Population Division
Department of Economic and Social Affairs
United Nations Secretariat

METHODS FOR ESTIMATING ADULT MORTALITY

ESA/P/WP.175
1 July 2002

English Only

Population Division
Department of Economic and Social Affairs
United Nations Secretariat

METHODS FOR ESTIMATING ADULT MORTALITY

NOTE

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

The designations "developed" and "developing" countries and "more developed" and "less developed" regions are intended for statistical convenience and do not necessarily express a judgement about the state reached by a particular country or area in the development process.

The term "country" as used in the text of this publication also refers, as appropriate, to territories or areas.

The present report has been reproduced without formal editing.

PREFACE

Progress in the measurement of adult mortality in developing countries has lagged far behind achievements in measuring infant and child mortality. Much of the difficulty in measuring adult mortality arises from the rarity of adult deaths, relative to the size of the population at risk. However, of even more significance in developing countries is the absence of a reliable civil registration system that records deaths and the demographic characteristics of the deceased. In such contexts, indirect methods of estimation such as those discussed in this volume are invaluable tools.

The preparation of this volume was driven by the Population Division's continued interest in fostering the development and sharing of skills for demographic analysis. The last major effort by the Division to collate and disseminate state of the art information on indirect methods of demographic estimation of adult mortality was *Manual X: Indirect Techniques for Demographic Estimation*, *1983*.[1] That report, which was prepared in collaboration with the Committee on Population and Demography of the National Research Council, United States National Academy of Sciences, covered a broad range of indirect methods. This volume is more restricted in coverage, focusing only on methods for estimating adult mortality or related parameters.

The need for a greater focus on adult mortality measurement in developing countries has been accentuated by recent evidence of increasing adult mortality in a number of countries as a result of the human immune deficiency virus (HIV) and acquired immune deficiency syndrome (AIDS) epidemic. Although this report does not deal with cause-specific mortality measurement, familiarity with and application of the techniques discussed herein should permit more ready use of census and survey data to assess levels and trends in overall mortality.

Acknowledgement is due to Mr. Griffith Feeney, who assisted the Population Division in the preparation of this report.

To discuss the present publication, or population issues in general, please contact the office of Mr. Joseph Chamie, Director, Population Division, United Nations, New York, NY 10017, USA, at telephone (212) 963-3179 or fax (212) 963-2147.

NOTE

[1] United Nations publication, Sales No. E.83.XIII.2.

CONTENTS

Page

CONTENTS (*continued*)

CONTENTS (*continued*)

TABLES

TABLES (*continued*)

No. *Page*

FIGURES

FIGURES (*continued*)

ANNEX TABLES

ANNEX FIGURES

Explanatory notes

Symbols of United Nations documents are composed of capital letters combined with figures. Various symbols have been used in the tables throughout this report.

Two dots (..) indicate that data are not available or are not separately reported.
An em dash (—) indicates that the population is less than 500 persons.
A hyphen (-) indicates that the item is not applicable.
A minus sign (-) before a figure indicates a decrease.
A full stop (.) is used to indicate decimals.
Use of a hyphen (-) between years, for example, 1995-2000, signifies the full period involved, from 1 July of the beginning year to 1 July of the end of the second year.

Numbers and percentages in tables do not necessarily add to totals because of rounding.

The following abbreviations have been used in the present document:

AIDS	Acquired immunodeficiency syndrome
HIV	Human immunodeficiency syndrome

INTRODUCTION

The level of mortality in a society is a fundamental indicator of health and development. The ageing of populations in both developed and developing countries, with the associated increasing share of mortality that occurs in adulthood, has accentuated the need to obtain better estimates of mortality at adult ages. In developed countries, adult mortality can be measured using data from civil registration systems and population estimates derived from censuses or population registers. In most developing countries, however, the estimation of adult mortality is seriously constrained by the absence of reliable, continuous, and complete data registration systems.

This manual brings together existing methods for adult mortality estimation in situations where reliable and complete data registration systems are not available. The manual explains the concepts behind each method, details the steps required for application, and discusses issues of analysis and interpretation.

The methods discussed in this volume are indirect methods, and they do not provide the same degree of accuracy as direct methods, which use complete registration statistics. However, each of the methods presented involves a standard series of calculations that will, in the best of circumstances, produce useful estimates of adult mortality. Unlike methods based on reliable civil registration data, however, the accuracy of the estimates produced by the methods discussed herein cannot be taken for granted, but must be established in each application. This validation requires knowledge and judgement that go well beyond the mechanical application of the equations that underpin each method and require a good understanding of the assumptions on which each method is based. A key strategy, in this regard, is to derive estimates from all data available for each particular case, to compare them, and to use the comparisons to make judgements on the accuracy of the different data sources and the validity of the assumptions underlying the various methods.

A. Overview of contents

Following a brief overview of mortality measurement in sections C and D of this Introduction, the manual is organised according to the data required for the application of the methods described. Chapter I discusses census survival methods which require, as input, age distributions derived from at least two consecutive censuses. Because nearly every country in the world has taken at least two population censuses, these methods are very widely applicable. Census survival methods yield fairly accurate results when the census data used are accurate in terms of both coverage and age reporting. However, the results are sensitive to certain kinds of data errors, and they are not applicable to populations that experience substantial migration.

In many countries data on the age distribution of the population from two or more consecutive censuses can be supplemented by data on the number of intercensal deaths by age and sex. These data may be derived from a civil registration system, even when the latter does not achieve complete coverage of events, or they may be obtained from field inquiries (censuses or surveys) using questions on the number and demographic characteristics of deaths occurring in each household over a given period. By combining age distributions obtained from censuses with data on intercensal deaths, it is possible to estimate the degree of under-reporting of deaths and, consequently, the number of deaths that were not reported. The reported number of deaths may then be adjusted and used to estimate a life-table. Estimates derived in this way are the subject of chapters II and III. The applicability of the methods described in those chapters, just as that of methods based solely on the estimation of intercensal survival, is limited to populations in which migration is negligible.

Chapters IV and V discuss the application of methods based on responses to retrospective questions on the survival status of specified relatives. Unlike the methods presented in Chapters II and III, the methods using information on the survival of a particular relative often do not require that the population be closed to migration. In Chapter IV, the focus is on methods based on responses to questions on parental survival and in Chapter V, methods that estimate adult mortality from information on the number of surviving siblings are discussed.

If the data used were always free from error, and if the assumptions on which the methods are based always held in practice, estimates derived using different approaches would coincide. Data are, however, frequently subject to different types of error and the assumptions on which the various methods are based are rarely perfectly met. As a result, the application of different methods to available data typically results in a range of estimates. To arrive at useful assessments of adult mortality it is necessary to interpret these estimates in light of other pertinent information, including typical errors in the data used, the behaviour of particular methods in other applications, and the demographic situation of the population.

Three annexes to this manual review tools and practical issues in mortality estimation. Annex 1 discusses practical considerations in data handling and processing. The annex is intended for those who need guidance on how to assess data quality and how to avoid common computational errors. Annex II provides an overview of the use of model life tables. The annex does not focus on the construction of model life tables but rather on the utility of these tables in adult mortality estimation. Annex III deals with line-fitting.

B. Scope and limitations

This manual is intended for users who have some basic knowledge of demography and demographic estimation. It assumes a fairly good grasp of the life table and the interrelations among its functions. However, the next two sections review basic mortality measurement and three annexes are provided as a reference for the user who needs to review these materials. Readers who require more detailed revision of basic demographic concepts may need to consult a standard demography text. Emphasis is placed in the presentation on how specific methods are applied and detailed applications are provided using data from Japan and Zimbabwe. Annotated tables demonstrate the detailed steps involved in each application.

C. Overview of basic mortality measurement

Two broad types of demographic statistics are used to measure mortality. The most common is the crude death rate, which is calculated by dividing the number of deaths that occur in a population during a given year or period by the average number of person-years lived by the population during that period.

The crude death rate is "crude" because it does not take account of the age distribution of the population. Age is fundamental to the study of mortality because the risk of death is very different at different ages. It is therefore important to control for age differences between populations, or for changing age distribution in a population over time, by computing "age-specific" death rates. These are defined in the same way as crude death rates, as number of deaths divided by the average number of person-years lived by the population over a particular period, except that deaths and population are restricted to a particular age group.

D. Life-table statistics

Age-specific death rates for males and females provide the essential information needed to study mortality risks. For many reasons, however, it is useful to transform them into life-table statistics, such as the expectation of life at different ages or the probability of survival over a particular age interval. A life-table is a more or less standard collection of statistics describing the age pattern of mortality in a population. Life-table statistics are the second broad type of statistics used to measure mortality.

Life-tables are of two types. Cohort or generation life-tables record the mortality experience of the group of persons born during a given year or other period. Period life-tables are synthetic constructs that show what the mortality experience of a hypothetical group of persons would be if they experienced the death rates observed in a population during a given year or other period.

Cohort life-tables have the advantage of conceptual simplicity, but the disadvantage of requiring data for, and referring to mortality risks over a very long time span. Since the upper limit of human life is about 100 years, a cohort life-table can be constructed only for groups of persons born at least one hundred years ago. Even when such life-tables can be constructed--and this is not possible for many countries of the world, including many developed countries—they represent an amalgam of the mortality experience over a very long period.

Period life-tables are conceptually more complex, but have the advantage of providing mortality measures localised in time. This makes it possible, for example, to talk about the change in expectation of life at birth from one year to the next. Most life-tables available for human populations are, in fact, period life-tables.

It is also possible to distinguish between period and cohort statistics in a more general way because life-table measures can be constructed on the basis of cohort experience over just a portion of the human life span. This manual, in particular, deals only with life-table measures for ages above age 5. Then, from an expanded perspective, period mortality statistics are those calculated on the basis of deaths observed during a given period and cohort statistics are those calculated on the basis of all deaths occurring to a particular group of persons followed over time.

I. CENSUS SURVIVAL METHOD

Census survival methods are the oldest and most widely applicable methods of estimating adult mortality. These methods assume that mortality levels can be estimated from the survival ratios for each age cohort over an intercensal period. Under optimal conditions, census survival methods provide excellent results. They are, however, applicable only to populations that experience negligible migration. They are also sensitive to age distribution errors and, in some cases, they give extremely poor results. Age reporting errors, in particular, can result in large variations in calculated survival ratios and inconsistent estimates of mortality. Census survival methods can also be seriously biased by relative differences in the completeness of censuses. It is therefore important to assess the input data carefully and to evaluate the results in whatever ways existing data sources allow.

A. DATA REQUIRED AND ASSUMPTIONS

Census survival methods require two age distributions for a population at two points in time. While variations for use with other age groups are possible, five-year age groups are nearly always the norm. It is desirable for the five-year age groups to extend into very old ages, with an open-ended group of 85+ or higher, although older age groups may be collapsed to reduce the effects of age exaggeration.

It is necessary to know the reference dates of the censuses producing the age distributions used. Reference dates often change from one census to the next and obtaining the correct length of the intercensal interval is critical.

Since census survival methods should be used only for populations in which migration is negligible, they can only be applied to national populations or to subpopulations whose characteristics do not change over time. In particular, census survival methods are generally not suitable for generating estimates of mortality for rural and urban areas, or for geographically defined subpopulations.

B. CENSUSES FIVE YEARS APART

This section considers the derivation of adult mortality estimates in a simple case of two censuses taken exactly five years apart. The objective is to derive the expectation of life at specific ages through adulthood.

Assume that people aged 0-4 at the first census are concentrated at the mid-point of the age group, i.e., that they are all aged 2.5 years exactly. They will then be 7.5 years exactly at the second census. Dividing the number of persons aged 5-9 at the second census by the number aged 0-4 at the first census therefore gives an estimate of the life-table conditional survival probability from age 2.5 years to 7.5 years, which is denoted by $l_{7.5}/l_{2.5}$. Similar quotients for subsequent age groups estimate the conditional survival probabilities $l_{12.5}/l_{7.5}$, $l_{17.5}/l_{12.5}$, and so on. In general, the life-table probability of surviving from the mid-point of one age group to the next is approximated by the census-survival ratio. That is,

$$l_{x+2.5}/l_{x-2.5} = P_2(x,5)/P_1(x-5,5)$$

for $x = 5, 10, 15 \ldots$, where $P_1(x-5,5)$ is the population aged $x-5$ to x in the first census and $P_2(x,5)$ is the population aged x to $x+5$ in the second census.

Cumulative multiplication of these probabilities gives the conditional survival schedule $l_x/l_{2.5}$. Thus, $l_{2.5}/l_{2.5} = 1$ and

$$l_{x+5}/l_{2.5} = (l_{x+5}/l_x)(l_x/l_{2.5}) \quad (1)$$

for $x = 2.5, 7.5, \ldots$. Interpolation is required to convert the non-standard ages, 2.5, 7.5 … to ages $x = 5, 10, \ldots$. Linear interpolation, for the conditional l_x values using the formula:

$$l_x/l_{2.5} = 0.5(l_{x-2.5}/l_{2.5} + l_{x+2.5}/l_{2.5}) \quad (2)$$

for $x = 5, 10, \ldots$ will usually suffice. However more elaborate interpolation methods can be applied, if warranted.

From the conditional l_x values given by formula (2) the conditional estimates of the number of person years lived in each age group (${}_5L_x$) can be calculated using

$$ {}_5L_x/l_{2.5} = 2.5(l_x/l_{2.5} + l_{x+5}/l_{2.5}), \quad (3)$$

and then, given a value of $T_x/l_{2.5}$ for some initial old age x, conditional T_x values can be calculated as:

$$T_{x-5}/l_{2.5} = T_x/l_{2.5} + {}_5L_{x-5}/l_{2.5}\,, \qquad (4)$$

The final result, the expectation of life at age x, is then computed as

$$e_x = (T_x/l_{2.5})/(\,l_x/l_{2.5}), \qquad (5)$$

where the $l_{2.5}$ values cancel out on division.

Census survival estimation, in this case, is a direct application of basic life-table concepts but for one detail: obtaining an initial value of the person years lived above age x (T_x) for some old age x. If the age distributions provide sufficient detail and age-reporting is accurate, T_x may simply be taken to be equal to zero for some very old age; $x=100$, for example. In contexts where there is severe age exaggeration at very old ages, however, this approach can result in major distortion of the mortality estimates. Special procedures for dealing with this problem, with an application to data for Zimbabwe, are discussed in section H.

C. Five-year intercensal interval method application: Japan, females, 1965-1970

To illustrate the application of the five-year intercensal survival method, the procedures discussed in the previous section have been applied to data on females enumerated in the 1965 and 1970 censuses of Japan. Japan has conducted a series of censuses at five-year intervals from 1920 through 1995, interrupted only during the 1940s. All censuses have a reference date of October 1 so adjustment of the intercensal period is not necessary.

