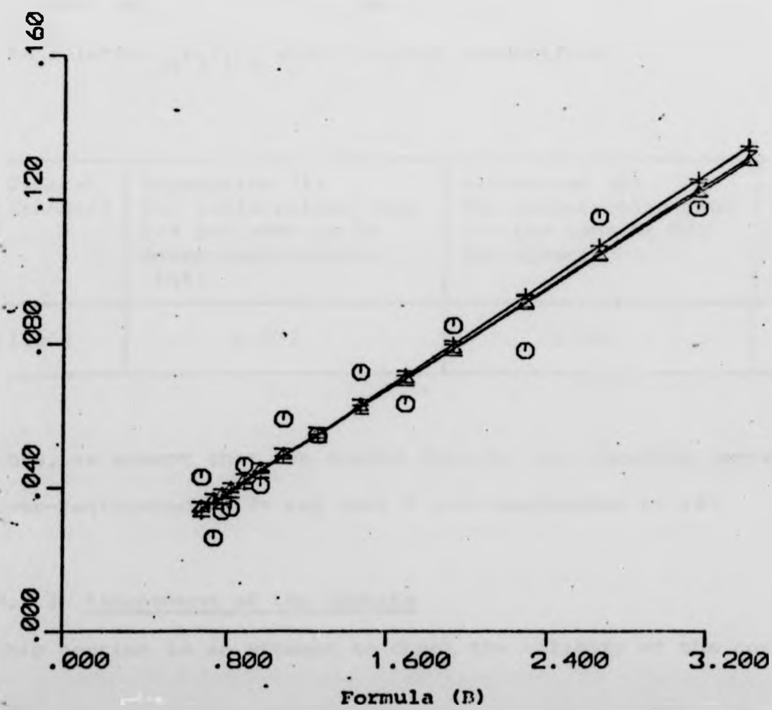


Table (8.7) The reported and estimated death rate for Guinea, both sexes 1954-55

reported CDR	.0421	
ESTIMATED CDR		
method of fit	Formula (A)	Formula (B)
Least square	.037	.035
Weighted least square	.037	.037
Wald	.038	.036

There is a strong indication of over-registration of deaths, Graph (8.12) suggests the use of least square or weighted least square and Formula (A).

To check for differential under-registration:

under-registration for ages 0.1 = 0.0

under-registration for ages 5+ = -.14

To check for the registration of deaths from 1-4:

the relation $_4P_1/_1P_0$ under several assumptions.

General Standard	Assumption (1) The registration from 1-4 the same as 5+ (over-registration 14%)	Assumption (2) The registration from 1-4 the same as 0-1 (No error)	Assumption (3) The registration from 1-4 as average from 0-1 and 5+ (over-registration 7%)
1.064	1.072	1.046	1.06

Thus, we accept that the deaths from 0-1 are reported correctly, from 1-4 over-registered by 7% and over 5 over-registered by 14%.

8.3.3 Assessment of the Results

This section is an attempt to check the validity of the corrections introduced in the previous parts.

In view of the fact that the reported points for males and both sexes exhibited more linearity than those for females and that both data suggested no error in reporting young deaths, one is suspicious of the estimated over-registration of female deaths aged 0-1. Also, the relation between the probabilities of dying for males and females assuming correct reporting is 1.2, the corresponding relation in the male and female standard is 1.2 while the relation under the assumption of over-registration of female young deaths is 1.4. Thus, it seems more likely that female deaths from ages 0-5 are reported correctly.

The suggested proportionate under-registration for different ages and sexes are summarized in Table (8.8).

Table (8.8) The estimated proportionate under-registration for different ages and sexes

age	males	females	both sexes
0-1	0.0	0.0	0.0
1-4	-.20	0.0	-.07
5+	-.20	-.04	-.14

A comparison between our suggested corrections and the results reached using a different procedure is very helpful. If both measures are similar, more weight is attached to the corrections suggested. In an earlier analysis of data for Guinea presented in Brass et al (1968), the retrospective reports of the proportions of children dead by age of mother were used to estimate the life table survivors at different ages. At the time of this analysis only provisional data were available. The following is an extract from the conclusions reached: 'In Guinea and its regions the differentials between reported current and retrospective childhood mortality are relatively small. Extremely high current death rates beyond childhood were recorded, particularly at ages 10 to 30 years, where the level is far above that for any other area. The pattern is reflected in the high values of β obtained when life tables from the model system are fitted to the observations. In the analysis of fertility it was shown that the P/F ratios for Guinea were low and the conclusion drawn that the births recorded were for a longer period than the preceding year. There seems a strong possibility that a similar lengthening of the reference period also occurred in the reporting of deaths but that it was offset, for young children at least, by omissions. Other evidence for such an effect exists.'

APPENDIX (A)

A.1 DEFINITIONS

M:	total number of age groups.
d_i:	actual number of deaths in age group i. i = 1, 2, ..., M.
u:	proportionate under-registration in age groups m to M ($0 \leq u < 1$). u = (under-registered deaths/actual deaths).
ou:	proportionate under-registration in the remaining age groups (1 to m).
m:	number of age groups experiencing under-report ou.
N_y:	actual and reported population proportion per year of age around the point y.
P_y:	actual and reported population proportion over age y.
D^r_y:	reported proportion of deaths over age y (reported deaths over age y/total reported deaths for all ages).
D_y:	actual proportion of deaths beyond age y.
Y_y:	N_y/P_y.
X^r_y:	D^r_y/P_y.
r:	growth rate.
CDR:	actual death rate.
X_y:	D_y/P_y

A.2 RESULTS

In section (4) we will prove that:

For i > m:

1. $Y_i = r + CDR^{`}.X^r_i$

where

$CDR^{`} = CDR.K(u, o)$

$$K(u,o) = 1 - \frac{u(o-1) \sum_{x=1}^{m} d_x}{(1-u) \sum_{x=1}^{M} d_x}$$

$$= 1 - \frac{u(o-1)}{(1-u) + (1-ou) \dfrac{D_m^r}{1-D_m^r}}$$

Also,

2. $u = 1 - \dfrac{\text{Total reported deaths}}{\text{CDR}^{\sim}.\text{Total population}}$

For $i < m$:

3. $v_i = \dfrac{(D_i^r - D_m^r)\, u(o-1)}{(1 - ou)}$

Also,

4. $v_i = \dfrac{Y_i - r}{\text{CDR}^{\sim}} \; P_i - D_i^r.$

A.3 METHOD

Using the reported population and deaths for age groups m to M, CDR˜, r and u are estimated using (1) and (2).

Using relation (3) and (4) o is estimated as:

$$o = \{\frac{v_i}{u} + (D_i^r - D_m^r)\} \frac{1}{(D_i^r - D_m^r + v_i)} .$$

Finally, the reported deaths are adjusted using u and o, thus:

$$\text{CDR} = (\frac{\text{Reported deaths from 1 to m}}{(1-ou)} + \frac{\text{Reported deaths from m to M}}{(1-u)} \text{ / total population}$$

or

CDR = CDR`/K(u,o).

A.4 PROOF

For $i > m$:

1. $$D_i = \frac{\sum_{x=i}^{M} d_x}{\sum_{x=1}^{M} d_x} = \frac{\sum_{x=i}^{M} d_x(1-u)}{\sum_{x=1}^{M} d_x(1-u)} . \qquad (a.1)$$

$$D_i^r = \frac{\sum_{x=i}^{M} d_x(1-u)}{\sum_{x=1}^{m} d_x(1-ou) + \sum_{x=m}^{M} d_x(1-u)}$$

$$D_i^r = \frac{\sum_{x=i}^{M} d_x(1-u)}{\sum_{x=1}^{m} d_x(1-u+u-ou) + \sum_{x=m}^{M} d_x(1-u)}$$

then

$$D_i^r = \frac{\sum_{x=i}^{M} d_x(1-u)}{\sum_{x=1}^{M} d_x(1-u) + u(1-o\ ) \sum_{x=1}^{m} d_x}$$

dividing the nominator and denominator by $\sum_{x=1}^{M} d_x(1-u)$ and using (a.1) we get:

$$D_i^r = \frac{D_i}{1 - \dfrac{u(o-1) \sum_{x=1}^{m} d_x}{(1-u) \sum_{x=1}^{M} d_x}}$$

Thus,

$$D_i^r = D_i / K(u,o) \tag{a.2}$$

where

$$K(u,o) = 1 - \frac{u(o-1) \sum_{x=1}^{m} d_x}{(1-u) \sum_{x=1}^{M} d_x}$$

but

$$x_i^r = \frac{D_i^r}{P_i}$$

then

$$X_i^r = \frac{D_i}{K(u,o)P_i} = \frac{X_i}{K(u,o)} \tag{a.3}$$

since

$$Y_i = r + CDR.X_i$$

using (a.3)

$$Y_i = r + CDR.K(u,o).X_i^r.$$

Finally,

$$Y_i = r + CDR^{\backprime}.X_i^r$$

where $\quad CDR^{\backprime} = CDR.K(u,o)$

$$K(u,o) = 1 - \frac{u(o-1) \sum_{x=1}^{m} d_x}{(1-u) \sum_{x=1}^{M} d_x}.$$

To show that $K(u,o)$ may be re-expressed in terms of the reported deaths as:

$$K(u,o) = 1 - \frac{u(o-1)}{(1-u) + (1-ou)\left(\dfrac{D_m^r}{1-D_m^r}\right)}$$

since

$$K(u,o) = 1 - \frac{u(o-1) \sum_{x=1}^{m} d_x}{(1-u) \sum_{x=1}^{M} d_x}$$

$$= 1 - \frac{u(o-1) \sum_{x=1}^{m} d_x}{(1-u)\left[\sum_{x=1}^{m} d_x + \sum_{x=m}^{M} d_x\right]}$$

$$= 1 - \frac{u(o-1)}{(1-u)\left[1 + \frac{\sum_{x=m}^{M} d_x}{\sum_{x=1}^{m} d_x}\right]}$$

$$= 1 - \frac{u(o-1)}{(1-u)\left[1 + \frac{(1-ou)}{(1-u)} \cdot \frac{\sum_{x=m}^{M} d_x(1-u)}{\sum_{x=1}^{m} d_x(1-ou)}\right]}$$

and

$$K(u,o) = 1 - \frac{u(o-1)}{(1-u) + (1-ou)\frac{D_m^r}{1 - D_m^r}}$$

This completes the proof of (A.1).

2. Using the reported number of deaths for ages over m, and the relation (2.3)

$$\frac{n_y}{P_y} = r + f\,\frac{d_y^r}{P_y}$$

where f is the ratio of the true deaths over age m to the reported deaths over age m.

Thus, $f = \dfrac{\text{true deaths over age m}}{\text{true deaths over age m - under-registered deaths over age m}}$

$$f = \frac{1}{1-u}$$

since

$$Y_i = r + CDR^{*}.X_i^r.$$

Then

$$\frac{CDR^{*}\ (\text{total population})}{(\text{total reported deaths})} = \frac{1}{1-u}$$

and

$$u = 1 - \frac{\text{total reported deaths}}{CDR^{*}.\text{total population}}\ .$$

For i < m:

3. $$D_i^r = \frac{\sum_{x=i}^{m} d_x(1-o\ u) + \sum_{x=m}^{M} d_x(1-u)}{\sum_{x=1}^{m} d_x(1-o\ u) + \sum_{x=m}^{M} d_x(1-u)} \qquad \text{(a.4)}$$

$$D_i^r = \frac{\sum_{x=i}^{m} d_x(1-u+u-ou) + \sum_{x=m}^{M} d_x(1-u)}{\sum_{x=1}^{m} d_x(1-u+u-ou) + \sum_{x=m}^{M} d_x(1-u)}$$

$$= \frac{\sum_{x=i}^{M} d_x(1-u) - u(o-1) \sum_{x=i}^{m} d_x}{\sum_{x=1}^{M} d_x(1-u) - u(o-1) \sum_{x=1}^{m} d_x}$$

dividing the nominator and denominator by $(1-u) \sum_{x=1}^{M} d_x$, we get:

$$D_i^r = \frac{D_i - \dfrac{u(o-1)\sum_{x=i}^{m} d_x}{(1-u)\sum_{x=1}^{M} d_x}}{1 - \dfrac{u(o-1)\sum_{x=1}^{m} d_x}{(1-u)\sum_{x=1}^{M} d_x}}$$

$$D_i^r = \frac{D_i}{K(u,o)} - v_i$$

Thus

$$v_i = \frac{u(o-1)\sum_{x=i}^{m} d_x}{k(u,o)(1-u)\sum_{x=1}^{M} d_x} \tag{a.5}$$

$$v_i = \frac{D_i}{K(u,o)} - D_i^r \tag{a.6}$$

Rewriting (a.5)

$$v_i = \frac{\dfrac{u(o-1)\sum_{x=i}^{m} d_x}{(1-u)\sum_{x=1}^{M} d_x}}{1 - \dfrac{u(o-1)\sum_{x=1}^{m} d_x}{(1-u)\sum_{x=1}^{M} d_x}}$$

Multiplying the nominator and denominator by $\dfrac{(1-u)\sum_{x=1}^{M} d_x}{u(o-1)\sum_{x=1}^{m} d_x}$, we get

$$v_i = \cfrac{\cfrac{\sum_{x=i}^{m} d_x}{\sum_{x=1}^{m} d_x}}{\cfrac{(1-u)\sum_{x=1}^{M} d_x}{u(o-1)\sum_{x=1}^{m} d_x} - 1}$$

$$v_i = \cfrac{\cfrac{\sum_{x=i}^{m} d_x(1-ou)}{\sum_{x=1}^{m} d_x(1-ou)}}{\cfrac{(1-u)}{u(o-1)}\left\{1 + \cfrac{(1-ou)\sum_{x=m}^{M} d_x(1-u)}{(1-u)\sum_{x=1}^{m} d_x(1-ou)}\right\} - 1} \tag{a.7}$$

but

$$\frac{\sum_{x=i}^{m} d_x(1-ou)}{\sum_{x=1}^{m} d_x(1-ou)} = \frac{\left[\sum_{x=i}^{m} d_x(1-ou) + \sum_{x=m}^{M} d_x(1-u)\right] - \sum_{x=m}^{M} d_x(1-u)}{\left[\sum_{x=1}^{m} d_x(1-ou) + \sum_{x=m}^{M} d_x(1-u)\right] - \sum_{x=m}^{M} d_x(1-u)}$$

$$= \frac{D_i^r - D_m^r}{1 - D_m^r}$$

substituting in (a.7)

$$v_i = \cfrac{\cfrac{D_i^r - D_m^r}{1 - D_m^r}}{\cfrac{(1-u)}{u(o-1)}\left\{1 + \cfrac{(1-ou)D_m^r}{(1-u)(1-D_m^r)}\right\} - 1}$$

$$= \frac{(D_i^r - D_m^r)}{\frac{(1-u)(1-D_m^r)}{u(o-1)} + \frac{(1-ou)D_m^r}{u(o-1)} - (1-D_m^r)}$$

and finally,

$$v_i = \frac{(D_i^r - D_m^r)u(o-1)}{(1-ou)} .$$

4. From (a.6)

$$v_i = \frac{D_i}{K(u,o)} - D_i^r$$

$$\left(\frac{Y_i - r}{CDR}\right) = \left(\frac{(Y_i - r)}{CDR.K(u,o)}\right) = \frac{X_i}{K(o,u)} = \frac{D_i}{P_i.K(o,u)}$$

then

$$\frac{D_i}{K(o,u)} = \left(\frac{Y_i - r}{CDR}\right).P_i$$

and

$$v_i = \frac{Y_i - r}{CDR} \; P_i - D_i^r.$$

APPENDIX (B)

Table (F.1) The effect of age error - when both the population and death distribution are subject to the same kind of age error - on the age and death distribution and on the estimate of the crude death rate.

SINGLE YEAR AGE GROUP

AGE DISTRIBUTION

AGE	STABLE AGE DIST.	RANDOM AGE DIST.	%STABLE AGE DIST.	%RANDOM AGE DIST.
0=	3971.1	4431.1	3.97	4.43
1=	3340.3	2519.3	3.34	2.52
2=	3085.9	3734.9	3.09	3.73
3=	2938.3	2584.3	2.94	2.58
4=	2825.5	3383.5	2.83	3.38
5=	2742.6	2426.6	2.74	2.43
6=	2678.6	2742.6	2.68	2.74
7=	2615.8	2225.8	2.62	2.23
8=	2554.3	3216.3	2.55	3.22
9=	2494.1	1802.1	2.49	1.80
10=	2438.5	3106.5	2.44	3.11
11=	2387.3	1571.3	2.39	1.57
12=	2337.2	1705.2	2.34	1.71
13=	2288.0	1778.0	2.29	1.78
14=	2239.7	1727.7	2.24	1.73
15=	2190.1	1716.1	2.19	1.72
16=	2139.2	1830.2	2.14	1.83
17=	2089.3	1880.3	2.09	1.88
18=	2040.5	1936.5	2.04	1.94
19=	1992.6	1947.6	1.99	1.95
20=	1942.7	3135.7	1.94	3.14
21=	1890.7	1914.7	1.89	1.91
22=	1839.9	2573.9	1.84	2.57
23=	1790.2	1854.2	1.79	1.85
24=	1741.5	2281.5	1.74	2.28
25=	1693.4	2040.4	1.69	2.04
26=	1645.8	1956.8	1.65	1.96
27=	1599.3	1451.3	1.60	1.45
28=	1553.7	1950.7	1.55	1.95
29=	1509.2	1090.2	1.51	1.09
30=	1464.9	1977.9	1.46	1.98
31=	1420.7	1042.7	1.42	1.04
32=	1377.5	1672.5	1.38	1.67
33=	1335.3	1088.3	1.34	1.09
34=	1294.1	1385.1	1.29	1.39
35=	1252.9	1236.9	1.25	1.24
36=	1211.9	1259.9	1.21	1.26
37=	1171.8	942.8	1.17	0.94
38=	1132.6	1317.6	1.13	1.32
39=	1094.4	753.4	1.09	0.75
40=	1056.0	1431.0	1.06	1.43
41=	1017.4	713.4	1.02	0.71
42=	979.9	1089.9	0.98	1.09
43=	943.2	792.2	0.94	0.79
44=	907.3	983.3	0.91	0.98
45=	872.0	857.0	0.87	0.86
46=	837.1	877.1	0.84	0.88
47=	803.0	647.0	0.80	0.65
48=	769.8	905.8	0.77	0.91
49=	737.4	519.4	0.74	0.52
50=	704.7	925.7	0.70	0.93
51=	671.9	456.9	0.67	0.46
52=	639.8	460.8	0.64	0.46
53=	608.6	454.6	0.61	0.45

Table (D.1) (continued)