Table I.1 shows the results of the application. The calculations are based on the equations presented in section B. Further details of the procedure are provided in the notes to the table.

The estimated expectations of life (e_x) for $x=5,10,...75$ are given in column 11. These estimates from the application of the intercensal survival method are compared with values for e_x from life-tables derived from registered deaths (column 12). Since the quality of age-reporting in Japan is very high, the results of applying the five-year intercensal survival method are comparable to life-table estimates obtained from deaths registered through a civil registration system. The median deviation of the results from estimates derived from the civil registration data is 0.4 per cent (column 13). More precise estimates are unlikely in other applications of the census survival method and, even for Japan, results for males or for other intercensal periods are less accurate.

D. Censuses *t* years apart

The calculations of the preceding section may be adapted, with modest effort, for use with censuses 10 years apart. However, they do not readily extend to other intercensal intervals. Preston and Bennett (1983) have developed a different approach that works with any intercensal interval, although very short or very long intervals are likely to give poor results. This section presents a formulation that is similar to the Preston-Bennett method, but is simpler.

To apply this method - the synthetic survival ratio method it is necessary to first calculate the intercensal rate of growth of each age group from the age distributions produced by two consecutive censuses as follows:

$$r(x,5) = ln[P_2(x,5)/P_1(x,5)]/t, \qquad (6)$$

where $r(x,5)$ denotes the growth rate for the x to $x+5$ age group, $P_1(x,5)$ and $P_2(x,5)$ denote, respectively, the numbers of persons aged x to $x+5$ at the first and second censuses, and t denotes the length of the intercensal interval. Next, calculate the average annual number of person-years lived by persons in the x to $x+5$ age group, $N(x,5)$, during the intercensal period using

$$N(x,5) = [P_2(x,5) - P_1(x,5)]/[tr(x,5)] \qquad (7)$$

This number is an approximation of the number of persons aged x to x+5 at the midpoint of the intercensal period.

The synthetic survival ratios

$$\frac{N(x+5,5)exp\{2.5r(x+5,5)\}}{N(x,5)exp\{-2.5r(x,5)\}} \qquad (8)$$

can be calculated where the numerator here may be thought of as an interpolated number of persons aged

$x+5$ to $x+10$ at time $m+2.5$ years. The value of m denotes the mid-point of the intercensal period. This number is obtained by projecting the mid-period number of persons in this age group forward by 2.5 years using the age-specific growth rate $r(x+5,5)$. Similarly, the denominator in (8) may be thought of as an interpolated number of persons aged x to $x+5$ at time $m-2.5$ years. The persons represented in the numerator are thus, on the assumption that no migration occurs, the survivors of the persons represented in the denominator.

The synthetic survival ratios in (8) thus estimate the life table probabilities of survival from age x to $x+5$ (l_{x+5}/l_x) exactly as in the case of censuses five years apart. The remainder of the calculation is the same as in the case of censuses five years apart presented in section B.

If the intercensal interval is five years, the denominator of equation (8) equals the number of persons in the x to $x+5$ age group at the first census and the numerator is the number in the $x+5$ to $x+10$ age group at the second census. When censuses are five years apart, then, the method for arbitrary intercensal intervals described in this section is identical to the method for censuses five years apart described in section B.

E. Arbitrary intercensal interval method application: Japan, females, 1960-1970

Table I.2 illustrates the application of the census survival method for arbitrary intercensal intervals to data on females enumerated in the 1960 and 1970 censuses of Japan. Detailed procedures for the application of this method are provided with the table.

Columns 2 and 3 of the table show the age distributions of females enumerated in the two censuses, and column 4 shows the age-specific intercensal growth rates calculated using formula (6). The average annual person-years lived by persons in each age group during the intercensal period (column 5) may be thought of as an interpolated mid-period age distribution. Column 7 shows the synthetic survival ratios calculated according to formula (8), and subsequent columns show the same calculations as columns 6 to 13 of table I.1. The calculation assumes the expectation of life at age 80 (e_{80}) to be 5.99 years. This figure is obtained by interpolating between data on expectation of life from official Japanese sources (Japan Statistical Association, 1987, pp. 270-271). The last two columns compare the estimated expectations of life at birth with values from life-tables derived from registered deaths.

Although the estimates of life expectancy produced by the arbitrary intercensal interval method are in reasonably good agreement with those derived from vital registration data, they are not as good as the estimates obtained from applying the five-year intercensal interval method (section C). This is because the generalisation that allows estimation when intercensal intervals have any arbitrary length comes at a cost. When age-specific growth rates change substantially from one five-year age group to another, as they do in this example, the growth rates of the number of persons at different ages within each age group will also be far from constant. Errors in the synthetic survival ratios will therefore occur because the interpolation that produces the numerators and denominators of those ratios assumes a constant rate of growth within each five-year age group during the intercensal period.

In this example, the sharply lower size of the cohort aged 10-14 in 1970 relative to the cohort the same age in 1960 (3.9 million and 5.4 million, respectively), results in a large negative growth rate (-3.4 per cent) for 10-14 year olds during the intercensal period. Growth rates for the 0-4, 5-9 and 15-19 age groups, in contrast, are considerably higher. This variability of growth rates results in a synthetic survival ratio from age 17.5 to age 22.5 that is much too high, with the result that errors in the estimated expectations of life at ages 5 and 10 are relatively large.

F. Censuses ten years apart

When censuses are exactly ten years apart, ten-year intercensal survival ratios can be calculated by dividing the number of persons aged 10-14 at the second census by the number aged 0-4 at the first census; the number aged 15-19 at the second census by the number aged 5-9 at the first census, and so on. Assuming, as in the case of censuses five years apart, that persons are concentrated at the mid-points of age groups, the intercensal survival ratios for age groups 0-4, 10-14, 20-24, etc., give estimates of the conditional probabilities of survival, $l_{12.5}/l_{2.5}$, $l_{22.5}/l_{12.5}$, ...

and the ratios for age groups 5-9, 15-19, ... give estimates of the conditional survival probabilities $l_{17.5}/l_{7.5}$, $l_{27.5}/l_{17.5}$, and so on.

This results in two series of conditional l_x values. The first consists of the conditional survival probabilities $l_x/l_{2.5}$, computed by noting that $l_{2.5}/l_{2.5} = 1$ and using the formula

$$l_{x+10}/l_{2.5} = (l_{x+10}/l_x)\,(l_x/l_{2.5}) \qquad (9)$$

for $x = 2.5, 12.5, 22.5, ...$ The second consists of the conditional survival probabilities $l_x/l_{7.5}$, computed by noting that $l_{7.5}/l_{7.5} = 1$ and using the formula

$$l_{x+10}/l_{7.5} = (l_{x+10}/l_x)(l_x/l_{7.5}) \qquad (10)$$

for $x = 7.5, 17.5, ...$

While it would be possible to carry out subsequent calculations independently on both of these series, this procedure would have the dual disadvantage of working with ten-year, rather than five-year age intervals, and of providing two different sets of estimates. It is preferable to merge the two series, thus giving survival values at five-year intervals. Averaging the first two terms of the first series gives a value of $l_{7.5}/l_{2.5}$,

$$l_{7.5}/l_{2.5} = 0.5(l_{2.5}/l_{2.5} + l_{12.5}/l_{2.5}) \qquad (11)$$

Multiplying the second series by $l_{7.5}/l_{2.5}$ results in a series with $l_{2.5}$ values in the denominator, that may be merged with the first series so that

$$l_x/l_{2.5} = (l_{7.5}/l_{2.5})(l_x/l_{7.5}), (12)$$

$x = 7.5, 17.5, ...$ Once the merged series is available, subsequent calculations are the same as for the two previous methods.

G. Ten-year intercensal interval method application: Japan, females, 1960-1970

Table I.3 illustrates the application of the method for ten year intercensal intervals to data for females enumerated in the 1960 and 1970 censuses of Japan. Details of the calculation are given in the notes to the table.

As with previous applications, the last two columns of table I.3 compare the estimated expectations of life to those derived from the deaths recorded by the civil registration system. The percentage deviations are similar to those displayed in table I.1, with a median error of 0.4 per cent. Note that the results of the arbitrary intercensal interval method in table I.2 show much wider deviations with a median error of 1.1 per cent. This outcome suggests that when censuses at exact ten-year intervals are available, the ten-year method should be used in preference to the arbitrary intercensal interval method presented in section E.

H. Ten-year intercensal interval method application: Zimbabwe, females, 1982-1992

The preceding examples show that estimates derived from census survival methods can be very accurate when the age distribution data used as input are reliable, as is the case with Japan. Much of the data to which the indirect estimates discussed in this manual will be applied, however, will come from contexts where reliable civil registration statistics are lacking and where census age distributions are less accurate.

Table I.4 therefore illustrates a more typical application using the example of census data for Zimbabwe. The ten-year intercensal method is applied to the data on females enumerated in the 1982 and 1992 censuses of Zimbabwe. Because the census reference dates are the same, the ten-year census survival method can be used. However, special procedures discussed in this section, have to be adopted to estimate life expectancy for the uppermost age group because, unlike Japan, good life table estimates are not available for Zimbabwe. Further, in the absence of accurate life table estimates with which to compare the results of this application, careful examination of the survival ratios becomes important in assessing the reliability of the life expectancy estimates. Approaches to this evaluation are also discussed in this section.

1. *Estimating the uppermost expectation of life*

Columns 1-13 of table I.4 show calculations for Zimbabwe that are identical to those in table I.3 for Japan. However, because the open-ended interval starts at age 75, a value for e_{70} is needed in order to calculate the T_x values in column 12. In the absence of reliable life table estimates, the simplest way to estimate the uppermost expectation of life, e_{70}, is to make an initial guess about the likely level of the expectation of life at birth and determine the corresponding value of e_{70} using model life tables. Even a rough guess of the life expectancy at birth will usually work reasonably well for two reasons. First, the range of variation in expectation of life at older ages is not large. The Brass model life-tables shown in annex table I.3, for example, suggest that increasing e_0 from 50 to 65 years increases e_{70} only from 8.25 to 9.76 years. Second, the estimated expectations of life at younger ages are relatively insensitive to the value of the expectation of life that is used to start the T_x calculation. This robustness may be illustrated with the example worked out for Zimbabwe in table I.4.

Assuming that the expectation of life at birth for Zimbabwe females during 1982-1992 is 60 years, a corresponding model life table value of e_{70} can be determined. The Brass model life-tables shown in annex table II.3 show that the corresponding e_{70} value, given a female e_0 of 60 years, is 9.09. The principle behind this is discussed in annex II. Using 9.09 provisionally as the uppermost expectation of life for purposes of our calculation in table I.4, would yield an expectation of life at age 5 of 63.3 years. However, as can be seen from the table of e_x values in annex table II.3, an e_5 of 63.3 years is closer to the model life table with e_0 of 62.5 years (column 18). The initial value of e_{70} should be replaced by the value from this table, which is 9.39 years. This gives an estimated e_5 of 63.7 years. The procedure for interpolating the e_x . . values is discussed in annex II.

2. Evaluating the census survival ratios

Since accurate life tables derived from vital registration statistics are not available for Zimbabwe, to assess the quality of the estimates derived using the census survival method, it is necessary to use a different approach from that used in the case of Japan.

The first step is to evaluate the levels and trends in the survival ratios. Figure I-1 plots the conditional survival ratios shown in column 5 of table I-4. It is important to look at these values, rather than the interpolated values in column 8, because the interpolation has a strong smoothing effect that obscures patterns resulting from age distribution errors.

The survival ratios plotted in figure I.1 show fluctuations from the third ratio through the end of the series, with more pronounced swings over age 50. Some of this variation is certainly due to imperfect merging of the two series of survival ratios (the one beginning with age group 0-4, the other with 5-9). However, the larger fluctuations for older ages cannot be accounted for in this way.

Another key observation is that the first three survival ratios are greater than one. This is impossible if the age data are accurate and the population was indeed closed to migration. Column 5 of table I.4 shows that this is due mainly to the first survival ratio, which is just over 1.1. The survival ratios of the 5-9 and 10-14 year age groups are also slightly above one. The high value of the first survival ratio might reflect substantial under enumeration of the 0-4 age group in 1982. This is generally believed to be a common problem in census enumeration, although it is difficult to know for certain whether the deficit is due to under enumeration or to age misreporting. In this case, however, a transfer of 0-4 year olds into the 5-9 age group would be expected to result in a second survival ratio less than one, contrary to what is observed here.

The first three survival ratios being greater than one might be interpreted to mean either that the 1992 census enumerated the population somewhat more completely than the 1982 census, or that there was net immigration into the affected age groups during the intercensal period. However, the 1992 Zimbabwe census was a less complete enumeration than the 1982 census, and for at least one category of international migrants, Europeans, net migration during the intercensal decade was negative, not positive. It is possible, therefore, that the survival ratios above one in table I-4 reflect differences in age misreporting or differential completeness of enumeration by age in the two censuses.

Another feature worth noting is the sharp fluctuation in the survival ratios for ages 50 and over that is exhibited by the Zimbabwe data. Such fluctuation, commonly observed in other populations, is most likely to result from age heaping. Despite the obvious distortions in survival ratios that they cause, age heaping errors cause relatively few problems for the estimation of overall mortality levels because the effect of higher values at some ages tends to be cancelled out by lower values at other ages.

Age exaggeration, in contrast to age heaping, may play an important role in biasing estimates derived from the use of census survival methods. One way of thinking about the effect of age exaggeration is to imagine what would happen to reported age distributions and survival ratios if everyone were to overstate their age by exactly five years. The survival ratio identified with, for instance, the 50-54 age group at the first census, would then refer, in fact, to the 45-49 age group. Since survival for the younger age group is higher, the survival ratio identified with the 50-54 age group would be too high. The same would be true for every other age group, and the result would be that the data, as reported, would overstate the estimated expectation of life.

Empirical patterns of age exaggeration are complex and not well understood. In some cases, they are pronounced enough to have important effects on estimates derived from census survival and other indirect methods. Systematic and substantial overstatement of age tends to begin only in the adult ages. The youngest age groups affected will lose persons by transference of some persons to older age groups. Older age groups will gain persons transferred from younger age groups and lose persons transferred to older age groups. If the population is young, as is the case in most developing countries, the number of persons will decline sharply from one age group to the next, at least for older age groups. If fixed proportions of persons in each age group overstate their ages, all age groups beyond the youngest one affected will tend to gain more persons than they lose. The effect on survival ratios is not immediately clear, since both the numerator and the denominator increase.

I. Translation to a common mortality indicator using model life-tables

One way to detect the presence of age exaggeration in an application of the census survival method is to transform the estimated expectations of life at each age to a common indicator, such as expectation of life at age 5, using a model life-table family. On the assumption that the age pattern of mortality in the population is represented by the model used, biases due to age exaggeration will be revealed by a tendency for estimates of the expectation of life at age 5 derived from data on older age groups to be higher. This method is discussed below, with an application to the Zimbabwe data.