54	578.1	447.1	0.58	0.45
55	548.1	414.1	0.55	0.41
56	518.7	442.7	0.52	0.44
57	490.0	441.0	0.49	0.44
58	462.0	449.0	0.46	0.45
59	434.8	416.8	0.43	0.42
60	407.5	664.5	0.41	0.66
61	380.2	406.2	0.38	0.41
62	353.6	528.6	0.35	0.53
63	327.7	355.7	0.33	0.36
64	302.5	415.5	0.30	0.42
65	278.4	354.4	0.28	0.35
66	255.3	316.3	0.26	0.32
67	232.8	214.8	0.23	0.21
68	210.9	272.9	0.21	0.27
69	189.6	143.6	0.19	0.14
70	170.0	252.0	0.17	0.25
71	152.1	108.1	0.15	0.11
72	134.6	157.6	0.13	0.16
73	117.7	104.7	0.12	0.10
74	101.2	136.2	0.10	0.14
75	87.1	107.1	0.09	0.11
76	75.2	89.2	0.08	0.09
77	63.7	64.7	0.06	0.06
78	52.6	73.6	0.05	0.07
79	41.7	39.7	0.04	0.04
80	134.4	277.4	0.13	0.28
TOTAL	100000.0	100000.0		

DEATH DISTRIBUTION

AGE	STABLE DEATH DIST.	RANDOM DEATH DIST.	%STABLE DEATH DIST.	%RANDOM
0	1314.3	1288.3	38.83	38.06
1	289.0	264.0	8.54	7.80
2	124.7	186.7	3.68	5.51
3	80.5	60.5	2.38	1.79
4	58.9	77.9	1.74	2.30
5	23.6	16.6	0.70	0.49
6	23.2	21.2	0.69	0.63
7	22.9	22.9	0.68	0.68
8	22.5	25.5	0.67	0.75
9	22.2	14.2	0.66	0.42
10	15.1	25.1	0.45	0.74
11	14.9	10.9	0.44	0.32
12	14.6	10.6	0.43	0.31
13	14.4	10.4	0.43	0.31
14	14.2	10.2	0.42	0.30
15	18.6	13.6	0.55	0.40
16	18.3	16.3	0.54	0.48
17	18.0	20.0	0.53	0.59
18	17.7	12.7	0.52	0.38
19	17.5	17.5	0.52	0.52
20	23.4	29.4	0.69	0.87
21	23.0	24.0	0.68	0.71
22	22.7	26.7	0.67	0.79
23	22.3	27.3	0.66	0.81
24	22.0	27.0	0.65	0.80
25	22.7	20.7	0.67	0.61
26	22.4	25.4	0.66	0.75

27=	22.0	18.0	0.65	0.53
28=	21.7	29.7	0.64	0.88
29=	21.4	14.4	0.63	0.43
30=	22.7	25.7	0.67	0.76
31=	22.3	16.3	0.66	0.48
32=	22.0	23.0	0.65	0.68
33=	21.7	14.7	0.64	0.43
34=	21.4	28.4	0.63	0.84
35=	22.7	20.7	0.67	0.61
36=	22.4	26.4	0.66	0.78
37=	22.1	18.1	0.65	0.53
38=	21.7	27.7	0.64	0.82
39=	21.4	20.4	0.63	0.60
40=	23.1	29.1	0.68	0.86
41=	22.8	16.8	0.67	0.50
42=	22.4	29.4	0.66	0.87
43=	22.1	23.1	0.65	0.68
44=	21.8	20.8	0.64	0.61
45=	22.2	27.2	0.66	0.80
46=	21.9	18.9	0.65	0.56
47=	21.6	18.6	0.64	0.55
48=	21.3	27.3	0.63	0.81
49=	21.0	16.0	0.62	0.47
50=	22.7	26.7	0.67	0.79
51=	22.4	21.4	0.66	0.63
52=	22.0	17.0	0.65	0.50
53=	21.7	11.7	0.64	0.35
54=	21.4	21.4	0.63	0.63
55=	21.6	15.6	0.64	0.46
56=	21.3	19.3	0.63	0.57
57=	21.0	23.0	0.62	0.68
58=	20.7	16.7	0.61	0.49
59=	20.4	17.4	0.60	0.51
60=	21.6	31.6	0.64	0.93
61=	21.3	23.3	0.63	0.69
62=	20.9	24.9	0.62	0.74
63=	20.6	21.6	0.61	0.64
64=	20.3	25.3	0.60	0.75
65=	19.3	16.3	0.57	0.48
66=	19.0	15.0	0.56	0.44
67=	18.7	21.7	0.55	0.64
68=	18.4	15.4	0.54	0.46
69=	18.1	13.1	0.54	0.39
70=	15.6	22.6	0.46	0.67
71=	15.4	12.4	0.46	0.37
72=	15.2	12.2	0.45	0.36
73=	15.0	8.0	0.44	0.24
74=	14.7	13.7	0.44	0.41
75=	10.7	9.7	0.32	0.29
76=	10.5	9.5	0.31	0.28
77=	10.4	9.4	0.31	0.28
78=	10.2	11.2	0.30	0.33
79=	10.1	7.1	0.30	0.21
80=	34.7	54.7	1.02	1.61
TOTAL=	3385.0	3385.0		

Table (F.1) (continued)

FIVE YEAR AGE GROUP

AGE	STABLE AGE DIST.	RANDOM AGE DIST.	%STABLE AGE DIST.	%RANDOM AGE DIST.
0=	16161.0	16653.0	16.16	16.65
5=	13085.4	12413.4	13.09	12.41
10=	11690.6	9888.6	11.69	9.89
15=	10451.7	9310.7	10.45	9.31
20=	9205.1	11760.1	9.21	11.76
25=	8001.4	8489.4	8.00	8.49
30=	6892.5	7166.5	6.89	7.17
35=	5863.6	5510.6	5.86	5.51
40=	4903.7	5009.7	4.90	5.01
45=	4019.3	3806.3	4.02	3.81
50=	3203.0	2745.0	3.20	2.75
55=	2453.7	2163.7	2.45	2.16
60=	1771.6	2370.6	1.77	2.37
65=	1166.9	1301.9	1.17	1.30
70=	675.6	758.6	0.68	0.76
75=	320.3	374.3	0.32	0.37
80=	134.4	277.4	0.13	0.28
OTAL=	100000.0	100000.0		

DEATH DISTRIBUTION

AGE	STABLE DEATH DIST.	RANDOM DEATH DIST.	%STABLE DEATH DIST.	%RANDOM DIST.
0=	1867.2	1877.2	55.16	55.46
5=	114.4	100.4	3.38	2.97
10=	73.2	67.2	2.16	1.98
15=	90.1	80.1	2.66	2.37
20=	113.4	134.4	3.35	3.97
25=	110.3	108.3	3.26	3.20
30=	110.1	108.1	3.25	3.19
35=	110.3	113.3	3.26	3.35
40=	112.2	119.2	3.32	3.52
45=	108.0	108.0	3.19	3.19
50=	110.3	98.3	3.26	2.90
55=	104.9	91.9	3.10	2.71
60=	104.7	126.7	3.09	3.74
65=	93.5	81.5	2.76	2.41
70=	75.9	68.9	2.24	2.04
75=	51.9	46.9	1.53	1.39
80=	34.7	54.7	1.02	1.61
OTAL=	3385.0	3385.0		

Table (F.1) (continued)

FORMULA (A)

AGE	XS	YS	XR	YR
5	0.53482	0.03488	0.55443	0.03487
10	0.56595	0.03502	0.58613	0.03146
15	0.66534	0.03749	0.64858	0.03145
20	0.75366	0.04044	0.71958	0.04073
25	0.84470	0.04366	0.83196	0.05066
30	0.95619	0.04743	0.95470	0.04973
35	1.09239	0.05204	1.10475	0.05213
40	1.26110	0.05774	1.25044	0.05594
45	1.46978	0.06492	1.44914	0.06389
50	1.74922	0.07426	1.68195	0.06557
55	2.10877	0.08673	1.91847	0.06774
60	2.61098	0.10385	2.20112	0.08921
65	3.29211	0.12792	2.74481	0.13541
70	4.24731	0.16300	3.57164	0.14610
75	5.62530	0.21901	4.60484	0.17383

FORMULA (B)

AGE	XS	YS	XR	YR
5	0.55822	0.03386	0.55320	0.03218
10	0.62267	0.03602	0.61501	0.02997
15	0.70521	0.03883	0.68115	0.03302
20	0.79442	0.04183	0.76857	0.05129
25	0.89415	0.04520	0.85604	0.04752
30	1.01589	0.04931	1.02009	0.05137
35	1.16529	0.05434	1.16829	0.05111
40	1.34965	0.06055	1.33453	0.06146
45	1.53557	0.06850	1.54692	0.06400
50	1.89355	0.07885	1.78138	0.06370
55	2.30478	0.09267	2.03499	0.07020
60	2.86188	0.11132	2.37029	0.12165
65	3.60711	0.13618	3.02767	0.12632
70	4.64263	0.17049	3.87819	0.14716
75	6.07928	0.21746	4.96751	0.16114

ACTUAL DEATH RATE .03385

LEAST SQUARE FIT

ACTUAL DEATH RATE(1) .03578   ESTIMATED(1) .03654

ACTUAL DEATH RATE(2) .03334   ESTIMATED(2) .03200

(1) stable distribution: corresponding to model west, males, mortality level 6, growth rate = 15% given in Coale & Demeny (1966)

(2) random distribution: the resulting distribution when the model of error is applied to the stable distribution

Table (F.2) The effect of age error - when the population and death distribution are subject to different age error - on the age and death distribution and on the estimate of the crude death rate.

SINGLE YEAR AGE GROUP

AGE DISTRIBUTION

AGE	STABLE AGE DIST.	RANDOM AGE DIST.	%STABLE AGE DIST.	%RANDOM AGE DIST.
0-	3971.3	4444.3	3.97	4.44
1-	3340.4	2588.4	3.34	2.59
2-	3086.0	3643.0	3.09	3.64
3-	2938.4	2405.4	2.94	2.41
4-	2825.6	3379.6	2.83	3.38
5-	2742.8	2734.8	2.74	2.73
6-	2678.7	2740.7	2.68	2.74
7-	2615.9	2084.9	2.62	2.08
8-	2554.4	3161.4	2.55	3.16
9-	2494.2	1837.2	2.49	1.84
10-	2438.6	3121.6	2.44	3.12
11-	2387.4	1618.4	2.39	1.62
12-	2337.3	1736.3	2.34	1.74
13-	2288.1	1745.1	2.29	1.75
14-	2239.8	1714.8	2.24	1.71
15-	2190.2	1790.2	2.19	1.79
16-	2139.3	1819.3	2.14	1.82
17-	2089.4	1888.4	2.09	1.89
18-	2040.6	1947.6	2.04	1.95
19-	1992.7	1881.7	1.99	1.88
20-	1942.8	3106.8	1.94	3.11
21-	1890.8	1946.8	1.89	1.95
22-	1840.0	2661.0	1.84	2.66
23-	1790.3	1738.3	1.79	1.74
24-	1741.6	2146.6	1.74	2.15
25-	1693.5	2227.5	1.69	2.23
26-	1645.9	1980.9	1.65	1.98
27-	1599.3	1399.3	1.60	1.40
28-	1553.8	1918.8	1.55	1.92
29-	1509.3	1082.3	1.51	1.08
30-	1464.9	2004.9	1.46	2.00
31-	1420.8	1006.8	1.42	1.01
32-	1377.6	1687.6	1.38	1.69
33-	1335.4	1002.4	1.34	1.00
34-	1294.2	1367.2	1.29	1.37
35-	1253.0	1339.0	1.25	1.34
36-	1212.0	1306.0	1.21	1.31
37-	1171.9	918.9	1.17	0.92
38-	1132.7	1358.7	1.13	1.36
39-	1094.4	795.4	1.09	0.80
40-	1056.0	1379.0	1.06	1.38
41-	1017.5	716.5	1.02	0.72
42-	979.9	1144.9	0.98	1.14
43-	943.2	696.2	0.94	0.70
44-	907.4	987.4	0.91	0.99
45-	872.0	958.0	0.87	0.96
46-	837.1	895.1	0.84	0.90
47-	803.1	612.1	0.80	0.61
48-	769.8	911.8	0.77	0.91
49-	737.4	527.4	0.74	0.53
50-	704.8	906.8	0.70	0.91
51-	671.9	448.9	0.67	0.45
52-	639.8	471.8	0.64	0.47
53-	608.6	456.6	0.61	[illegible]

Table (E.2) (continued)

54=	578.1	436.1	0.58	0.44
55=	548.2	439.2	0.55	0.44
56=	518.7	428.7	0.52	0.43
57=	490.0	421.0	0.49	0.42
58=	462.1	431.1	0.46	0.43
59=	434.8	426.8	0.43	0.43
60=	407.5	634.5	0.41	0.63
61=	380.2	410.2	0.38	0.42
62=	353.6	550.6	0.35	0.55
63=	327.7	369.7	0.33	0.37
64=	302.5	400.5	0.30	0.40
65=	278.4	356.4	0.28	0.36
66=	255.3	290.3	0.26	0.29
67=	232.8	203.8	0.23	0.20
68=	210.9	246.9	0.21	0.25
69=	189.6	142.6	0.19	0.14
70=	170.0	233.0	0.17	0.23
71=	152.1	119.1	0.15	0.12
72=	134.6	157.6	0.13	0.16
73=	117.7	96.7	0.12	0.10
74=	101.2	119.2	0.10	0.12
75=	87.1	106.1	0.09	0.11
76=	75.2	100.2	0.08	0.10
77=	63.7	67.7	0.06	0.07
78=	52.6	69.6	0.05	0.07
79=	41.7	42.7	0.04	0.04
80=	129.8	299.8	0.13	0.30
TOTAL=	100000.0	100000.0		

DEATH DISTRIBUTION

AGE	STABLE DEATH DIST.	RANDOM DEATH DIST.	%STABLE DEATH DIST.	%RANDOM
0=	1314.3	1271.3	38.83	37.56
1=	289.0	288.0	8.54	8.51
2=	124.7	170.7	3.68	5.04
3=	80.5	68.5	2.38	2.02
4=	58.9	62.9	1.74	1.86
5=	23.6	25.6	0.70	0.76
6=	23.2	22.2	0.69	0.66
7=	22.9	21.9	0.68	0.65
8=	22.5	27.5	0.67	0.81
9=	22.2	17.2	0.66	0.51
10=	15.1	24.1	0.45	0.71
11=	14.9	9.9	0.44	0.29
12=	14.6	14.6	0.43	0.43
13=	14.4	14.4	0.43	0.43
14=	14.2	14.2	0.42	0.42
15=	18.6	11.6	0.55	0.34
16=	18.3	15.3	0.54	0.45
17=	18.0	18.0	0.53	0.53
18=	17.7	13.7	0.52	0.41
19=	17.5	16.5	0.52	0.49
20=	23.4	31.4	0.69	0.93
21=	23.0	20.0	0.68	0.59
22=	22.7	28.7	0.67	0.85
23=	22.3	22.3	0.66	0.66
24=	22.0	25.0	0.65	0.74
25=	22.7	24.7	0.67	0.73
26=	22.4	34.4	0.66	1.02

Table (1.2) (continued)

27=	22.1	16.1	0.65	0.47
28=	21.7	24.7	0.64	0.73
29=	21.4	15.4	0.63	0.45
30=	22.7	27.7	0.67	0.82
31=	22.3	19.3	0.66	0.57
32=	22.0	26.0	0.65	0.77
33=	21.7	11.7	0.64	0.35
34=	21.4	20.4	0.63	0.60
35=	22.7	21.7	0.67	0.64
36=	22.4	26.4	0.66	0.78
37=	22.1	19.1	0.65	0.56
38=	21.7	28.7	0.64	0.85
39=	21.4	13.4	0.63	0.40
40=	23.1	35.1	0.68	1.04
41=	22.8	16.8	0.67	0.50
42=	22.4	27.4	0.66	0.81
43=	22.1	22.1	0.65	0.65
44=	21.8	22.8	0.64	0.67
45=	22.2	32.2	0.66	0.95
46=	21.9	17.9	0.65	0.53
47=	21.6	16.6	0.64	0.49
48=	21.3	18.3	0.63	0.54
49=	21.0	19.0	0.62	0.56
50=	22.7	28.7	0.67	0.85
51=	22.4	17.4	0.66	0.51
52=	22.1	17.1	0.65	0.50
53=	21.7	12.7	0.64	0.38
54=	21.4	18.4	0.63	0.54
55=	21.6	18.6	0.64	0.55
56=	21.3	18.3	0.63	0.54
57=	21.0	15.0	0.62	0.44
58=	20.7	22.7	0.61	0.67
59=	20.4	18.4	0.60	0.54
60=	21.6	29.6	0.64	0.87
61=	21.3	19.3	0.63	0.57
62=	20.9	24.9	0.62	0.74
63=	20.6	16.6	0.61	0.49
64=	20.3	17.3	0.60	0.51
65=	19.3	21.3	0.57	0.63
66=	19.0	26.0	0.56	0.77
67=	18.7	13.7	0.55	0.40
68=	18.4	22.4	0.54	0.66
69=	18.1	10.1	0.54	0.30
70=	15.6	24.6	0.46	0.73
71=	15.4	[illegible]	0.46	0.78
72=	15.2	14.2	0.45	0.42
73=	15.0	12.0	0.44	0.35
74=	14.7	16.7	0.44	0.49
75=	10.7	17.7	0.32	0.52
76=	10.5	13.5	0.31	0.49
77=	10.4	7.4	0.31	0.22
78=	10.2	10.2	0.30	0.30
79=	10.1	6.1	0.30	0.18
80=	34.7	51.7	1.02	1.53
TOTAL=	3385.2	3385.2		

Table (E.2) (continued)

FIVE YEAR AGE GROUP

AGE	STABLE AGE DIST.	RANDOM AGE DIST.	%STABLE AGE DIST.	%RANDOM AGE DIST.
0=	16161.3	16460.8	16.16	16.46
5=	13086.1	12559.1	13.09	12.56
10=	11691.2	9936.2	11.69	9.94
15=	10452.2	9327.2	10.45	9.33
20=	9205.5	11599.5	9.21	11.60
25=	8001.3	8608.8	8.00	8.61
30=	6892.8	7068.8	6.89	7.07
35=	5863.9	5717.9	5.86	5.72
40=	4904.0	4924.0	4.90	4.92
45=	4019.4	3904.4	4.02	3.90
50=	3203.2	2720.2	3.20	2.72
55=	2453.8	2146.8	2.45	2.15
60=	1771.7	2374.7	1.77	2.37
65=	1166.9	1239.9	1.17	1.24
70=	675.6	725.6	0.68	0.73
75=	320.3	386.3	0.32	0.39
80=	129.8	299.8	0.13	0.30
OTAL=	100000.0	100000.0		

DEATH DISTRIBUTION

AGE	STABLE DEATH DIST.	RANDOM DEATH DIST.	%STABLE DEATH DIST.	%RANDOM DIST
0=	1867.3	1861.3	55.16	54.98
5=	114.4	114.4	3.38	3.38
10=	73.2	77.2	2.16	2.28
15=	90.1	75.1	2.66	2.22
20=	113.4	127.4	3.35	3.76
25=	110.3	115.3	3.26	3.41
30=	110.1	105.1	3.25	3.10
35=	110.3	109.3	3.26	3.23
40=	112.3	124.3	3.32	3.67
45=	108.0	104.0	3.19	3.07
50=	110.3	94.3	3.26	2.78
55=	104.9	92.9	3.10	2.74
60=	104.7	107.7	3.09	3.18
65=	93.5	93.5	2.76	2.76
70=	75.9	76.9	2.24	2.27
75=	51.9	54.9	1.53	1.62
80=	34.7	51.7	1.02	1.53
OTAL=	3385.2	3385.2		