Begin by taking the estimates of e_x for Zimbabwe females shown in table I.4, column 13, and compute the implied value of e_5 using the interpolation procedures described in section C of annex II. The result of the application of the method is shown in column 15 of table I.4 and in figure I.2. The values of e_5 range from a low of 59.1 years to a high of 68.4 years. The estimated e_5 values fall from ages 5 to 20, then rise from ages 20 to 45, followed by a levelling off, although downward spikes are evident for ages 50 and 70. This pattern suggests that although age exaggeration is undoubtedly present to some degree, it is not playing a major role in distorting the census survival ratios. If it were, there would be a clear increase in e_5 values above age 50. Further, an unsuitable choice of a model life-table would produce a set of e_5 values that increase or decrease smoothly with x. In contrast, a tendency for the e_5 values to rise as x increases, but only beyond the young adult ages, may indicate an upward bias in the survival ratios for older age groups due to age exaggeration.

The median of all the estimated e_5 values for Zimbabwe is 64.6. A useful indicator of the error associated with this estimate is one half the inter-quartile range of the distribution of the e_5 values, (2.8 years in this case). To indicate relative error it is useful to express this as a per cent of the estimated e_5, (4.3 per cent in this case).

J. Methodological note

The outcome of the use of synthetic survival ratios is equivalent to that of the Preston-Bennett method as originally formulated, but there is a difference that must be noted. Census survival ratios may be calculated with ratios of l_x values or with ratios of ${}_5L_x$ values. In the first case it is logical to assume that persons in each age group are concentrated at the mid-point of the group and thus, to begin the life table calculations at $x=2.5$ years with $l_{2.5}/l_{2.5}=1$. Conditional ${}_5L_x/l_{2.5}$ values are then calculated in the usual way, using equation (3).

The alternative, calculating survival ratios with ${}_5L_x$ values leads to the series

$$ {}_5L_5\,/{}_5L_0\,,\ {}_5L_{10}\,/{}_5L_0\,,\ {}_5L_{15}\,/{}_5L_0\,,\ \ldots, \qquad (13) $$

which, by analogy with the $l_x/l_{2.5}$ series, may be thought of as ${}_5L_x$ values "conditioned on" ${}_5L_0$. With this approach, l_x values are similarly conditioned, being calculated as

$$ l_x\,/{}_5L_0 = ({}_5L_x\,/{}_5L_0\ +\ {}_5L_{x\text{-}5}/{}_5L_0\,)/10 \qquad (14) $$

The ${}_5L_0$ term in the denominator cancels out when calculating e_x, just as the $l_{2.5}$ term in the denominator of equation (5), in section B, cancels out.

Section D uses the l_x ratio approach in preference to the ${}_5L_x$ ratio approach, and will accordingly yield slightly different results from the original Preston-Bennett formulation. It would, of course, be possible to use the ${}_5L_x$ ratio approach with the synthetic survival ratio method, but the l_x ratio approach has several advantages. The resulting statistics are directly interpretable as conditional survival probabilities, and there is a naturally available radix, the value one, with which to initiate the series. More importantly, the l_x ratio approach greatly simplifies census survival calculations for intercensal intervals that are ten years in length.

TABLE I.1. FIVE YEAR INTERCENSAL SURVIVAL METHOD APPLIED TO JAPAN: FEMALES, 1965-1970

	Census population				*Estimated conditional life table functions*							
Age group(i)	*1965* [a]	*1970* [b]	*Midpoint of age group*	*Census survival ratio*	*Probability of survival*	*Interpo-lated age*	*Probability of survival to age x*	*Person years lived between exact age x and x+5*	*Total person years lived above age x*	*Life expectancy at age x*	*Estimated life expectancy from civil registration data*	*Deviation (col11-col.12) per cent*
	P_1	P_2		$P_2(x,5)/P_1(x-5,5)$	$l_x/l_{2.5}$	x	$l_x/l_{2.5}$	${}_5L_x/l_{2.5}$	$T_x/l_{2.5}$	e_x	$e_{x(R)}$	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
0-4	3,983,902	4,292,503	2.5	1.0011	1.0000	NA	NA	NA	NA	NA	NA	NA
5-9	3,854,281	3,988,292	7.5	0.9994	1.0011	5	1.0006	5.0034	70.0627	70.02	70.19	-0.2
10-14	4,513,237	3,852,101	12.5	0.9953	1.0005	10	1.0008	4.9975	65.0593	65.01	65.33	-0.5
15-19	5,373,547	4,492,096	17.5	0.9951	0.9958	15	0.9982	4.9790	60.0617	60.17	60.41	-0.4
20-24	4,572,392	5,347,327	22.5	0.9999	0.9910	20	0.9934	4.9609	55.0827	55.45	55.54	-0.2
25-29	4,206,801	4,571,868	27.5	0.9961	0.9909	25	0.9909	4.9497	50.1218	50.58	50.74	-0.3
30-34	4,110,076	4,190,340	32.5	0.9940	0.9870	30	0.9889	4.9324	45.1722	45.68	45.96	-0.6
35-39	3,751,030	4,085,338	37.5	0.9795	0.9811	35	0.9840	4.8876	40.2397	40.89	41.21	-0.8
40-44	3,231,736	3,674,127	42.5	0.9899	0.9609	40	0.9710	4.8177	35.3522	36.41	36.52	-0.3
45-49	2,697,217	3,198,934	47.5	0.9819	0.9512	45	0.9561	4.7466	30.5345	31.94	31.89	0.1
50-54	2,485,095	2,648,360	52.5	0.9588	0.9340	50	0.9426	4.6432	25.7879	27.36	27.39	-0.1
55-59	2,071,540	2,382,691	57.5	0.9512	0.8955	55	0.9147	4.4709	21.1446	23.12	23.05	0.3
60-64	1,719,370	1,970,485	62.5	0.9217	0.8518	60	0.8736	4.2302	16.6738	19.09	18.89	1.0
65-69	1,343,444	1,584,699	67.5	0.8725	0.7851	65	0.8184	3.8836	12.4436	15.20	14.99	1.4
70-74	955,567	1,172,155	72.5	0.7705	0.6850	70	0.7350	3.3535	8.5599	11.65	11.45	1.7
75-79	644,043	736,258	77.5	0.6338	0.5278	75	0.6064	2.5938	5.2064	8.59	8.43	1.9
80-84	341,170	408,191	82.5	NA	0.3345	80	0.4311	NA	2.6127	6.06	6.06	NA
85+	176,068	206,511	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
									Median absolute per cent deviation			0.4

Source: Population distribution for 1965 and 1970 from: *Historical Statistics of Japan*, volume 1, table 2-9, pp. 66-83 (Japan Statistical Association, Tokyo, 1987).

[a] Reference date: 1 October 1965 (1965.751).

[b] Reference date: 1 October 1970 (1970.751).

Procedure

Columns 1-3. Record the age distributions of the two censuses as shown in columns 1 to 3, taking care to calculate exact reference dates of censuses.

Columns 4-5. Record mid-points of age groups and compute census survival ratios. Record these in columns 4 and 5 respectively. Note that the first census survival ratio is the number of persons aged 5-9 at the second census divided by the number aged 0-4 at the first census, and similarly for higher age groups. Note also, that the last ratio calculated takes the number of persons in the last five-year age group, 80-84 in this case, as its numerator. The numbers of persons in the open-ended age groups are not used here.

Column 6. Compute the conditional survival schedule $l_x/l_{2.5}$, noting that $l_{2.5}/l_{2.5} = 1$ and using the equation

$$l_{x+5}/l_{2.5} = (l_{x+5}/l_x)(l_x/l_{2.5}) \quad (1)$$

where the l_{x+5}/l_x denotes the survival ratios in column 5. Enter these values in column 6.

Column 7-8. Interpolate the conditional survival schedule $l_x/l_{2.5}$ for $x = 5, 10, \ldots, 80$. Using the linear interpolation formula

$$l_x/l_{2.5} = 0.5(l_{x-2.5}/l_{2.5} + l_{x+2.5}/l_{2.5}) \quad (2)$$

or, if desired, other more elaborate methods may be applied. The interpolated values are entered in column 8 along with their corresponding ages in column 7.

Column 9. Compute the conditional ${}_5L_x$ values (${}_5L_x/l_{2.5}$) and enter them in column 9. The equation applied in this calculation is:

$${}_5L_x/l_{2.5} = 2.5(l_x/l_{2.5} + l_{x+5}/l_{2.5}) \quad (3)$$

for $x = 5, 10, \ldots, 75$.

Column 10. Given $e_{80} = 6.06$, compute $T_{80}/l_{2.5} = (l_{80}/l_{2.5})e_{80}$ and enter this value in column 10 for age 80. Now fill in T_x values in column 10 for other ages using the equation

$$T_{x-5}/l_{2.5} = T_x/l_{2.5} + {}_5L_{x-5}/l_{2.5}, \quad (4)$$

Column 11. Compute e_x for ages $x = 5, 10, \ldots, 75$ using the equation

$$e_x = (T_x/l_{2.5})/(l_x/l_{2.5}) \quad (5)$$

Enter these values in column 11.

Columns 12-13. Evaluate the accuracy of the estimates of life expectancy. In this example, the estimated values are compared with estimates obtained from civil registration data (column 12) and the deviation between these estimates is shown in column 13.

Note: In this example the expectation of life at age 80, required to initiate the calculation of the $T_x/l_{2.5}$ values, is taken from life tables derived from the registered deaths. See section H for a discussion of how to proceed when an estimate of the uppermost expectation of life has to be obtained from other sources.

TABLE I.2. CENSUS SURVIVAL METHOD FOR ARBITRARY INTERCENSAL INTERVALS APPLIED TO JAPAN: FEMALES, 1960-1970

	Census population						*Estimated conditional life table functions*							
Age group(i)	*1960[a] $P_1(x,5)$*	*1970[b] $P_2(x,5)$*	*Age specific growth rate $r(x,5)$*	*Average annual person years lived $N(x,5)$*	*Age(x)*	*Synthetic survival ratio l_{x+5}/l_x*	*Probability of survival $l_x/l_{2.5}$*	*Age x*	*Probability of survival to age x $l_x/l_{2.5}$*	*Person years lived in age group ${}_5L_x/l_{2.5}$*	*Total person years lived above age x $T_x/l_{2.5}$*	*Life expectancy at age x e_x*	*Life expectancy from civil registration $e_{x(R)}$*	*Deviation (col.13-col.14)*
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*	*(12)*	*(13)*	*(14)*	*(15)*
0-4	3,831,870	4,292,503	0.011352	4,057,830	2.5	1.0429	1.0000	NA	NA	NA	NA	NA	NA	NA
5-9	4,502,304	3,988,292	-0.012123	4,240,107	7.5	0.9635	1.0429	5	1.0215	5.1133	69.0551	67.60	69.45	-2.7
10-14	5,397,061	3,852,101	-0.033724	4,581,245	12.5	0.9082	1.0048	10	1.0238	4.9563	63.9419	62.45	64.62	-3.4
15-19	4,630,775	4,492,096	-0.003040	4,561,084	17.5	1.0976	0.9125	15	0.9587	4.7893	58.9856	61.53	59.72	3.0
20-24	4,193,184	5,347,327	0.024314	4,746,894	22.5	0.9973	1.0016	20	0.9570	4.8932	54.1963	56.63	54.87	3.2
25-29	4,114,704	4,571,868	0.010535	4,339,273	27.5	0.9661	0.9989	25	1.0002	4.9556	49.3031	49.29	50.11	-1.6
30-34	3,770,907	4,190,340	0.010547	3,976,938	32.5	1.0000	0.9650	30	0.9820	4.8676	44.3475	45.16	45.37	-0.5
35-39	3,274,822	4,085,338	0.022114	3,665,156	37.5	0.9884	0.9651	35	0.9651	4.8113	39.4800	40.91	40.65	0.6
40-44	2,744,786	3,674,127	0.029161	3,186,904	42.5	1.0233	0.9539	40	0.9595	4.8112	34.6686	36.13	35.99	0.4
45-49	2,559,755	3,198,934	0.022291	2,867,481	47.5	0.9297	0.9761	45	0.9650	4.7670	29.8574	30.94	31.40	-1.5
50-54	2,160,716	2,648,360	0.020350	2,396,274	52.5	0.9834	0.9075	50	0.9418	4.6043	25.0905	26.64	26.94	-1.1
55-59	1,839,025	2,382,691	0.025899	2,099,137	57.5	0.9375	0.8924	55	0.8999	4.4111	20.4862	22.76	22.64	0.6
60-64	1,494,043	1,970,485	0.027679	1,721,288	62.5	0.9116	0.8366	60	0.8645	4.1603	16.0752	18.59	18.52	0.4
65-69	1,133,409	1,584,699	0.033516	1,346,473	67.5	0.8820	0.7626	65	0.7996	3.7932	11.9148	14.90	14.67	1.6
70-74	870,238	1,172,155	0.029783	1,013,714	72.5	0.7383	0.6726	70	0.7176	3.2556	8.1217	11.32	11.20	1.0
75-79	577,972	736,258	0.024206	653,925	77.5	0.6227	0.4966	75	0.5846	2.4688	4.8661	8.32	8.25	0.9
80-84	313,781	408,191	0.026304	358,919	82.5	NA	0.3092	80	0.4029	NA	2.3973	5.95	5.95	NA
85+	131,547	53,116	-0.090689	86,484	87.5	NA	NA	NA	NA	NA	NA	NA	6.00	NA
													Median absolute per cent deviation	1.1

Source: Population distribution for 1960 and 1970 from :*Historical Statistics of Japan*, volume 1, table 2-9, pp. 66-83 (Japan Statistical Association, Tokyo, 1987).

[a] Reference date: 1 October 1960 (1965.751).

[b] Reference date: 1 October 1970 (1970.751).

Procedure

Columns 1-3. Record the age distribution of the two censuses as shown in columns 1 to 3, taking care to calculate the exact duration of the intercensal period.

Column 4. Compute the age-specific growth rates $r(x,5)$, $x=0, 5, ..., 70$, using the equation

$$r(x,n) = ln[P_2(x,n)/P_1(x,n)]/t, \tag{6}$$

where $N_i(x,n)$ denotes the number of persons aged x to $x+5$ at the I-th census and t denotes the length of the intercensal period. The growth rate for the open-ended interval 85+ may also be calculated, though it is not required in this example. Enter the age specific growth rates in column 4.

Column 5. Compute the average number $N(x,5)$ of person-years lived by each age group during the intercensal period using the formula

$$N(x,5) = [P_2(x,5) - P_1(x,5)]/[tr(x,5)] \tag{7}$$

Enter these in column 5.