Table (F.2) (continued)

FORMULA (A)

AGE	XS	YS	XR	YR
5	0.53483	0.03489	0.53887	0.03474
10	0.58597	0.03502	0.58659	0.03169
15	0.66537	0.03749	0.64473	0.03156
20	0.75371	0.04044	0.71813	0.04046
25	0.84479	0.04367	0.83196	0.05037
30	0.95632	0.04743	0.95119	0.04976
35	1.09258	0.05205	1.09929	0.05232
40	1.26141	0.05775	1.26254	0.05684
45	1.47028	0.06494	1.44798	0.06398
50	1.75008	0.07430	1.70771	0.06696
55	2.11835	0.08679	1.96707	0.06785
60	2.62213	0.10397	2.26133	0.08996
65	3.29915	0.12317	3.08643	0.13632
70	4.26575	0.16368	3.84098	0.13923
75	5.68598	0.22128	4.59124	0.16207

FORMULA (B)

AGE	XS	YS	XR	YR
5	0.55824	0.03386	0.56079	0.03251
10	0.62210	0.03602	0.61347	0.03010
15	0.70525	0.03383	0.67839	0.03309
20	0.79449	0.04184	0.76735	0.05052
25	0.89425	0.04520	0.88441	0.04803
30	1.01605	0.04931	1.01589	0.05054
35	1.16553	0.05435	1.17010	0.05299
40	1.35803	0.06057	1.34034	0.06057
45	1.58621	0.06853	1.55592	0.06592
50	1.89469	0.07890	1.81672	0.06376
55	2.30691	0.09275	2.08331	0.07039
60	2.86630	0.11143	2.54629	0.12371
65	3.61747	0.13655	3.34858	0.12206
70	4.67140	0.17150	4.08635	0.13836
75	6.18092	0.22096	4.74414	0.15674

ACTUAL DEATH RATE .03385

LEAST SQUARE FIT

ACTUAL DEATH RATE(1) .03580   ESTIMATED(1) .03296

ACTUAL DEATH RATE(2) .03335   ESTIMATED(2) .03036

(1) stable distribution: corresponding to model west, males, mortality level 6, growth rate = 15% given in Coale & Demeny (1966).

(2) random distribution: the resulting distribution when the model of error is applied to the stable distribution.

APPENDIX (C)

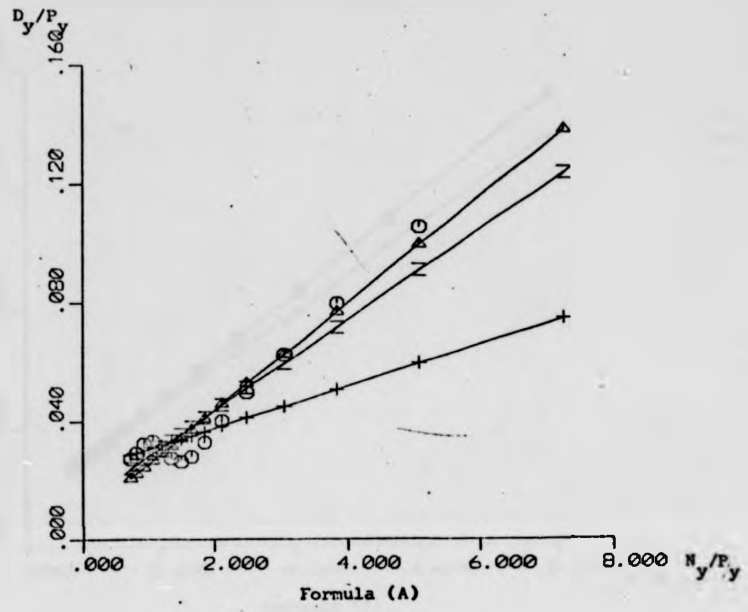
Table (C.1) The actual and estimated death rate after 20 years of mortality decline and migration

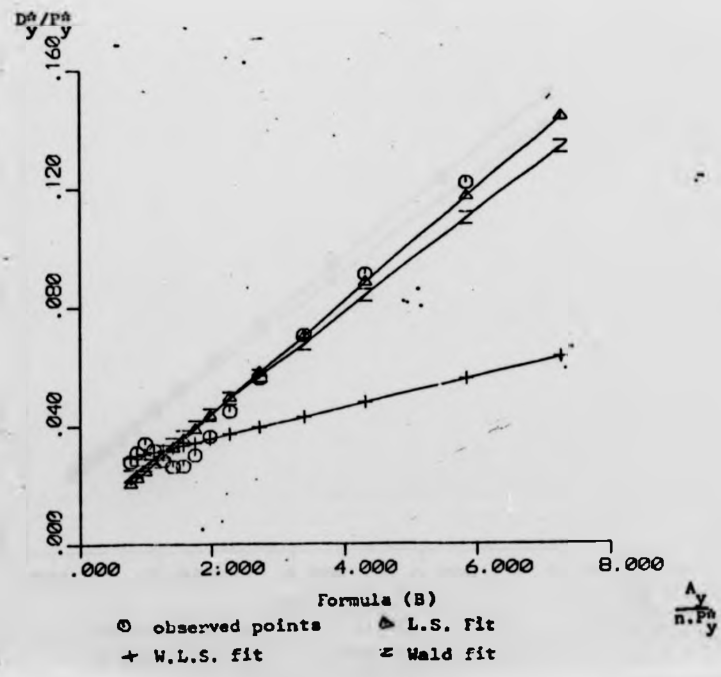
	Out migration	No migration	In migration
Actual CDR	.019	.018	.017

ESTIMATED CDR

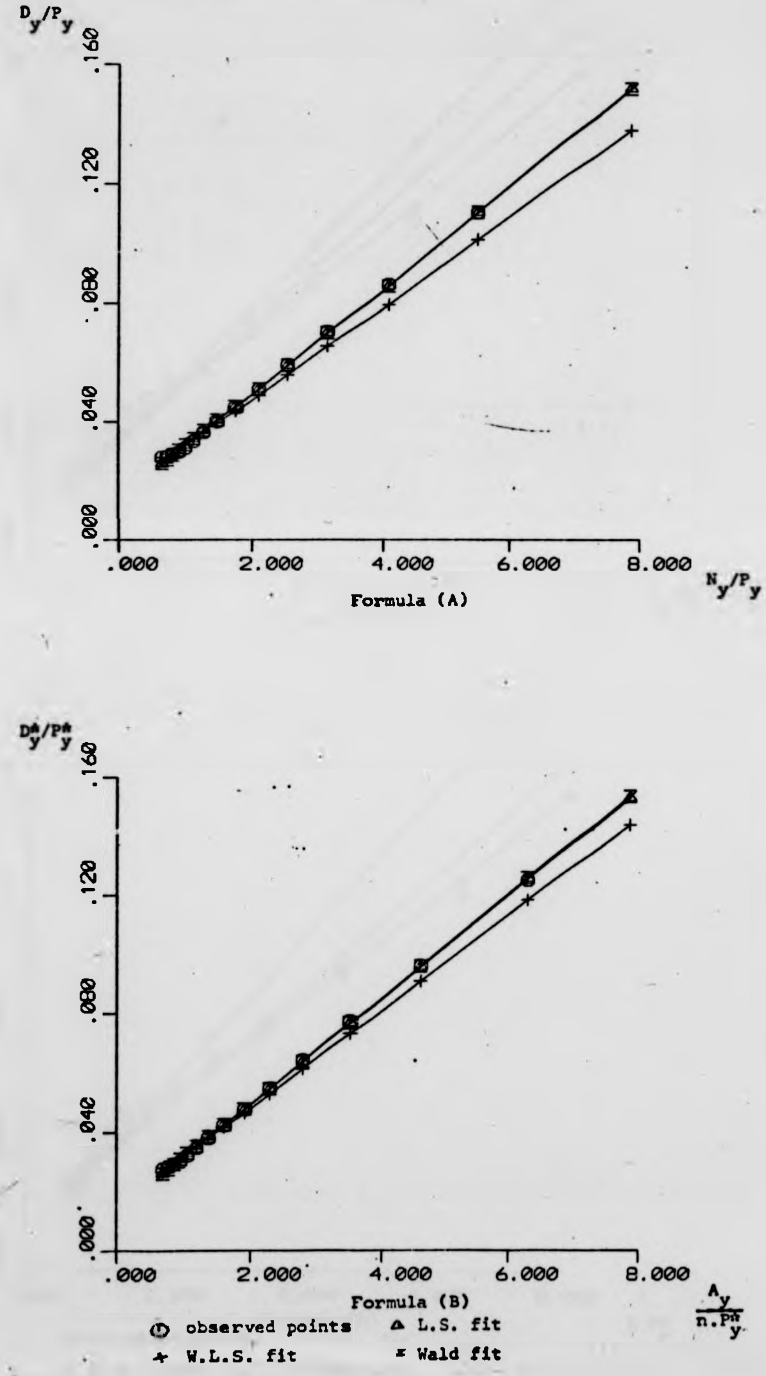
Formula	(A)	(B)	(A)	(B)	(A)	(B)
Method of fit						
Least square	.018	.019	.018	.018	.018	.017
Weighted least square	.007	.005	.016	.016	.023	.025
Wald	.015	.017	.018	.018	.020	.019

Graph (C.1)
Changing mortality and out migration

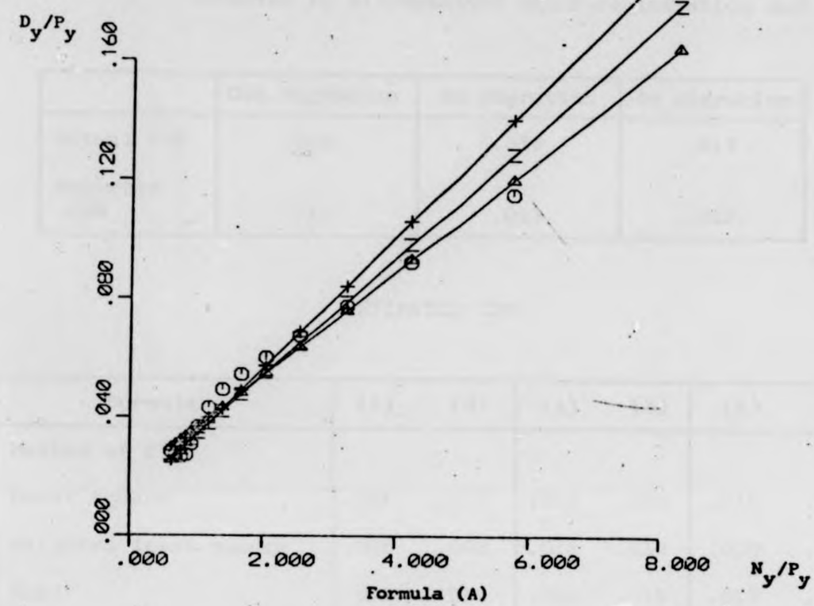




Graph (C.2) Changing mortality and no migration



Graph (C.3) Changing mortality and in migration



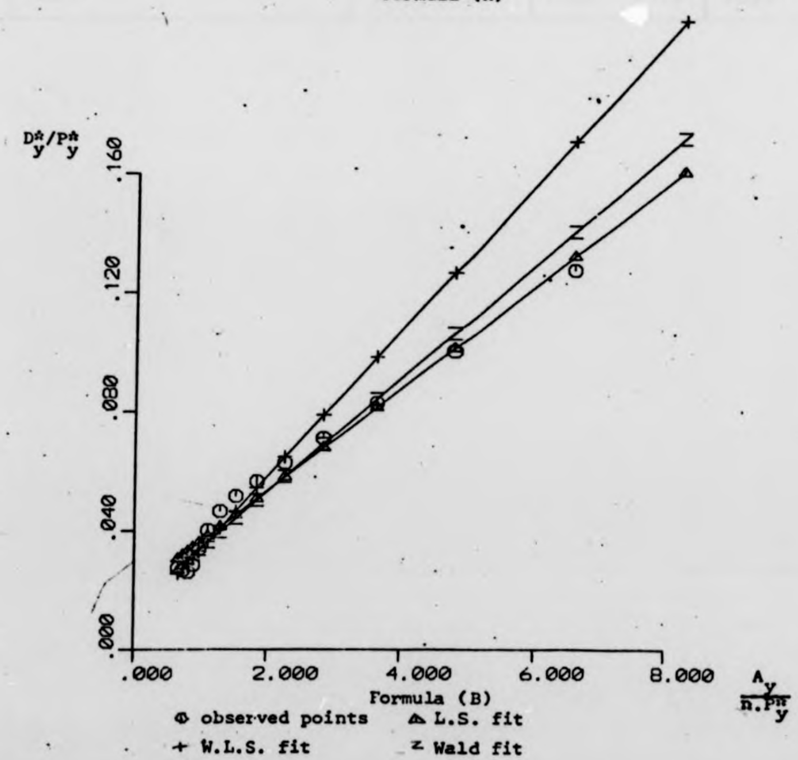


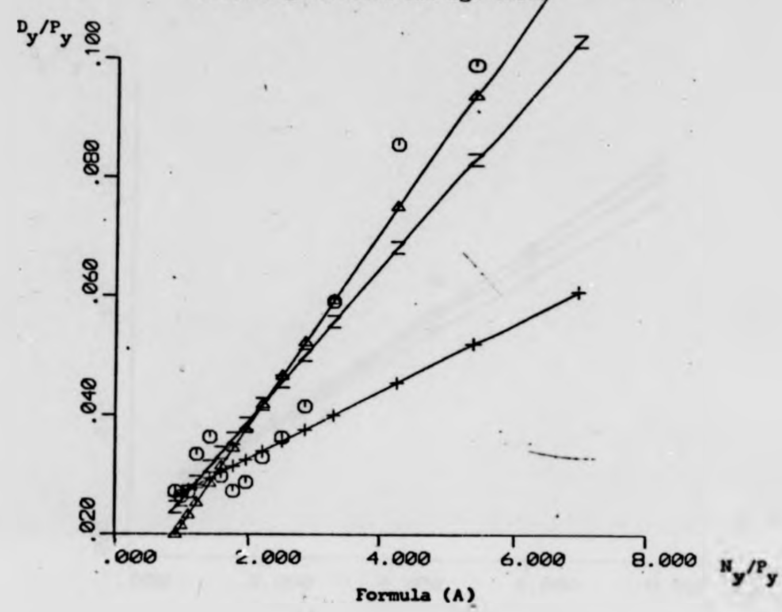
Table (C.2) The actual, reported and estimated death rate, after 20 years of mortality decline and migration, for data affected by differential under-registration and age error

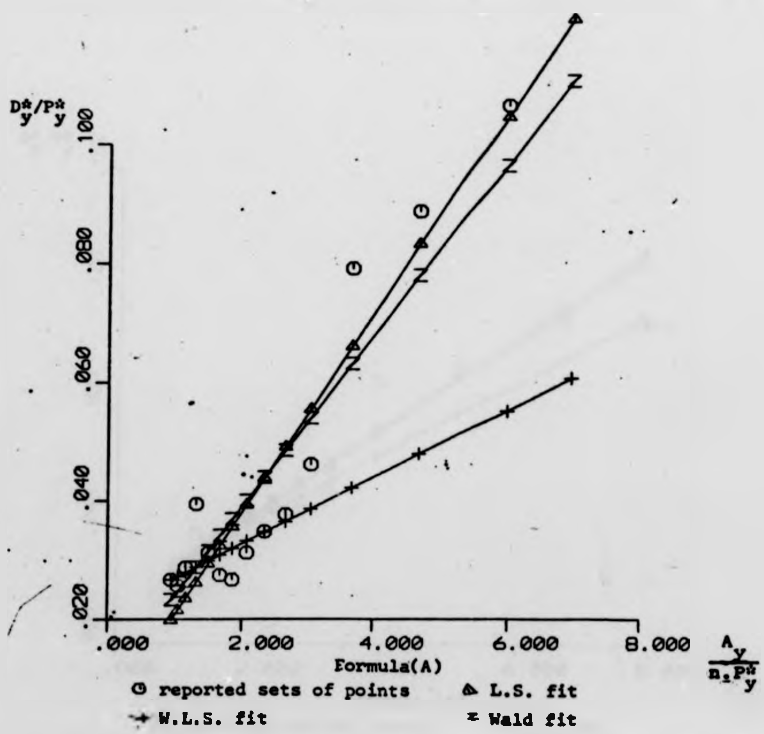
	Out migration	No migration	In migration
Actual CDR	.019	.018	.017
Reported CDR	.015	.013	.012

ESTIMATED CDR

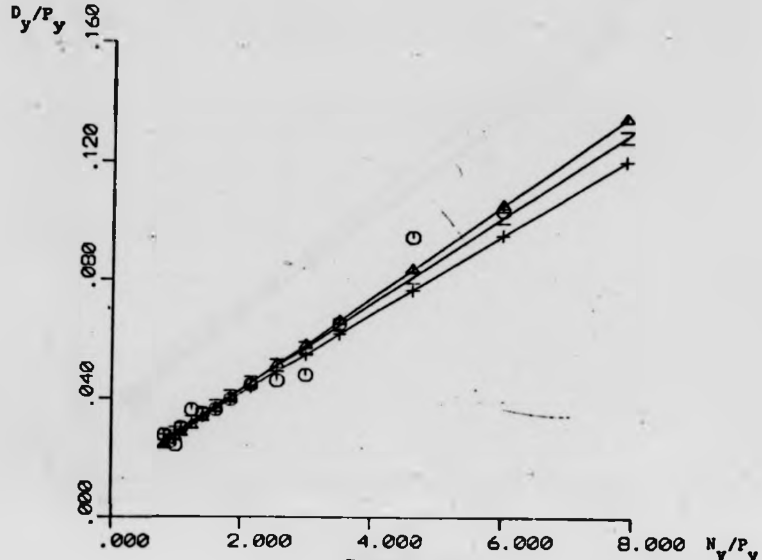
Formula	(A)	(B)	(A)	(B)	(A)	(B)
Method of fit						
Least square	.016	.017	.016	.015	.016	.015
Weighted least square	.006	.008	.013	.013	.018	.015
Wald	.013	.014	.015	.015	.017	.016

Graph (C.4) Mortality decline, out migration, differential under-registration and age error

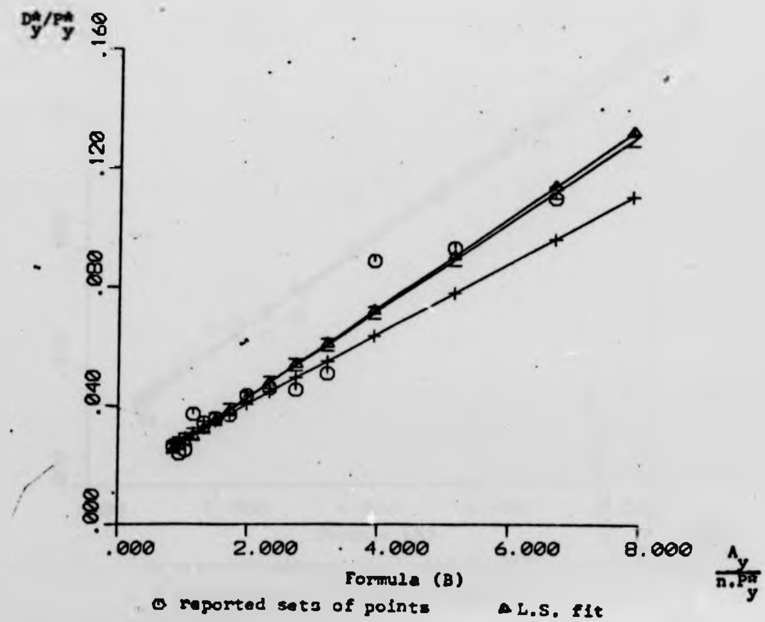




Graph (C.5) Mortality decline, no migration, differential under-registration and age error