Columns 6-7. Compute and enter in columns 6 and 7, the synthetic survival ratios, by age, using the formula

$$\frac{N(x+5,5)exp[(2.5r(x+5,5)]}{N(x,5)exp[(-2.5r(x,5)]} \tag{8}$$

and so on.

Column 8. Compute the conditional survival schedule $l_x/l_{2.5}$, noting that $l_{2.5}/l_{2.5} = 1$ and using the equation

$$l_{x+5}/l_{2.5} = (l_{x+5}/l_x)(l_x/l_{2.5}) \tag{1}$$

where the l_{x+5}/l_x denotes the survival ratios in column 7. Enter these values in column 8.

Columns 9-10. Interpolate the conditional survival schedule $l_x/l_{2.5}$ for $x = 5, 10, ..., 80$. Using the linear interpolation formula

$$l_x/l_{2.5} = 0.5(l_{x-2.5}/l_{2.5} + l_{x+2.5}/l_{2.5}) \tag{2}$$

or, if desired, other more elaborate methods may be applied. The interpolated values are entered in column 10, along with their corresponding ages in column 9.

Column 11. Compute the conditional $_5L_x$ values ($_5L_x/l_{2.5}$) and enter them in column 11. The equation applied in this calculation is:

$$_5L_x/l_{2.5} = 2.5(l_x/l_{2.5} + l_{x+5}/l_{2.5}) \tag{3}$$

for $x = 5, 10, ..., 75$.

Column 12. Given $e_{80} = 6.06$, compute $T_{80}/l_{2.5} = (l_{80}/l_{2.5})e_{80}$ and enter this value in column 12 for age 80. Now fill in T_x values in column 12 for other ages using the equation

$$T_{x-5}/l_{2.5} = T_x/l_{2.5} + {_5L_{x-5}}/l_{2.5}, \tag{4}$$

Column 13. Compute e_x for ages $x = 5, 10, ..., 75$ using the equation

$$e_x = (T_x/l_{2.5})/(l_x/l_{2.5}) \tag{5}$$

Enter these values in column 13.

Columns 14-15. Evaluate the accuracy of the estimates of life expectancy. In this example, the estimated values are compared with estimates obtained from civil registration data (column 14) and the deviation between these estimates is shown in column 15.

Note: In this example the expectation of life at age 80 is given. See section H for a discussion of how to proceed when an estimate of the uppermost expectation of life is not directly available.

TABLE I.3. CENSUS SURVIVAL METHOD FOR TEN YEAR INTERCENSAL INTERVALS APPLIED TO JAPAN: FEMALES, 1960-1970

Age group	*Census population*									*Conditional life table functions*				
	1960 [a] P_1	*1970* [b] P_2	*Age* x	*Census survival ratio* $P_2(x,10)/P_1(x-10)$	*Probability of survival from age 2.5 years* $l_x/l_{2.5}$	*Probability of survival from age 7.5 years* $l_x/l_{7.5}$	*Merged probability of survival* $l_x/l_{2.5}$	*Age* x	*Probability of survival to age x* $l_x/l_{2.5}$	*Person years lived between age x and x+5* ${}_5L_x/l_{2.5}$	*Total person years lived above age x* $T_x/l_{2.5}$	*Estimated life expectancy at age x* e_x	*Life expectancy from civil registration* $e_{x(R)}$	*Deviation (col.13-col.14) per cent*
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*	*(12)*	*(13)*	*(14)*	*(15)*
0-4	3,831,870	4,292,503	2.5	1.0053	1.0000	NA	1.0000	NA	NA	NA	NA	NA	NA	NA
5-9	4,502,304	3,988,292	7.5	0.9977	NA	1.0000	1.0026	5	1.0013	5.0132	69.7500	69.66	69.45	-0.3
10-14	5,397,061	3,852,101	12.5	0.9908	1.0053	NA	1.0053	10	1.0040	5.0170	64.7368	64.48	64.62	0.2
15-19	4,630,775	4,492,096	17.5	0.9873	NA	0.9977	1.0004	15	1.0028	5.0025	59.7198	59.55	59.72	0.3
20-24	4,193,184	5,347,327	22.5	0.9993	0.9960	NA	0.9960	20	0.9982	4.9750	54.7173	54.82	54.87	0.1
25-29	4,114,704	4,571,868	27.5	0.9929	NA	0.9850	0.9876	25	0.9918	4.9583	49.7422	50.15	50.11	-0.1
30-34	3,770,907	4,190,340	32.5	0.9743	0.9953	NA	0.9953	30	0.9915	4.9486	44.7839	45.17	45.37	0.4
35-39	3,274,822	4,085,338	37.5	0.9768	NA	0.9780	0.9806	35	0.9880	4.9079	39.8353	40.32	40.65	0.8
40-44	2,744,786	3,674,127	42.5	0.9649	0.9698	NA	0.9698	40	0.9752	4.8476	34.9274	35.82	35.99	0.5
45-49	2,559,755	3,198,934	47.5	0.9308	NA	0.9553	0.9579	45	0.9638	4.7766	30.0798	31.21	31.40	0.6
50-54	2,160,716	2,648,360	52.5	0.9120	0.9357	NA	0.9357	50	0.9468	4.6512	25.3032	26.73	26.94	0.8
55-59	1,839,025	2,382,691	57.5	0.8617	NA	0.8893	0.8916	55	0.9137	4.4654	20.6521	22.60	22.64	0.2
60-64	1,494,043	1,970,485	62.5	0.7846	0.8533	NA	0.8533	60	0.8725	4.2083	16.1867	18.55	18.52	-0.2
65-69	1,133,409	1,584,699	67.5	0.6496	NA	0.7663	0.7683	65	0.8108	3.8243	11.9784	14.77	14.67	-0.7
70-74	870,238	1,172,155	72.5	0.4691	0.6695	NA	0.6695	70	0.7189	3.2580	8.1541	11.34	11.20	-1.3
75-79	577,972	736,258	77.5	NA	NA	0.4978	0.4991	75	0.5843	2.4771	4.8962	8.38	8.25	-1.6
80-84	313,781	408,191	82.5	NA	0.3140	NA	0.3140	80	0.4066	NA	2.4190	5.95	5.95	NA
85+	131,547	53,116	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
													Median absolute per cent deviation	0.4

Source: Population distribution for 1960 and 1970 from: *Historical Statistics of Japan*, volume 1, table 2-9, pp. 66-83 (Japan Statistical Association, Tokyo, 1987).

[a] Reference date: 1 October 1960 (1960.751).

[b] Reference date: 1 October 1970 (1970.751).

Procedure

Columns 1-3. Record the age distribution of the two censuses as shown in columns 1 to 3.

Columns 4-5. Record the mid-points of the age groups in column 4 and compute census survival ratios, entering them in column 5. The first census survival ratio is the number of persons aged 10-14 at the second census divided by the number aged 0-4 at the first census, and similarly for higher age groups. Note that the last ratio calculated takes the number of persons in the last five-year age group, 80-84 in this case, as its numerator.

Column 6. Compute the conditional survival probabilities $l_x/l_{2.5}$ for $x = 2.5, 12.5, 22.5, ...$ noting that $l_{2.5}/l_{2.5} = 1$ and using the formula

$$l_{x+10}/l_{2.5} = (l_{x+10}/l_x)\,(l_x/l_{2.5}) \qquad (9)$$

for $x = 2.5, 12.5, 22.5, ...$ Enter these values in column 6. Note that x increases by 10 years each time this formula is applied, so that entries are made in every other row.

Column 7. Compute the conditional survival probabilities $l_x/l_{7.5}$ for $x = 7.5, 17.5, 27.5, ...$ noting that $l_{7.5}/l_{7.5} = 1$ and using the formula

$$l_{x+10}/l_{7.5} = (l_{x+10}/l_x)\,l_x/l_{7.5} \qquad (10)$$

for $x = 7.5, 17.5, 27.5, ...$ Enter these values in column 7. Note that x increases by 10 and entries are therefore made in every other row.

Column 8. Compute $l_{7.5}/l_{2.5}$ by interpolating between the first two entries in column 6, *i.e.*, using the formula

$$l_{7.5}/l_{2.5} = 0.5(l_{2.5}/l_{2.5} + l_{12.5}/l_{2.5}) \qquad (12)$$

In this case, the result is (1 + 1.0053)/2 = 1.0026. Column 8 is obtained by multiplying the number resulting from the application of equation (12) by the corresponding value in column 7. Note that this corresponds to recording the estimates of $l_x/l_{2.5}$ from column 6 and obtaining missing values by multiplying the entries in column 7 by 1.0026.

Columns 9-10. Interpolate the conditional survival schedule $l_x/l_{2.5}$ for $x = 5, 10, ..., 80$. Using the linear interpolation formula

$$l_x/l_{2.5} = 0.5(l_{x-2.5}/l_{2.5} + l_{x+2.5}/l_{2.5}) \qquad (2)$$

or, if desired, other more elaborate methods may be applied. The interpolated values are entered in column 10, along with their corresponding ages in column 9.

Column 11. Compute the conditional ${}_5L_x$ values (${}_5L_x/l_{2.5}$) and enter them in column 11. The equation applied in this calculation is:

$${}_5L_x/l_{2.5} = 2.5(l_x/l_{2.5} + l_{x+5}/l_{2.5}) \qquad (3)$$

for $x = 5, 10, ..., 75$.

Column 12. Given $e_{80} = 6.06$, compute $T_{80}/l_{2.5} = (l_{80}/l_{2.5})e_{80}$ and enter this value in column 12 for age 80. Now fill in T_x values in column 12 for other ages using the equation

$$T_{x-5}/l_{2.5} = T_x/l_{2.5} + {}_5L_{x-5}/l_{2.5}, \qquad (4)$$

Column 13. Compute e_x for ages $x = 5, 10, ..., 75$ using the equation

$$e_x = (T_x/l_{2.5})/(l_x/l_{2.5}) \qquad (5)$$

Enter these values in column 13.

Columns 14-15. Evaluate the accuracy of the estimates of life expectancy. In this example, the estimated values are compared with estimates obtained from civil registration data (column 14) and the deviation between these estimates is shown in column 15.

TABLE I.4. CENSUS SURVIVAL METHOD FOR TEN YEAR INTERCENSAL INTERVALS APPLIED TO ZIMBABWE : FEMALES, 1982-1992

Age group	Census population		Age	Census survival ratio	Probability of survival from age 2.5 years	Probability of survival from age 7.5 years	Merged probability of survival	Estimated conditional life table functions				Estimated life expectancy at age x	Age	
	1982[a]	1992[b]	Age					Age	Probability of survival to age x	Person years lived between age x and x+5	Total person years lived above age x			
	P_1	P_2	x	$P_2(x,10)/P_1(x-10)$	$l_x/l_{2.5}$	$l_x/l_{7.5}$	$l_x/l_{2.5}$	x	$l_x/l_{2.5}$	$_5L_x/l_{2.5}$	$T_x/l_{2.5}$	e_x	x	e_5
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
0-4	666,513	798,430	2.5	1.1018	1.0000	NA	1.0000	0	NA	NA	NA	NA	0	NA
5-9	620,383	835,296	7.5	1.0230	NA	1.0000	1.0509	5	1.0254	5.2544	66.5386	64.89	5	64.9
10-14	519,647	734,331	13	1.0100	1.1018	NA	1.1018	10	1.0763	5.4118	61.2842	56.94	10	61.4
15-19	413,331	634,658	18	0.9140	NA	1.0230	1.0751	15	1.0884	5.4558	55.8724	51.33	15	60.3
20-24	364,837	524,836	23	0.8974	1.1128	NA	1.1128	20	1.0939	5.3539	50.4166	46.09	20	59.1
25-29	281,551	377,773	28	0.9250	NA	0.9350	0.9826	25	1.0477	5.0956	45.0627	43.01	25	60.3
30-34	207,121	327,407	33	0.9181	0.9986	NA	0.9986	30	0.9906	4.8608	39.9671	40.35	30	62.4
35-39	170,467	260,436	38	0.8443	NA	0.8649	0.9089	35	0.9537	4.6664	35.1063	36.81	35	63.4
40-44	139,774	190,152	43	1.0577	0.9168	NA	0.9168	40	0.9128	4.3873	30.4399	33.35	40	64.7
45-49	110,583	143,928	48	0.7869	NA	0.7302	0.7674	45	0.8421	4.2765	26.0526	30.94	45	67.8
50-54	91,039	147,839	53	0.9282	0.9697	NA	0.9697	50	0.8685	4.1383	21.7761	25.07	50	65.3
55-59	60,906	87,023	58	0.8386	NA	0.5747	0.6039	55	0.7868	3.8468	17.6378	22.42	55	68.1
60-64	65,374	84,499	63	0.9590	0.9000	NA	0.9000	60	0.7520	3.6379	13.7910	18.34	60	68.3
65-69	38,928	51,075	68	NA	NA	0.4819	0.5064	65	0.7032	3.4699	10.1530	14.44	65	68.4
70-74	30,553	62,691	73	NA	0.8631	NA	0.8631	70	0.6847	NA	6.6831	9.76	70	64.4
75+	46,842	68,635	NA	NA	NA	NA	NA	75	NA	NA	NA	NA	75	NA
													Median	64.6
													0.5 x interquartile range	2.8
													Per cent	4.3

Source: Age distribution data available from http://www.census.gov/ipc/www/idbprint.html. See also, for the 1992 census, *Census 1992: Zimbabwe National Report* (Harare, Central Statistical Office, n.d.), table A1.2, p. 9 and 177. For the 1982 Census, see *1988 Demographic Yearbook*, table 7, pp. 252-253.

[a] Reference date: 18 August 1982.

[b] Reference date: 18 August 1992.

Procedure

Columns 1-3. Record the age distributions from the two censuses as shown in columns 1 to 3.

Columns 4-5. Record the midpoints of the age groups in column 4 and compute census survival ratios, entering them in column 5.

Column 6. Compute the conditional survival probabilities $l_x/l_{2.5}$ for x = 2.5, 12.5,… and enter these in column 6.

Column 7. Compute the conditional survival probabilities $l_x/l_{7.5}$ for $x = 7.5, 17.5, 27.5, \ldots$ noting that $l_{7.5}/l_{7.5} = 1$ and using the formula

$$l_{x+10}/l_{7.5} = (l_{x+10}/l_x)\, l_x/l_{7.5} \qquad (10)$$

for $x = 7.5, 17.5, 27.5, \ldots$. Note that x increases by 10 each time this formula is applied, so that entries are made in every other row.

Column 8. Compute $l_{7.5}/l_{2.5}$ by interpolating between the first two entries in column 6, *i.e.*, using the formula

$$l_{7.5}/l_{2.5} = 0.5(l_{2.5}/l_{2.5} + l_{12.5}/l_{2.5}) \qquad (12)$$

In this case, the result is (1 + 1.1018)/2 = 1.0509.