Formula (A)

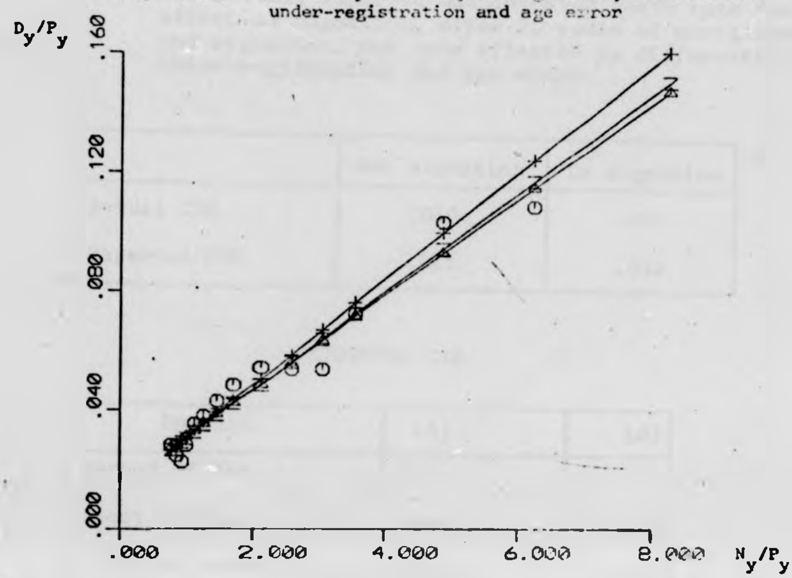


Formula (B)

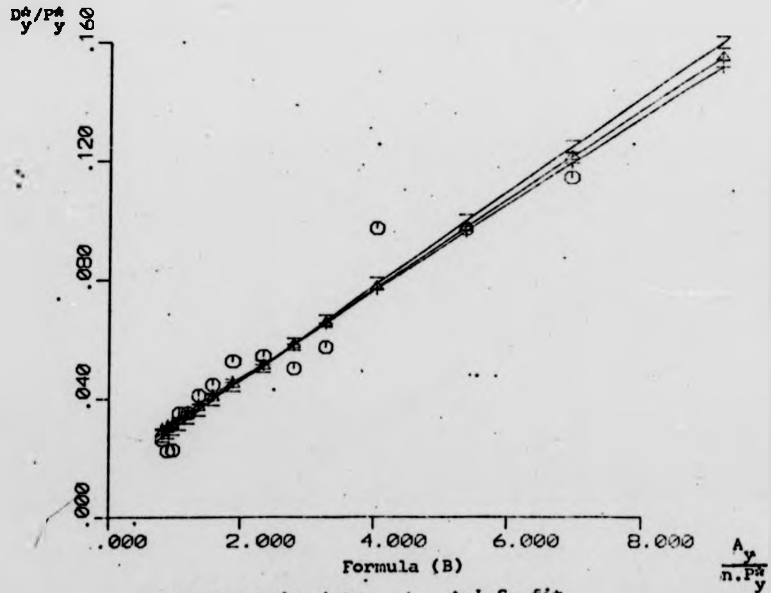
⊙ reported sets of points   ▲ L.S. fit

\+ W.L.S. fit   z Wald fit

Graph (C.6) Mortality decline, in migration, differential under-registration and age error



Formula (A)



Formula (B)

ⵁ reported points ▲ L.S. fit
+ W.L.S. fit ⵉ Wald fit

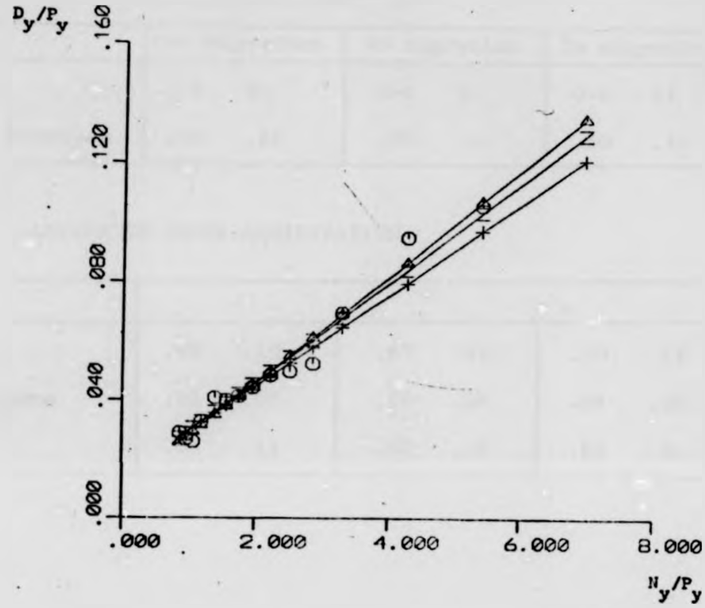
Table (C.3) The actual, reported and adjusted death rate for the effect of migration, after 20 years of mortality decline and migration, for data affected by differential under-registration and age error.

	Out migration	In migration
Actual CDR	.019	.017
Reported CDR	.015	.012

ADJUSTED CDR

Formula	(A)	(B)
Method of fit		
Least square	.018	.016
Weighted least square	.015	.013
Wald	.017	.015

Graph (C.7) Adjustment for out migration



Graph (C.8) Adjustment for in migration

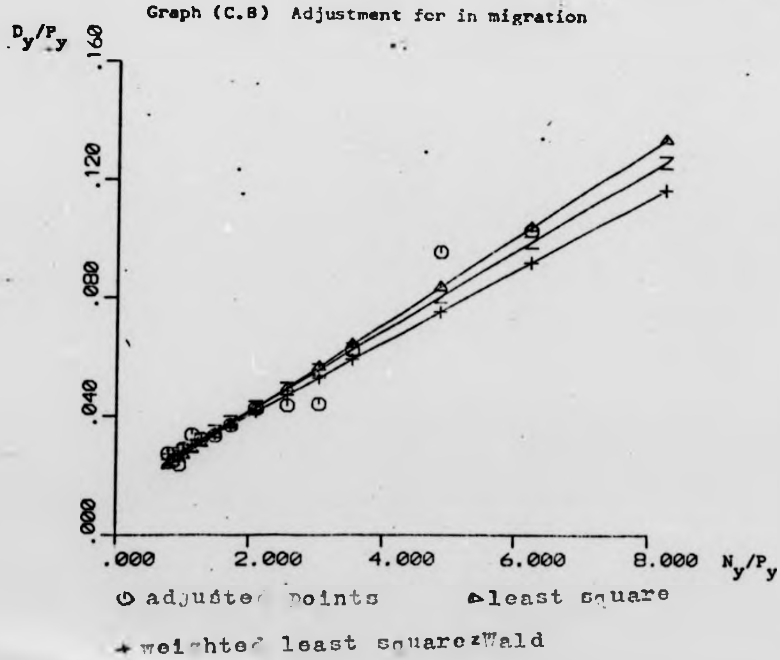


Table (C.4) The estimated proportionate under-registration for different ages

	Out migration		No migration		In migration	
Ages	0-5	5+	0-5	5+	0-5	5+
Actual under-registration	.50	.10	.50	.10	.50	.10
ESTIMATED UNDER-REGISTRATION						
Method of fit						
Least square	.52	.19	.67	.16	.70	.18
Weighted least square	.33	.05	.59	.04	.64	.02
Wald	.38	.13	.61	.12	.65	.11

APPENDIX (D)

Table (D.1) Number of males and females population

Age	Males	Females	Total
0-	63131	66180	129311
1-	29683	29082	58765
2-	37321	37334	74655
3-	52869	53567	106436
4-	49839	49649	99488
5-9	210647	195734	406381
10-	115037	91773	206810
15-	101113	127909	229022
20-	74841	123892	198733
25-	85766	133706	219472
30-	63973	90858	154831
35-	77743	98272	176015
40-	59008	63027	122035
45-	58219	59762	117981
50-	38529	34840	73369
55-	37596	30888	68484
60-	21997	19466	41463
65-	20213	17951	38164
70-	10741	10629	21370
75-	7782	6176	13958
80+	6764	5049	11813
N.D.	486	1167	1653
TOTAL	1223298	1346911	2570219

Table (D.2)* Number of males and females deaths

Age	Males	Females	Total
0-	18526	15823	34349
1-	2619	1347	3966
2-	3049	2651	5700
3-	2931	2807	5738
4-	1521	1712	3233
5-9	3292	2542	5834
10-	1480	1979	3459
15-	2601	2756	5357
20-	1812	2161	3973
25-	1849	3088	4937
30-	1516	2243	3759
35-	1842	2327	4169
40-	1639	1589	3228
45-	1761	1746	3507
50-	1455	1113	2568
55-	1582	1355	2937
60-	1484	1268	2752
65-	1496	1332	2828
70-	1080	1236	2316
75-	1050	640	1690
80+	1227	734	1961
N.D.	5	10	15
TOTAL	55808	52449	108158

* Source: "Etudes demographiques par sondage en Guinee 1954-55, Resultats definitifs - 1" reproduced from table 3.23.1.