Column 8 is obtained by recording the estimates of $l_x/l_{2.5}$ from column 6 and, to obtain missing estimates, by multiplying the entries in column 7 by 1.0509.

Column 9-10. Interpolate the conditional survival schedule $l_x/l_{2.5}$ for $x = 5, 10, \ldots, 80$. Using the linear interpolation formula

$$l_x/l_{2.5} = 0.5(l_{x-2.5}/l_{2.5} + l_{x+2.5}/l_{2.5}) \qquad (2)$$

or, if desired, other more elaborate methods may be applied. The interpolated values are entered in column 10, along with their corresponding ages in column 9.

Column 11. Compute the conditional ${}_5L_x$ values (${}_5L_x/l_{2.5}$) and enter them in column 11. The equation applied in this calculation is:

$${}_5L_x/l_{2.5} = 2.5(l_x/l_{2.5} + l_{x+5}/l_{2.5}) \qquad (3)$$

for $x = 5, 10, \ldots, 75$.

Column 12. Given $e_{80} = 6.06$, compute $T_{80}/l_{2.5} = (l_{80}/l_{2.5})e_{80}$ and enter this value in column 12 for age 80. Now fill in T_x values in column 12 for other ages using the equation

$$T_{x-5}/l_{2.5} = T_x/l_{2.5} + {}_5L_{x-5}/l_{2.5}, \qquad (4)$$

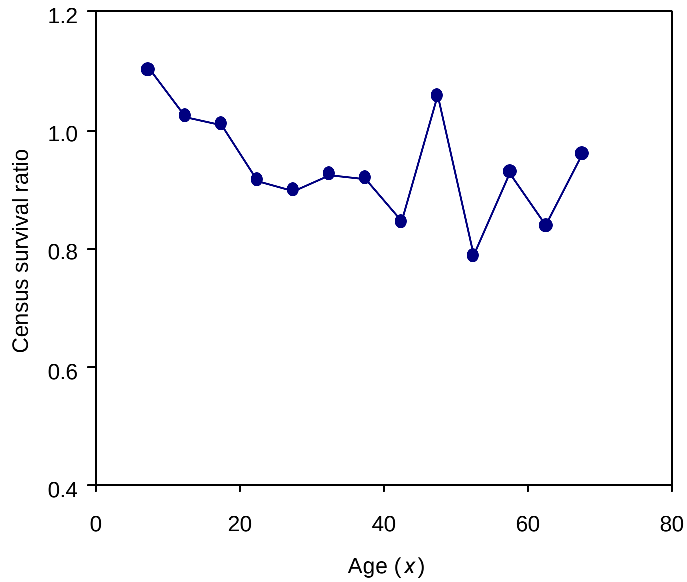
Column 13. Compute e_x for ages $x = 5, 10, \ldots, 75$ using the equation

$$e_x = (T_x/l_{2.5})/(l_x/l_{2.5}) \qquad (5)$$

Enter these values in column 13.

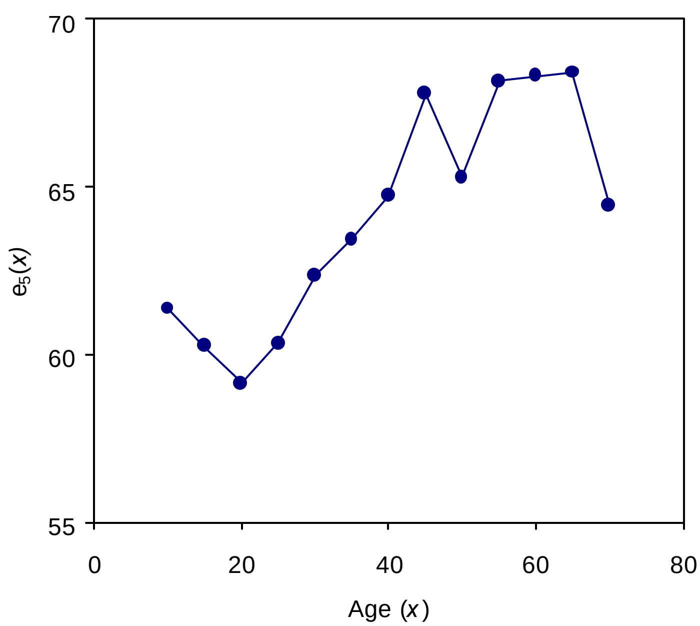
The estimated expectation of life at each age can be translated to a common denominator (in this case expectation of life at age 5 (e_5)) using methods that are described in annex II.

Figure I.1. Census survival method for ten-year intercensal intervals applied to Zimbabwe: Females, 1982-1992: plot of census survival ratios



Source: Survival ratios from column 5 of table I.4.

Figure I.2 Census survival method for ten-year intercensal intervals applied to Zimbabwe: Females, 1982-1992: plot of estimated life expectancy at age 5 years



Source: Column 15 of table I.4.

II. GROWTH BALANCE METHODS

Growth balance techniques are important tools in the adult mortality estimation process because they permit an evaluation of the completeness of death registration data. The original growth balance method formulated by Brass is based on the assumption of a stable, closed population. In that context, the rate of entry into the population aged x and over by those reaching age x is equal to the rate of departure from the same population through death, plus the stable population growth rate, which is the same for all values of x. If it is also assumed that the completeness of death reporting does not vary by age, then an estimate of the completeness of death reporting can be obtained (United Nations, 1983, pp. 139-146). While the Brass formulation has the advantage of requiring, as input, only a single population age distribution and the corresponding distribution of deaths by age, the assumption that the population is stable is often inappropriate in many contexts because of changing fertility and mortality levels and non-negligible levels of migration.

If two census age distributions and a distribution of intercensal deaths are available, a simple reformulation of the original growth balance method eliminates the need for the assumption that the population is stable. The two-census formulation has the further advantage of allowing the estimation of the differential completeness of enumeration between two censuses.

This chapter presents two versions of the growth balance method. The first, the simple growth balance method, uses two age distributions and the distribution of intercensal deaths by age to estimate completeness of death reporting. The second, the general growth balance method (Hill, 1987), utilises the same input data and estimates both the completeness of death reporting and the relative completeness of enumeration of the two censuses.

A. Data required and assumptions

Both methods presented here require two census age distributions and the distribution of intercensal deaths by age. If registered deaths are available for all years of the intercensal period, they may be summed, with interpolation as required, to obtain intercensal registered deaths. If death registration data are missing for some intercensal years, they may be estimated from the available data, either by interpolation between data for available years, or by using the available data to calculate age-specific death rates and then applying these death rates to intercensal person years lived. The latter approach may be used when retrospectively reported deaths from a census or survey are available, although it must be noted that these deaths generally do not refer to calendar years, but to an interval of time (most often 12 months) prior to the census or survey. In this approach the true number of deaths is not estimated but rather, the number that would have been registered or reported in the missing years if data had been available for these years.

Both the simple and general growth balance methods assume that the population experiences no or negligible migration during the intercensal period, at least of persons above some specified lower age limit. This lower age limit can vary and useful results may sometimes be obtained when the age limit is as high as 50 years. Since migration is generally concentrated at young adult ages, the "no migration" assumption is not as limiting as it would otherwise be if this age limit is set above young adulthood. In principle, of course, the method may be applied to populations that are open to migration but for which numbers of intercensal migrants by age are known and can, therefore, be adjusted for. In practice, this data is rarely available.

The simple and general growth balance methods also assume that completeness of death reporting is the same for all age groups above a specified lower limit, and provide estimates of completeness of death reporting only for deaths occurring at or above this age. The general growth balance method further assumes that the completeness of enumeration in the two censuses does not vary by age.

B. The simple growth balance method

The familiar demographic or balancing equation may be written for any time period as

$$P_2 = P_1 + B - D \qquad (1)$$

where P_1 and P_2 denote the number of persons in a population at the beginning and end of some time period, respectively; B denotes the number of births during the period, and D the number of deaths during the period. If the number of births during an intercensal period is known, the number of deaths can be computed directly by rearranging terms in equation (1) to give

$$D = P_1 + B - P_2 \qquad (2)$$

Equation (2) is generally not useful in contexts where deaths are incompletely reported because in these situations, births are likely to be under-reported too.

The balancing equation applies not just to the entire population, but also to the population of persons over any given age. Formula (2), in this instance, can be rewritten as

$$D(x+) = P_1(x+) + N(x) - P_2(x+), \qquad (3)$$

where $P_1(x+)$ and $P_2(x+)$ denote the numbers of persons aged x and over in the population at the beginning and ending of some time period, respectively, $D(x+)$ denotes the number of deaths during the period to persons aged x and over, and $N(x)$ denotes the number of persons reaching exact age x during the period. For x sufficiently above zero, $N(x)$ may be obtained by interpolation between the census age distributions using the approximation

$$N(x) = t0.2[P_1(x-5,5)P_2(x,5)]^{0.5} \qquad (4)$$

where t denotes the length of the intercensal period.

The rationale for formula (4) is as follows. The number $P_1(x-5,5)$ may be taken as an estimate of the number of persons reaching exact age x during the five years following the first census. The estimate is high, if the age data are accurate, because $P_1(x-5,5)$ includes persons who die before reaching exact age x. Similarly, $P_2(x,5)$ provides an estimate of the number of persons reaching exact age x during the five years preceding the second census. This estimate is low if the age data are accurate, because $P_2(x,5)$ excludes persons reaching exact age x during the five years preceding the second census and who die before the second census.

The geometric mean of $P_1(x-5,5)$ and $P_2(x,5)$ in formula (4) therefore estimates the average number of persons reaching exact age x during any five-year period within the intercensal period. The errors in the component terms tend to cancel each other out.

Multiplying this average by 0.2 gives an average number of persons reaching exact age x during any one year of the intercensal period. Multiplying this by the length of the period gives formula (4).

Now let $D^*(x+)$ and $D^c(x+)$ denote, respectively, the reported number of deaths of persons aged x and the number of deaths implied by the census age distributions using formula (3). If reported deaths are a fraction c, constant over all ages, of true deaths, and if the age distribution data are perfectly accurate, then the ratios

$$c(x) = D^*(x+)/D^c(x+) \qquad (5)$$

for all values of x ($5, 10, \ldots$.) will be identical. In practice, there will be some dispersion of values and the completeness of death reporting may be estimated as the median over all or a subset of the $c(x)$ values.

An alternative and essentially equivalent approach is to write equation (5) as

$$D^c(x+) = (1/c)D^*(x+) \qquad (6)$$

and estimate $1/c$ as the slope of a line fitted to the xy-points $(D^*(x+), D^c(x+))$ and passing through the origin. This line-fitting approach is used for the general growth balance method discussed below.

The balancing equation may also be applied to the population of persons aged x to $x+n$. In this case, formula (3) generalises to

$$D(x,n) = [P_1(x,n)+N(x)] - [P_2(x,n)+N(x+n)] \qquad (7)$$

where $P_1(x,n)$ and $P_2(x,n)$ denote, respectively, persons aged x to $x+n$ at the beginning and end of the period, $D(x,n)$ denotes deaths during the period to persons aged x to $x+n$, and $N(x)$ and $N(x+n)$ denote, respectively, the number of persons reaching exact ages x and $x+n$ during the period. The estimated number of deaths calculated from formula (7) is not very robust unless the age interval, n, is large.

The simple growth balance method, like methods based on census survival, is sensitive to differential coverage of the two censuses. If the second census is more (less) completely enumerated than the first, the right hand side of formula (3) will be too small (large).

C. THE GENERAL GROWTH BALANCE METHOD

The general growth balance method proposed by Hill (1987), simultaneously estimates the completeness of death reporting and the relative completeness of enumeration in the two censuses. It is assumed that the completeness of enumeration in the two censuses, like completeness of death reporting, is independent of age.

To apply this method, equation (3) above can be rewritten in the form

$$N(x) - [P_2(x+) - P_1(x+)] = D(x+) \tag{8}$$

and each side of the equation can be divided by the number of person years lived during the intercensal period by persons aged x and over ($PYL(x+)$). Person years lived may be approximated in various ways, but for the present purposes it is necessary to use the geometric mean formula

$$PYL(x+) = t[P_1(x+)P_2(x+)]^{0.5} \tag{9}$$

where t denotes the length of the interval between the two censuses. Dividing through by $PYL(x+)$, reduces equation (8) to:

$$n(x) - r(x+) = d(x+), \tag{10}$$

where, $$n(x) = N(x)/PYL(x+) \tag{11}$$

denotes the rate at which persons enter the population group aged x and over, and

$$r(x+) = [P_2(x+) - P_1(x+)]/PYL(x+) \tag{12}$$

denotes the growth rate of the population aged x and over, and

$$d(x+) = D(x+)/PYL(x+) \tag{13}$$

is the death rate of the population aged x and over.

Formula (12) is equivalent to the standard formula for calculating the growth rate of a population, ($ln[P_2(x+)/P_1(x+)]/t$) if $PYL(x+)$ is calculated by exponential interpolation between $P_2(x+)$ and $P_1(x+)$. The use of formula (9) to compute person years lived requires that the same denominator be used in (12) as in (11) and (13), however, otherwise the identity will not be preserved. The difference between the two approximations for person years lived is generally quite small.

These equations are not immediately useful because the terms refer to true rather than observed quantities. To obtain an equation containing observed quantities, let k_1 and k_2 denote the completeness of enumeration at the first and second censuses, respectively, and let c denote the completeness of reporting of deaths. In view of the uniformity assumptions, the result is the following:

$$P^*_1(x+) = k_1P_1(x+) \tag{14a}$$

$$P^*_2(x+) = k_2P_2(x+) \tag{14b}$$

$$D^*(x+) = cD(x+) \tag{14c}$$

for all x, where $P^*_1(x+)$ denotes the observed value of $P_1(x+)$, $P^*_2(x+)$ the observed value of $P_2(x+)$ and $D^*(x+)$ the observed value of $D(x+)$. From this it follows that

$$P_1(x+) = P^*_1(x+)/k_1 \tag{15a}$$

$$P_2(x+) = P^*_2(x+)/k_2 \tag{15b}$$

$$D(x+) = D^*(x+)/c \tag{15c}$$

for all x.

Now substitute the expressions on the right in (15a-c) in equations (4), (9) and (11-13) above and manipulate as indicated below in formulas (16-21) to arrive at formula (22), which contains only the observed values and parameters.

Substitution in formula (4) gives

$$N(x) = 0.2t\{[P^*_1(x-5,5)/k_1][P^*_2(x,5)/k_2]\}^{0.5}$$

$$= 0.2t\{[(P^*_1(x-5,5)P^*_2(x,5)]/[k_1k_2]\}^{0.5}$$

$$= 0.2\ t[(P^*_1(x-5,5)P^*_2(x,5)]^{0.5}/[k_1k_2]^{0.5}$$

$$= N^*(x)/[k_1k_2]^{0.5}, \qquad (16)$$

where $N^*(x)$ denotes the number of persons reaching exact age x during the intercensal period calculated from the observed population numbers $P^*_1(x-5,5)$ and $P^*_2(x,5)$.

Substitution in formula (9) and similar manipulation gives

$$PYL(x+) = PYL^*(x+)/[k_1k_2]^{0.5} \qquad (17)$$

where $PYL^*(x+)$ denotes persons years lived by the population aged x and over during the intercensal period calculated from the observed age distributions.

From formulas (11), (16) and (17) it can be seen that, subject to the uniformity assumptions, the entry rate $n^*(x) = N^*(x)/PYL^*(x+)$ calculated from the observed age distributions equals the true rate $n(x)$,

$$n(x) = n^*(x) \qquad (18)$$

because the $[k_1k_2]^{0.5}$ terms cancel out on division.

For the growth rate $r(x+)$, substitution in formula (12) and manipulation gives

$$(1/t)\ln\{[P^*_2(x+)/P^*_1(x+)][k_1/k_2]\}$$

$$= (1/t)\ln[P^*_2(x+)/P^*_1(x+)]$$

$$+ (1/t)\ln(k_1/k_2)$$

so that

$$r(x+) = r^*(x+) + (1/t)\ln(k_1/k_2) \qquad (19)$$

where $r^*(x+)$ denotes the growth rate of the population aged x and over calculated from the observed age distributions.

Substitution in formula (13) and manipulation gives

$$d(x+) = D(x+)/PYL(x+)$$

$$= [D^*(x+)/c]/[PYL^*(x+)/(k_1k_2)^{0.5}]$$

$$= [D^*(x+)/PYL^*(x+)][(k_1k_2)^{0.5}/c]$$

$$= d^*(x+)[(k_1k_2)^{0.5}/c] \qquad (20)$$

where $d^*(x+)$ denotes the death rate for the population aged x and over as calculated from the observed numbers of persons and deaths.

Substituting the expressions for $n(x)$, $r(x+)$ and $d(x+)$ given by formulas (18), (19) and (20), respectively, in the rate form of the balancing equation (10) and now gives

$$n(x) - [r^*(x+) + (1/t)\ln(k_1/k_2)] =$$

$$= d^*(x+)[(k_1k_2)^{0.5}/c] \qquad (21)$$

and rearranging terms gives

$$n^*(x) - r^*(x+) = a + bd^*(x+) \qquad (22)$$

where

$$a = \ln(k_1/k_2)]/t \qquad (22a)$$

and

$$b = (k_1k_2)^{0.5}/c. \qquad (22b)$$

Equation (22) contains only the observable quantities $n^*(x)$, $r^*(x+)$ and $d^*(x+)$ and the parameters c, k_1, and k_2.

To estimate values for c, k_1, and k_2 a straight line is fitted to the points

$$(n^*(x) - r^*(x+), d^*(x+)) \qquad (23)$$

to obtain values for the intercept a and the slope b. The ratio k_1/k_2 is then calculated by inverting formula (22a)

$$k_1/k_2 = \exp(ta). \qquad (24)$$

It is not possible to estimate k_1 and k_2 individually because there is no way to distinguish the situation in which both censuses and deaths are under-reported by precisely the same amount from the situation in which both censuses and deaths are completely reported. This is not generally problematic since our aim in the present context is usually to compute death rates, in which equal under-reporting in both censuses and deaths cancel out.

To calculate completeness of death reporting c, however, a value for the product k_1k_2 in the formula

$$c = (k_1k_2)^{0.5}/b, \qquad (25)$$

which follows from (22b), is needed. A convenient way to proceed is to ascertain which of the two k values is larger, arbitrarily set this value equal to one, and then determine the other k value by their ratio. Thus if $k_1/k_2 > 1$, then $k_1 > k_2$ then

$$k_1 = 1 \text{ and } k_2 = 1/(k_1/k_2) \qquad (25a)$$

If $k_1/k_2 < 1$, then $k_1 < k_2$ and we put

$$k_2 = 1 \text{ and } k_1 = k_1/k_2. \qquad (25b)$$

The product k_1k_2 is calculated as the product of these values.

D. Simple growth balance method application: Japan, females, 1960-1970

As in the case of census survival methods, an example is presented using very high quality data both as an illustration and as a test of the method. Census age distributions for females enumerated in the 1960 and 1970 censuses of Japan are used. Both censuses had a reference date of 1 October. Intercensal deaths are available from vital registration data. The data available online from the Berkeley Mortality Database (http://demog.berkeley.edu/wilmoth/mortality/) include, in addition to annual deaths, deaths during the last quarter of each year as well, allowing an exact calculation of numbers of intercensal deaths.

Table II.1 presents the results of applying the simple growth balance method calculations to Japan. Detailed calculations follow the methods and formulas derived in section B and step by step guidance is provided in the notes to table II.1. The ratios in column 11 vary only slightly, with a median of 0.987. This suggests that the registration of deaths is 98.7 per cent complete and that deaths need to be adjusted upwards by 1.3 per cent. Because the simple growth balance method is designed for use in situations in which under-reporting is much higher than the level found for Japan females, the method can be tested by applying the method to the synthetic data in annex table II.1, to determine the performance of the method under "perfect" data reporting conditions. The synthetic data represent approximately the same level of mortality as that of Japan. The application of the method to the synthetic data results in an adjustment factor of 1.0004, thus suggesting that the growth balance method performs well under conditions where the reporting of deaths is close to complete - - as is the case for Japan. Although the simple growth balance method suggests that the reporting of deaths for Japan is fairly complete, the general growth balance method is applied to the same data to assess whether our results were biased by differential completeness of the Japanese censuses.

E. General growth balance method application: Japan, females:1960-1970

Table II.2 shows the results of applying the general growth balance method to the data. The calculations follow the formulas developed in the preceding section, and are detailed in the notes to the table. Figure II.1 shows the scatter plot and residual plot of the (x,y) points $d(x+)$ and $n(x)$-$r(x+)$ for $x = 5, 10, \ldots$. These values are shown in the last two columns of table II.2. The procedure for fitting the line is presented in annex III.

The observed data points fall closely along the fitted line. The residual plot shows that the last two points are outliers, with values relatively far below the fitted line. The intercept and slope of the fitted line are $a = 0.00007$ and slope $b = 1.0070$. From the intercept, calculate, using formula (24),

$$k_1/k_2 = exp(10H0.00007) = 1.0007.$$

Since k_1/k_2 is greater than one, k_1 is bigger than k_2 and

$$k_1 = 1$$

and

$$k_2 = 1/(k_1/k_2)$$

$$= 1/1.0007$$
$$= 0.9993,$$

indicating that the 1970 census achieved a slightly less complete enumeration than the 1960 census. To adjust the 1970 census counts to the same level of completeness as the 1960 counts, based on these results, it is necessary to divide the 1970 counts by 0.9993, *i.e.*, increase them by about 0.07 per cent.

The implied completeness of death reporting, from formula (25), is then

$$c = (k_1 k_2)^{0.5}/b = 0.9930,$$

where, from the preceding paragraph,

$$k_1 k_2 = 0.9993,$$

suggesting that intercensal deaths are under-registered by 0.7 per cent compared to the 1.3 per cent estimated by the simple growth balance method.

The general growth balance method estimates a very slight relative underenumeration in the 1970 census, but mortality is so low that even this slight underenumumeration creates the appearance of many more intercensal deaths and a much higher level of under-registration than is really the case.

F. Simple growth balance method application: Zimbabwe, females, 1982-1992

Tables II.3 and II.4 show the results of the application of the simple growth balance method to census and vital registration data for Zimbabwe. As a preliminary step, table II.3 shows the calculation of estimated intercensal registered deaths.

Death registration data for Zimbabwe are available for 1982, 1986 and 1990-1992. Intermediate calculations are therefore required to obtain an estimate of the deaths that would have been registered over the entire intercensal period. First, it is necessary to estimate registered deaths for 1983-1985 as the average of registered deaths in 1982 and 1986 and registered deaths for 1987-1989 as the average of registered deaths for 1986 and 1990. Intercensal deaths are then estimated as the sum of deaths in the years 1983-1991, (1-0.630) times deaths in 1982 and 0.630 times deaths in 1992. The factor (1-0.630) represents the interval between the 1982 census and the end of calendar year 1982. The factor 0.630 represents the interval between the beginning of calendar year 1992 and the 1992 census. The fraction 0.630 is the decimal equivalent of 18 August, the reference date for both 1982 and 1992 censuses. The procedure for translating dates into decimal fractions of a year is described in annex 1.

Table II.4 shows the results of the application of the simple growth balance method for Zimbabwe. The growth balance calculations indicate an overall completeness of death registration for the intercensal period of 35.9 per cent. The plot of the ratios of reported to estimated deaths $c(x)$ by age is shown in figure II.2. The ratio for age $x=5$ is a clear outlier. The remaining points mostly fall in the range of 0.3 to 0.4. In an intensive analysis it would be desirable to explain the clear pattern of rise and fall in $c(x)$ values with increasing age. In the present context, however, the variation can be accepted as the range of possible error in estimated completeness.

Table II.5 shows a life-table for the intercensal period calculated from adjusted deaths. Calculations are based on standard life table techniques and are detailed in the notes to the table.

It should be noted that small variations in the completeness of death registration have a relatively small effect on the estimated expectation of life at age 5 years. A 10 per cent lower completeness of death registration, for example, decreases the estimated e_5 from 61.3 to 59.9 years, a drop of only 2.3 per cent. Conversely, a 10 per cent higher completeness increases e_5 from 61.3 to 62.6 years, an increase of only 2.1 per cent.

G. General growth balance method application: Zimbabwe, females, 1982-1992

The results of the application of the general growth balance method to data for females enumerated in the Zimbabwe census for 1982-1992 are presented in tables II.6 through II.9 and in figure II.3. Table II.6 shows the preliminary calculations, with the points $d(x+)$,and $n(x+)-r(x+)$ given in columns 13 and 14. Table II.7 shows calculations for obtaining the slope and intercept of the fitted line and the values for the parameters k_1, k_2 and c. The procedure used for fitting the line is described in annex III. Figure II.3 shows the data points, fitted line, and residuals. Table II.8 calculates the adjusted age-specific death rates for the intercensal period, adjusting both the intercensal deaths and the census age distributions. Table II.9 presents the life-table calculated from the adjusted intercensal death rates.

The intercept and slope of the fitted line are $a = 0.00268$ and $b = 2.229$, respectively. From the intercept, calculate, using formula (24),

$$k_1/k_2 = exp(10H0.00268) = 1.0272.$$

Since k_1/k_2 is greater than one, k_1 is bigger than k_2 set $k_1 = 1$ and $k_2 = 1/1.0272 = 0.9735$. The implied completeness of death reporting, from formula (25), is then

$$c = (k_1 k_2)^{0.5}/b = 0.443.$$

The estimated completeness of death registration is thus 44.3 per cent, as compared with 35.9 per cent from the simple growth balance method.

In table II.8 the calculation of adjusted intercensal death rates is complicated by the need to adjust for the completeness of the census count and for the completeness of death registration. In this case the numbers of persons in each age group at the second census are divided by k_2=0.9735 to adjust for the estimated lesser completeness of enumeration in the 1992 census. Estimated registered deaths for the intercensal period are also divided by $c=0.443$ to adjust for incomplete death registration. Death rates are then calculated in the usual way from the adjusted numbers of deaths and person years lived computed from the two-census age distributions. Table II.9 shows a life-table calculated from the adjusted death rates.

TABLE II.1. SIMPLE GROWTH BALANCE METHOD APPLIED TO JAPAN, FEMALES, 1960-1970

Age group	*Age*	*Census population*		*Deaths in intercensal period*	*Population aged x+ in 1960*	*Population aged x+ in 1970*	*Number of persons reaching age x*	*Estimated deaths from age distribution*	*Deaths from registration*	*Ratio of reported to estimated deaths*	*Adjusted deaths*	*Adjusted death rate*
		1960[a] $P_1(x,5)$	*1970*[b] $P2(x,5)$	$D(x,5)$	$P_1(x+)$	$P_2(x+)$	$N(x)$	$D(x+)$	$D^*(x+)$	$D^*(X+)/D(x+)$		
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*	*(12)*	*(13)*
0-4	0	3,831,870	4,292,503	184,456	47,540,899	52,802,276	NA	NA	3,163,894	NA	186,886	0.00461
5-9	5	4,502,304	3,988,292	18,690	43,709,029	48,509,773	7,818,597	3,017,853	2,979,438	0.987	18,936	0.00045
10-14	10	5,397,061	3,852,101	14,762	39,206,725	44,521,481	8,329,065	3,014,309	2,960,748	0.982	14,956	0.00033
15-19	15	4,630,775	4,492,096	24,849	33,809,664	40,669,380	9,847,663	2,987,947	2,945,986	0.986	25,176	0.00055
20-24	20	4,193,184	5,347,327	39,171	29,178,889	36,177,284	9,952,340	2,953,945	2,921,137	0.989	39,687	0.00084
25-29	25	4,114,704	4,571,868	45,996	24,985,705	30,829,957	8,756,868	2,912,616	2,881,966	0.989	46,602	0.00107
30-34	30	3,770,907	4,190,340	52,681	20,871,001	26,258,089	8,304,700	2,917,612	2,835,970	0.972	53,375	0.00134
35-39	35	3,274,822	4,085,338	63,353	17,100,094	22,067,749	7,849,950	2,882,295	2,783,289	0.966	64,187	0.00175
40-44	40	2,744,786	3,674,127	76,826	13,825,272	17,982,411	6,937,467	2,780,328	2,719,936	0.978	77,838	0.00245
45-49	45	2,559,755	3,198,934	99,895	11,080,486	14,308,284	5,926,344	2,698,546	2,643,110	0.979	101,211	0.00354
50-54	50	2,160,716	2,648,360	135,676	8,520,731	11,109,350	5,207,361	2,618,742	2,543,215	0.971	137,463	0.00575
55-59	55	1,839,025	2,382,691	176,369	6,360,015	8,460,990	4,537,981	2,437,006	2,407,539	0.988	178,692	0.00854
60-64	60	1,494,043	1,970,485	233,002	4,520,990	6,078,299	3,807,241	2,249,932	2,231,170	0.992	236,071	0.01376
65-69	65	1,133,409	1,584,699	314,309	3,026,947	4,107,814	3,077,407	1,996,540	1,998,168	1.001	318,449	0.02376
70-74	70	870,238	1,172,155	404,578	1,893,538	2,523,115	2,305,238	1,675,661	1,683,859	1.005	409,907	0.04059
75+	75	1,023,300	1,350,960	1,279,281	1,023,300	1,350,960	NA	NA	1,279,281	NA	1,296,131	0.11024
Total		47,540,899	52,802,276	3,163,894						NA	3,205,567	
									Median	0.987		
									0.5 Interquartile range	0.005		
									Per cent	0.5		

Source: Population age distribution for 1960 and 1970 from: Japan Statistical Association (1987) *Historical Statistics of Japan*, volume 1, tables 2-9, pp. 66-83.