Table (D.3)** Number of males and females live births in the last twelve months for different regions

Guinee Maritime		Fouta Djallon		Hante Guinee		Guinee Forestiere	
M	F	M	F	M	F	M	F
14530	15090	30572	30467	10065	9735	24162	24037

** Source: "Etude demographiques pare sondage en Guinee 1954-55, Resultats definitifs - 1" reproduced from table 3.21.1.

REFERENCES

ABOU GAMRAH, H. (1976) Methods of adjusting the stable estimates of fertility for the effects of mortality decline. Pop. Studies., 30, 1, 165.

ARRIAGA, E.E. (1968) New life tables for Latin American populations in the nineteenth and twentieth centuries. Institute of International Studies, University of California, Berkeley.

ARRIAGA, E.E. and DAVIS, K. (1969) The pattern of mortality change in Latin America. Demography, 6, 3, 223.

BASAVARAJAPPA, K.G. (1968) The significance of differences in patterns of mortality for population projections. Demography, 5, 1, 185.

BEHRMAN, S.J. and CORSA, L. and FREEDMAN, R. (1969) Editors. Fertility and family planning, a world view. Ann Arbor.

BLACKER, J.G.C. (1977) The estimation of adult mortality in Africa from data on orphanhood. Population Studies, 31, 1, 107.

BOURGEOIS-PICHAT, J. (1957) Utilisation de la notion stable pour mesurer la mortalité et la fecondité des populations des pays sous développés. Bulletin de l'institut International de Statistique, Congrès de Stockholm, 36, 2, 94.

BRASS, W. (1964) Use of census or survey data for the estimation of vital rates. African seminar on vital statistics, Addis Ababa: U.N. Economic Commission for Africa.

BRASS, W. (1971) On the scale of mortality.in biological aspects of demography, p.69, Taylor and Francis Limited, London.

BRASS, W. (1972) Mortality estimation by indirect means. Working paper for UNESOB expert group meeting on mortality, Beirut. Unpublished. (Abridged version published in Population Bulletin of the United Nations Economic and Social Office in Beirut, no. 4 (January 1973), p.9.)

BRASS, W. (1974) The estimation of fertility and mortality from defective vital registration records. Working paper for UNESOB meeting. Unpublished.

BRASS, W. (1975) Methods for estimating fertility and mortality from limited and defective data. Laboratories for Population Statistics, an occasional publication, The University of North Carolina at Chapel Hill.

BRASS, W. (1976) Indirect methods of estimating mortality illustrated by application to Middle East and North African data. Population Bulletin of the United Nations Economic Commission for Western Asia, nos. 10 and 11, Jordan.

BRASS, W. and COALE, A.J. and DEMENY, P. and HEISEL, D.F. and LORIMER, F. and ROMANIUK, A. and VAN DE WALLE, E. (1968) The demography of tropical Africa. Princeton University Press, Princeton, New Jersey.

BRASS, W. and HILL, K. (1973) Estimating adult mortality from orphanhood. International Population Conference, Liège, 1973, 3, 111.

CARRIER, N.H. (1958) A note on the estimation of mortality and other population characteristics given deaths by age. Pop. Studies, 12, 2, 149.

CARRIER, N.H. and FARRAG, A.M. (1959) The reduction of errors in census populations for statistically underdeveloped countries. Pop. Studies, 12, 3, 240.

CARRIER, N.H. and HOBCRAFT, J. (1971) Demographic estimation for developing societies. Population Investigation Committee, London.

CHANDRA SEKAR, C. and DEMING, W.E. (1949) On a method of estimating birth and death rates and the extent of registration. Journal of the American Statistical Association, 44, 101.

COALE, A.J. (1956) The effect of changes in mortality and fertility on age composition. Millbank Memorial Fund Quarterly, 34, 79.

COALE, A.J. (1963) Estimates of various demographic measures through the quasi stable age distributions. In Emerging Techniques in Population Research, Millbank Memorial Fund, 175.

COALE, A.J. (1972) The growth and structure of human populations, a mathematical investigation. Princeton University Press, Princeton, New Jersey.

COALE, A.J. and DEMENY, P. (1966) Regional model life tables and stable populations. Princeton University Press, Princeton, New Jersey.

DAVIS, K. (1956(a)) Amazing decline of mortality in underdeveloped areas. The American Economic Review, 46, 305.

DAVIS, K. (1956(b)) The unpredicted pattern of population change. The Annals American Academy of Political and Social Science, 305, 3, 53.

DEMENY, P. (1965) Estimation of vital rates for populations in the process of destabilization. Demography, 2, 203.

EL BADRY, M.A. (1965) Trends in the components of population growth in the Arab countries of the Middle East: a survey of present information. Demography, 2, 140.

FARGUES, P. and COURBAGE, Y. (1972) Some methodological elements proper to Chinese data (1970) in order to obtain basic indices on mortality. Working paper for UNESOB expert group meeting on mortality. Unpublished.

FREEDMAN, R. and TAKESHITA, J.Y. and SUN, T.H. (1962) Fertility trends in Taiwan, tradition and change. Pop. Studies, 16, 3, 219.

FREEDMAN, R. and TAKESHITA, J.Y. (1964) Fertility and family planning in Taiwan: a case study of the demographic transition. The American Journal of Sociology, 70, 16.

GILLE, H. (1949) The demographic history of the Northern European countries in the 18th century. Pop. Studies, 2, 2, 129.

GOODMAN, L.A. (1967) On the age sex composition of the population that would result from given fertility and mortality conditions. Demography, 4, 1, 423.

HAREWOOD, J. (1968) Recent population trends and family planning activity in the Caribbean. Demography, 5, 2, 874.

HILL, K.H. (1975) Indirect methods of estimating adult mortality. Ph.D. Thesis from University of London, London School of Hygiene and Tropical Medicine.

KEYFITZ, N. and FLIEGER, W. (1968) World population, an analysis of vital data. The University of Chicago Press, Chicago.

LEDERMANN, S. (1969) Nouvelles Tables - Types de Mortalité. (Cahier no. 53 de l'Institut National d'Etudes Démographiques.) Presses Universitaires de France.

MIRO, C.A. (1964) The population of Latin America. Demography, 1, 1, 15.

NAGI, M.H. and STOCKWELL, E.G. and SNAVELY, L.M. (1973) Digit preference and avoidance in the age statistics of some recent African censuses: some patterns and correlates. International Statistical Review, 41, 165.

NAMBOODIRI, N.K. (1969) On the dependence of age structure on a sequence of mortality and fertility. Demography, 6, 3, 287.

POTTER, J.E. and MYRIAM ORDÓÑEZ, G. (1976) The completeness of enumeration in the 1973 census of the population of colombia. pop. Index, 42, 3, 377.

PRESTON, S.H. (1974) Effect of mortality change on stable population parameters. Demography, 11, 1, 119.

PRESTON, S.H. and KEYFITZ, N. and SCHOEN, R. (1972) Causes of death, life tables for national populations. Seminar Press, New York and London.

SHRYOCK, H.S. and SIEGEL, J.S. (1973) The methods and materials of demography. U.S. Bureau of the Census, Washington, D.C.

STOLNITZ, G.J. (1955) A century of international mortality trends. Part I. Pop. Studies, 9, 3, 24.

STOLNITZ, G.J. (1956) A century of international mortality trends. Part II. Pop. Studies, 10, 1, 17.

STOLNITZ, G.J. (1956(b)). Mortality decline and age distribution. Millbank Memorial Fund, XXXIV, 2, 178.

STOLNITZ, G.J. (1965) Recent mortality trends in Latin America, Asia and Africa, review and interpretation. Pop. Studies, 19, 2, 117.

TORKI, F.C. (1975) Trends and differentials in mortality in Kuwait: 1965-70. Cairo Demographic Centre, Seminar on mortality trends and differentials in some Arab and African countries.

UNITED NATIONS (1952) Population Bulletin of the United Nations, no. 2, New York.

UNITED NATIONS (1955) Age and sex patterns of mortality: model life tables for underdeveloped countries (Population Studies no. 22) New York.

UNITED NATIONS (1956) Methods for population projections by sex and age. Manual III, ST/SOA/ Series A, Population Studies, no. 25, New York.

UNITED NATIONS (1962) Population Bulletin of the United Nations, no. 6 (with special reference to the situation and recent trends of mortality in the world) New York.

UNITED NATIONS (1963) Population Bulletin of the United Nations, no. 7 (conditions and trends of fertility in the world) New York.

UNITED NATIONS (1966) Demographic year book.

UNITED NATIONS (1967) Methods for estimating basic demographic measures from incomplete data. Pop. Studies, no. 42, Manual 4, New York.

UNITED NATIONS (1972) Mortality models and demographic projections. Ad Hoc Committee of Experts on Methods of Revising United Nations Model Life Tables.

UNITED NATIONS (1975) Techniques of evaluation of basic demographic data. African population studies series, no. 2. Addis Ababa.

VAN DE WALLE, E. (1966) Some characteristic features of census age distributions in illiterate populations. American Journals of Sociology.

VAZQUEZ, J.L. (1968) Fertility decline in Puerto Rico: extent and causes. Demography, 5, 2, 855.

YOU POH SENG (1959) Errors in age reporting in statistically under-developed countries. Pop. Studies, 13, 2, 164.

ZACHARIAH, K.C. (1971) Problems of applying stable population technique in estimating demographic measures for Arab countries. Paper presented at the annual meeting of the American Statistical Association. Fort Collins.

ÉTUDE DÉMOGRAPHIQUE PAR SONDAGE EN GUINÉE 1954-55, Résultats Définitifs, 1, Administration Générale des services de la France D'outre-Mer.