[a] Reference date: 1 October 1960

[b] Reference date: 1 October 1970

Procedure

Columns 1-5. Record the population age distribution at the two censuses and intercensal deaths as shown in table II.1. Intercensal deaths by age were calculated from files in Berkeley Mortality Data Base, http://demog.berkeley.edu/wilmoth/mortality/.

Columns 6-7. Cumulate the population age distributions and intercensal deaths from bottom-up to give the numbers of persons aged x and over at the first and second census.

Column 8. Compute the number of persons reaching exact age x during the intercensal period using the formula

$$N(x) = t0.2[P_1(x\text{-}5,5)P_2(x,5)]^{0.5}, \tag{4}$$

where $x = 5, 10, \ldots$.

Column 9. Compute the estimated number of deaths of persons aged x and over from the input age distributions using the formula

$$D(x+) = P_1(x+) + N(x) - P_2(x+), \tag{3}$$

$x = 5, 10, \ldots$.

Column 10. Enter the deaths by age from civil registration source.

Column 11. Compute the ratio of reported to estimated deaths,

$$c(x) = D^*(x+)/D^c(x+). \tag{5}$$

for ages $x = 5, 10, \ldots$.

Column 12. Calculate the adjusted deaths by dividing the registered intercensal deaths in column 5 by the estimated median ratio in column 11.

Column 13. Calculate the adjusted death rate by dividing the adjusted deaths by person years lived at each age.

TABLE II.2. GENERAL GROWTH BALANCE METHOD APPLIED TO JAPAN, FEMALES, 1960-1970

Age group	*Age x*	*Census population* 1960 [a] $P_1(x,5)$	1970 [b] $P_2(x,5)$	*Deaths in intercensal period* $D(x,5)$	*Population aged x+ in 1960* $P_1(x+)$	*Population aged x+ in 1970* $P_2(x+)$	*Deaths above age x* $D(x+)$	*Person years lived above age x* $PYL(x+)$	*Number of persons reaching age x* $N(x)$	*Entry rate into age x and over* $n(x+)$	*Growth rate of population aged x* $r(x+)$	*Death rate above age x* $d(x+)$	*Difference between entry and growth rate over age x* $n(x+)-r(x+)$
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*	*(12)*	*(13)*	*(14)*
0-4	0	3,831,870	4,292,503	184,456	47,540,899	52,802,276	3,163,894	501,025,715	NA	NA	0.01050	0.00631	NA
5- 9	5	4,502,304	3,988,292	18,690	43,709,029	48,509,773	2,979,438	460,468,791	7,818,597	0.01698	0.01043	0.00647	0.00655
10-14	10	5,397,061	3,852,101	14,762	39,206,725	44,521,481	2,960,748	417,796,776	8,329,065	0.01994	0.01272	0.00709	0.00721
15-19	15	4,630,775	4,492,096	24,849	33,809,664	40,669,380	2,945,986	370,812,361	9,847,663	0.02656	0.01850	0.00794	0.00806
20-24	20	4,193,184	5,347,327	39,171	29,178,889	36,177,284	2,921,137	324,901,978	9,952,340	0.03063	0.02154	0.00899	0.00909
25-29	25	4,114,704	4,571,868	45,996	24,985,705	30,829,957	2,881,966	277,544,269	8,756,868	0.03155	0.02106	0.01038	0.01049
30-34	30	3,770,907	4,190,340	52,681	20,871,001	26,258,089	2,835,970	234,100,962	8,304,700	0.03547	0.02301	0.01211	0.01246
35-39	35	3,274,822	4,085,338	63,353	17,100,094	22,067,749	2,783,289	194,257,711	7,849,950	0.04041	0.02557	0.01433	0.01484
40-44	40	2,744,786	3,674,127	76,826	13,825,272	17,982,411	2,719,936	157,674,260	6,937,467	0.04400	0.02637	0.01725	0.01763
45-49	45	2,559,755	3,198,934	99,895	11,080,486	14,308,284	2,643,110	125,913,756	5,926,344	0.04707	0.02563	0.02099	0.02143
50-54	50	2,160,716	2,648,360	135,676	8,520,731	11,109,350	2,543,215	97,293,259	5,207,361	0.05352	0.02661	0.02614	0.02692
55-59	55	1,839,025	2,382,691	176,369	6,360,015	8,460,990	2,407,539	73,356,679	4,537,981	0.06186	0.02864	0.03282	0.03322
60-64	60	1,494,043	1,970,485	233,002	4,520,990	6,078,299	2,231,170	52,421,302	3,807,241	0.07263	0.02971	0.04256	0.04292
65-69	65	1,133,409	1,584,699	314,309	3,026,947	4,107,814	1,998,168	35,262,069	3,077,407	0.08727	0.03065	0.05667	0.05662
70-74	70	870,238	1,172,155	404,578	1,893,538	2,523,115	1,683,859	21,857,754	2,305,238	0.10547	0.02880	0.07704	0.07666
75+	75	1,023,300	1,350,960	1,279,281	1,023,300	1,350,960	1,279,281	11,757,710	NA	NA	0.02787	NA	NA
Total		47,540,899	52,802,276	3,163,894									

Source: Population age distribution for 1960 and 1970 from: Japan Statistical Association (1987) *Historical Statistics of Japan*, volume 1, tables 2-9, pp. 66-83.

[a] Reference date: 1 October 1960

[b] Reference date: 1 October 1970

Procedure

Columns 1-5. Enter input data from columns 1-5 of table II.1.

Columns 6-8. Cumulate input age distributions and intercensal deaths from bottom to give numbers of persons aged x and over at the first and second census, and numbers of deaths to persons aged x and over during the intercensal period.

Column 9. Compute the number of person years lived by the population aged x and over using the formula

$$PYL(x+) = t[P_1(x+)P_2(x+)]^{0.5} \qquad (9)$$

$x = 0, 5, 10, \ldots.$

Column 10. Compute the number of persons reaching exact age x during the intercensal period using the formula

$$N(x) = t0.2[P_1(x\text{-}5,5)P_2(x,5)]^{0.5}, \qquad (4)$$

$x = 5, 10, \ldots.$

Column 11. Compute the entry rate $n(x+)$ into the population aged x and over by dividing $N(x)$ by the number of person years lived by the population aged x and over, $PYL(x+)$.

Column 12. Compute the growth rates of the population aged x and over using the formula

$$r(x+) = [P_2(x+) - P_1(x+)]/PYL(x+) \qquad (12)$$

$x = 0, 5, 10, \ldots$, where $P_1(x+)$ and $P_2(x+)$ denote the observed numbers of persons aged x and over at the first and second censuses, respectively.

Column 13. Compute the death rate $d^*(x+)$ for the population aged x and over by dividing $D(x+)$ by the number of person years lived by the population aged x and over, $PYL(x+)$.

Column 14. Compute $n(x) - r(x+)$ using the values for $n(x)$ and $r(x+)$ in columns 11 and 12, respectively. Columns 13 and 14 give the x and y points, respectively, for fitting a line to estimate the constant a and slope b of the equation

$$n^*(x) - r^*(x+) = a + bd^*(x+) \qquad (22)$$

TABLE II.3. ESTIMATION OF INTERCENSAL REGISTERED DEATHS, ZIMBABWE, FEMALES, 1982-1992

	Registered deaths					
Age group *(1)*	*1982* *(2)*	*1986* *(3)*	*1990* *(4)*	*1991* *(5)*	*1992* *(6)*	*Estimated total deaths in intercensal period* *(7)*
0-4	3,135	3,276	4,532	5,288	6,247	39,520
5-9	216	219	299	300	385	2,570
10-14	166	171	233	257	301	2,024
15-19	209	232	498	525	627	3,484
20-24	274	322	665	846	1,158	5,038
25-29	298	335	706	922	1,244	5,368
30-34	250	311	692	856	1,322	5,130
35-39	242	305	606	785	1,177	4,714
40-44	273	345	558	716	935	4,591
45-49	214	305	482	584	705	3,853
50-54	355	389	619	662	786	4,925
55-59	233	345	455	559	559	3,864
60-64	468	517	755	814	900	6,212
65-69	276	396	496	546	549	4,232
70-74	303	367	733	769	933	5,224
75+	517	709	913	1,007	1,155	7,820
Total	7,429	8,544	13,242	15,436	18,983	108,569

Source: Registered deaths for 1982 from: United Nations (1985). *Demographic Yearbook*, table 26, pp. 534-535. Registered deaths for 1990-1992 from: unpublished data at the Central Statistical Office, Harare, Zimbabwe.

NOTE: The estimated total deaths in the intercensal period (column 7), is the sum of the fraction of 1982 deaths that occurred during the intercensal period i.e. (1-0.630) multiplied by 7429, plus all deaths occurring between 1983 and 1991, plus the fraction of 1992 deaths, that occurred in the intercensal period; i.e. 0.630 * 18,983. Deaths for 1987-1989 are assumed to be an average of the 1986 and 1990 deaths.

TABLE II.4. SIMPLE GROWTH BALANCE METHOD APPLIED TO ZIMBABWE, FEMALES, 1982-1992

Age group	*Age*	*Census population*		*Deaths in intercensal period*	*Population aged x+ in 1982*	*Population aged x+ in 1992*	*Number of persons reaching age x*	*Estimated deaths from age distribution*	*Deaths from registration*	*Ratio of reported to estimated deaths*	*Adjusted deaths*	*Adjusted death rate*
		1982[a] $P_1(x,5)$	*1992[b]* $P_2(x,5)$	$D(x,5)$	$P_1(x+)$	$P_2(x+)$	$N(x)$	$D(x+)$	$D^*(x+)$	$D^*(x+)/D(x+)$		
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*	*(12)*	*(13)*
0-4	0	666,513	798,430	39,520	3,827,849	5,329,009	NA	NA	108,569	NA	110,084	0.01509
5-9	5	620,383	835,296	2,570	3,161,336	4,530,579	1,492,294	123,051	69,048	0.561	7,159	0.00099
10-14	10	519,647	734,331	2,024	2,540,953	3,695,283	1,349,913	195,583	66,478	0.340	5,637	0.00091
15-19	15	413,331	634,658	3,484	2,021,306	2,960,952	1,148,561	208,915	64,455	0.309	9,704	0.00189
20-24	20	364,837	524,836	5,038	1,607,975	2,326,294	931,517	213,198	60,971	0.286	14,035	0.00321
25-29	25	281,551	377,773	5,368	1,243,138	1,801,458	742,497	184,177	55,933	0.304	14,953	0.00458
30-34	30	207,121	327,407	5,130	961,587	1,423,685	607,229	145,131	50,565	0.348	14,291	0.00549
35-39	35	170,467	260,436	4,714	754,466	1,096,278	464,507	122,695	45,434	0.370	13,131	0.00623
40-44	40	139,774	190,152	4,591	583,999	835,842	360,081	108,238	40,720	0.376	12,787	0.00784
45-49	45	110,583	143,928	3,853	444,225	645,690	283,672	82,207	36,130	0.439	10,734	0.00851
50-54	50	91,039	147,839	4,925	333,642	501,762	255,722	87,602	32,276	0.368	13,717	0.01182
55-59	55	60,906	87,023	3,864	242,603	353,923	178,017	66,697	27,352	0.410	10,764	0.01479
60-64	60	65,374	84,499	6,212	181,697	266,900	143,478	58,275	23,487	0.403	17,303	0.02328
65-69	65	38,928	51,075	4,232	116,323	182,401	115,568	49,490	17,276	0.349	11,788	0.02644
70-74	70	30,553	62,691	5,224	77,395	131,326	98,802	44,871	13,044	0.291	14,551	0.03325
75+	75	46,842	68,635	7,820	46,842	68,635	NA	NA	7,820	NA	21,783	0.03842
Total		3,827,849	5,329,009	108,569							302,421	
								Median estimated completeness		0.359		
								0.5 Interquartile range		0.040		
								Per cent		11.1		

Source: Population age distribution for 1982 and 1992 from: http://www.census.gov/ipc/www/idbprint.html. See also, for the 1992 census: Central Statistical Office (n.d.) *Census 1992: Zimbabwe National Report*, Harare, Zimbabwe, table A1.2, p.9 and p. 177. For the 1982 census see United Nations (1988) *Demographic Yearbook*, table 7, pp. 252-253. Intercensal deaths by age from: http://demog.berkeley.edu/wilmoth/mortality.

[a] Reference date: 18 August 1982

[b] Reference date: 18 August 1992

Procedure

Columns 1-5. Record the population age distribution at the two censuses and intercensal deaths as shown in table II.4. Intercensal deaths by age were calculated from files in Berkeley Mortality Data Base, http://demog.berkeley.edu/wilmoth/mortality/.

Columns 6-7. Cumulate the population age distributions and intercensal deaths from bottom-up to give the numbers of persons aged x and over at the first and second censuses.

Column 8. Compute the number of persons reaching exact age x during the intercensal period using the formula

$$N(x) = t0.2[P_1(x\text{-}5,5)P_2(x,5)]^{0.5}, \qquad (4)$$

where $x = 5, 10, \ldots$

Column 9. Compute the estimated number of deaths of persons aged x and over from the age distributions using the formula

$$D(x+) = P_1(x+) + N(x) - P_2(x+), \qquad (3)$$

$x = 5, 10, \ldots$

Column 10. Enter the deaths by age from civil registration source.

Column 11. Compute the ratio of reported to estimated deaths using

$$c(x) = D^*(x+)/D^c(x+). \qquad (5)$$

for ages $x = 5, 10, \ldots$

Column 12. Calculate the adjusted deaths by dividing the intercensal deaths in column 5 by the median estimated completeness (column 11).

Column 13. Calculate the death rate, adjusted for under registration, by the dividing adjusted deaths in column 12 by the number of person years lived in the corresponding age group. This is calculated as the length of the intercensal period times the geometric mean of the number of persons in the age group at the beginning and end of the period.

TABLE II.5. LIFE-TABLE FOR ZIMBABWE : FEMALES, 1982-1992, BASED ON ADJUSTED DEATHS

Age group *(1)*	*Age specific death rate* $_5m_x$ *(2)*	*Age* x *(3)*	*Probability of dying at age x* $_5q_x$ *(4)*	*Survivors at age x* l_x/l_5 *(5)*	*Person years lived between age x and x+5* $_5L_x/l_5$ *(6)*	*Total person years expected to be lived at above age x* T_x/l_5 *(7)*	*Life expectancy at age x* e_x *(8)*
0-4	0.015090	0	NA	NA	NA	NA	NA
5- 9	0.000990	5	0.004962	1.000000	4.9876	61.3114	61.3
10-14	0.000910	10	0.004560	0.995038	4.9638	56.3238	56.6
15-19	0.001890	15	0.009495	0.990500	4.9290	51.3599	51.9
20-24	0.003210	20	0.016180	0.981095	4.8658	46.4309	47.3
25-29	0.004580	25	0.023165	0.965221	4.7702	41.5652	43.1
30-34	0.005490	30	0.027832	0.942862	4.6487	36.7949	39.0
35-39	0.006230	35	0.031643	0.916620	4.5106	32.1462	35.1
40-44	0.007840	40	0.039984	0.887616	4.3494	27.6357	31.1
45-49	0.008510	45	0.043475	0.852125	4.1680	23.2863	27.3
50-54	0.011820	50	0.060900	0.815079	3.9513	19.1183	23.5
55-59	0.014790	55	0.076789	0.765441	3.6803	15.1670	19.8
60-64	0.023280	60	0.123593	0.706664	3.3150	11.4867	16.3
65-69	0.026440	65	0.141557	0.619325	2.8775	8.1718	13.2
70-74	0.033250	70	0.181322	0.531655	2.4173	5.2943	10.0
75+	0.038420	75	1.000000	0.435254	NA	2.8770	6.61

Source: Age specific death rates from Table II.4, column 13.

Procedure

Columns 1-2. Record ages and age-specific death rates for 5-9 and older age groups from column 13 of table II.4.

Columns 3-4. Compute life table $_5q_x$ values for age intervals x = 5, 10, 15 …. 75 using the formula

$$_5q_x = 5\,_5m_x /[1 - 2.5\,_5m_x]$$

Column 5. Compute l_x/l_5 values by noting that $l_5/l_5=1$ and using the formula

$$l_{x+5} = l_x(1-_5q_x)$$

Column 6. Compute $_5L_x/l_5$ where :

$$_5L_x/l_5 = 2.5(l_x/l_5 + l_{x+5}/l_5)$$

Column 7. Based on a preliminary estimate of e_0 of 57.5 years, put $e_{75} = 6.5$ years. Then compute T_{75}/l_5 as $e_{75}(l_{75}/l_5)$. Now compute T_x/l_5 using the formula

$$T_{x-5}/l_5 = T_x/l_5 + {}_5L_x/l_5,$$

Column 8. Compute e_x for $x = 5, 10, ..., 70$ using the formula

$$e_x = (T_x/l_5)/(l_x/l_5)$$

TABLE II.6. GENERAL GROWTH BALANCE METHOD APPLIED TO ZIMBABWE, FEMALES, 1982-1992

Age group	*Age*	*Census population*		*Deaths in intercensal period*	*Population aged x+ in 1982*	*Population aged x+ in 1992*	*Deaths above age x*	*Person years lived above age x*	*Number of Persons reaching age x*	*Entry rate into age x and over*	*Growth rate of population aged x*	*Death rate above age x*	*Difference between entry and growth rate over age x*
		1982[a]	*1992*[b]										
	x	$P_1(x,5)$	$P_2(x,5)$	$D(x,5)$	$P_1(x+)$	$P_2(x+)$	$D(x+)$	$PYL(x+)$	$N(x)$	$n(x+)$	$r(x+)$	$d(x+)$	$n(x+)-r(x+)$
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*	*(12)*	*(13)*	*(14)*
0-4	0	666,513	798,430	39,520	3,827,849	5,329,009	108,569	45,164,856	NA	NA	0.03324	0.00240	NA
5-9	5	620,383	835,296	2,570	3,161,336	4,530,579	69,048	37,845,320	1,492,294	0.03943	0.03618	0.00182	0.00325
10-14	10	519,647	734,331	2,024	2,540,953	3,695,283	66,478	30,642,357	1,349,913	0.04405	0.03767	0.00217	0.00638
15-19	15	413,331	634,658	3,484	2,021,306	2,960,952	64,455	24,464,239	1,148,561	0.04695	0.03841	0.00263	0.00854
20-24	20	364,837	524,836	5,038	1,607,975	2,326,294	60,971	19,340,689	931,517	0.04816	0.03714	0.00315	0.01102
25-29	25	281,551	377,773	5,368	1,243,138	1,801,458	55,933	14,964,828	742,497	0.04962	0.03731	0.00374	0.01231
30-34	30	207,121	327,407	5,130	961,587	1,423,685	50,565	11,700,414	607,229	0.05190	0.03949	0.00432	0.01240
35-39	35	170,467	260,436	4,714	754,466	1,096,278	45,434	9,094,528	464,507	0.05108	0.03758	0.00500	0.01349
40-44	40	139,774	190,152	4,591	583,999	835,842	40,720	6,986,636	360,081	0.05154	0.03605	0.00583	0.01549
45-49	45	110,583	143,928	3,853	444,225	645,690	36,130	5,355,667	283,672	0.05297	0.03762	0.00675	0.01535
50-54	50	91,039	147,839	4,925	333,642	501,762	32,276	4,091,563	255,722	0.06250	0.04109	0.00789	0.02141
55-59	55	60,906	87,023	3,864	242,603	353,923	27,352	2,930,235	178,017	0.06075	0.03799	0.00933	0.02276
60-64	60	65,374	84,499	6,212	181,697	266,900	23,487	2,202,156	143,478	0.06515	0.03869	0.01067	0.02646
65-69	65	38,928	51,075	4,232	116,323	182,401	17,276	1,456,620	115,568	0.07934	0.04536	0.01186	0.03398
70-74	70	30,553	62,691	5,224	77,395	131,326	13,044	1,008,165	98,802	0.09800	0.05349	0.01294	0.04451
75+	75	46,842	68,635	7,820	46,842	68,635	7,820	567,010	NA	NA	0.03843	NA	NA
Total		3,827,849	5,329,009	108,569									

Source: Population age distribution for 1982 and 1992 from:http://www.census.gov/ipc/www/idbprint.html. See also, for the 1992 census: Central Statistical Office (n.d.) *Census 1992: Zimbabwe National Report*, Harare, Zimbabwe, table A1.2, p.9 and p. 177. For the 1982 census see United Nations (1988) *Demographic Yearbook*, table 7, pp. 252-253. Intercensal deaths by age from: http://demog.berkeley.edu/wilmoth/mortality.

[a] Reference date: 18 August 1982

[b] Reference date: 18 August 1992

Procedure

Columns 1-5. Enter census age distributions and intercensal deaths as shown in Table II.6.

Columns 6-8. Cumulate input age distributions and intercensal deaths from bottom to give the numbers of persons aged x and over at the first and second censuses, and the numbers of deaths to persons aged x and over during the intercensal period.

Column 9. Compute the number of person years lived by the population aged x and over using the formula

$$PYL(x+) = t[P_1(x+)P_2(x+)]^{0.5} \qquad (9)$$

$x = 0, 5, 10, \ldots.$

Column 10. Compute the number of persons reaching exact age x during the intercensal period using the formula

$$N(x) = t0.2[P_1(x-5,5)P_2(x,5)]^{0.5} \qquad (4)$$

$x = 5, 10, \ldots.$

Column 11. Compute the entry rate $n(x+)$ into the population aged x and over by dividing $N(x)$ by the number of person years lived by the population aged x and over, $PYL(x+)$.

Column 12. Compute the growth rates of the population aged x and over using the formula

$$r(x+) = [P_2(x+) - P_1(x+)]/PYL(x+) \qquad (12)$$

$x = 0, 5, 10, \ldots$, where $P_1(x+)$ and $P_2(x+)$ denote the observed numbers of persons aged x and over at the first and second censuses, respectively.

Column 13. Compute the death rate $d^*(x+)$ for the population aged x and over by dividing $D(x+)$ by the number of person years lived by the population aged x and over, $PYL(x+)$.

Column 14. Compute $n(x) - r(x+)$ using the values for $n(x)$ and $r(x+)$ in columns 11 and 12, respectively. Columns 13 and 14 give the x and y points, respectively, for fitting a line to estimate the constant a and slope b of the equation

$$n^*(x) - r^*(x+) = a + bd^*(x+) \qquad (22)$$

TABLE II.7. GENERAL GROWTH BALANCE METHOD APPLIED TO ZIMBABWE, FEMALES, 1982-1992: FITTING A STRAIGHT LINE TO THE DATA POINTS

Index (1)	*Age(x) (2)*	*x-point (3)*	*y-point (4)*	*Intercepts y-bx (5)*	*Slopes (6)*	*y-fitted a+bx (7)*	*Residuals y-(a+bx) (8)*	*Per cent deviation (9)*
1	5	0.00182	0.00325	-0.00082	0.313	0.00675	-0.00350	-107.5
2	10	0.00217	0.00638	0.00155	1.707	0.00752	-0.00113	-17.8
3	15	0.00263	0.00854	0.00267	2.224	0.00855	-0.00001	-0.2
4	20	0.00315	0.01102	0.00400	2.647	0.00971	0.00132	11.9
5	25	0.00374	0.01231	0.00398	2.576	0.01101	0.00130	10.5
6	30	0.00432	0.01240	0.00277	2.250	0.01231	0.00009	0.7
7	35	0.00500	0.01349	0.00236	2.164	0.01382	-0.00032	-2.4
8	40	0.00583	0.01549	0.00250	2.198	0.01567	-0.00018	-1.2
9	45	0.00675	0.01535	0.00031	1.878	0.01772	-0.00237	-15.4
10	50	0.00789	0.02141	0.00383	2.374	0.02026	0.00115	5.4
11	55	0.00933	0.02276	0.00196	2.151	0.02349	-0.00072	-3.2
12	60	0.01067	0.02646	0.00269	2.230	0.02645	0.00001	0.0
13	65	0.01186	0.03398	0.00754	2.639	0.02912	0.00486	14.3
14	70	0.01294	0.04451	0.01567	3.233	0.03152	0.01299	29.2
			Median	0.00268	2.227			
		0.5* Interquartile range		0.00094	0.185			
			Per cent	35.1	8.3			

Source: Age specific estimates of *x* and *y* points from columns 13 and 14 of Table II.6.

Procedure

Columns 1-4. Copy age schedule and x and y points from columns 2, 13 and 14 of table II.6. Note that the entries for age 5 years are indexed as the first record.

Column 5. Calculate the intercepts $y-bx$ for each point, where b denotes the slope.

Column 6. For each point, calculate the slope of the line connecting each point and the point at which the fitted line intersects the y axis. This slope is $(y-a)/x$, where a denotes the y intercept. The median of these values will, in general, be very close, though not necessarily identical to the slope of the fitted line. Their variation is an indicator of how closely the points conform to the fitted line (see details on calculation of slope below).

Columns 7-9. Calculate the fitted y value, $a+bx$, for each point (column 7), the residual, $y-(a+bx)$ (column 8) and the residual as a per cent of the observed y value (column 9).

Calculation of adjustment factors. Calculate k_1, k_2 and c from a and b using formulas (24-26).

Calculation of error indicators. The error indicator for the intercept a is one half the interquartile range of the intercepts in column 5. The error indicator for the slope is taken to be one half the interquartile range of the slopes in column 6. The error indicator for the ratio k_2/k_1 is calculated as one half the absolute value of the difference between the ratio calculated from the intercept minus its error indicator and the ratio calculated from intercept plus its error indicator. The same procedure is used to calculate the error indicators for k_1 and k_2. The error indicator for c is calculated as one half the absolute value of the difference between c, calculated using the ratio k_2/k_1, plus its error indicator divided by the slope b minus its error indicator and the ratio k_2/k_1 minus its error divided by the slope b plus its error indicator.

Calculation of Slope

Group Of Points	Median x-point	Median y-point
Lower 3rd	0.00263	0.00854
Upper 3rd	0.01067	0.02646
Slope	2.229	

Calculation of Adjustment Factors And Error Indicators

Formula	Factor	Error Indicator	Per cent
Slope (b)=[k1*k2)ˆ0.5]/c=	2.229	0.185	8.3
Intercept=In(k1/k2/t=	0.00268	0.00094	35.1
t=	10		
k1/k2=exp(t*Intercept)=	1.0272	0.0190	1.9
k1=	1.0000		
k2=	0.9735		
k1*k2=	0.9735		
c=[(k1*k2)ˆ0.5]/Slope=	0.443	0.074	16.7

TABLE II.8. GENERAL GROWTH BALANCE METHOD APPLIED TO ZIMBABWE, FEMALES, 1982-1992: CALCULATION OF ADJUSTED INTERCENSAL DEATH RATES

Age group	*Census population*		*Deaths in intercensal period*	*Adjusted population*		*Adjusted intercensal deaths*	*Adjusted intercensal person years lived*	*Adjusted intercensal death rate*
	1982	*1992*		*1982*	*1992*			
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*
0-4	666,513	798,430	39,520	666,513	820,164	89,210	7,393,578	0.012066
5- 9	620,383	835,296	2,570	620,383	858,034	5,801	7,295,956	0.000795
10-14	519,647	734,331	2,0 24	519,647	754,320	4,568	6,260,832	0.000730
15-19	413,331	634,658	3,484	413,331	651,934	7,864	5,190,997	0.001515
20-24	364,837	524,836	5,038	364,837	539,123	11,373	4,434,997	0.002564
25-29	281,551	377,773	5,368	281,551	388,056	12,117	3,305,413	0.003666
30-34	207,121	327,407	5,130	207,121	336,319	11,581	2,639,294	0.004388
35-39	170,467	260,436	4,714	170,467	267,525	10,641	2,135,514	0.004983
40-44	139,774	190,152	4,591	139,774	195,328	10,362	1,652,325	0.006271
45-49	110,583	143,928	3,853	110,583	147,846	8,698	1,278,642	0.006803
50-54	91,039	147,839	4,925	91,039	151,863	11,116	1,175,817	0.009454
55-59	60,906	87,023	3,864	60,906	89,392	8,723	737,869	0.011822
60-64	65,374	84,499	6,212	65,374	86,799	14,022	753,286	0.018614
65-69	38,928	51,075	4,232	38,928	52,465	9,553	451,924	0.021138
70-74	30,553	62,691	5,224	30,553	64,398	11,792	443,571	0.026584
75+	46,842	68,635	7,820	46,842	70,503	17,652	574,674	0.030717

Source: Population age distribution for 1982 and 1992 from: http://www.census.gov/ipc/www/idbprint.html. See also, for the 1992 census: Central Statistical Office (n.d.) *Census 1992: Zimbabwe National Report*, Harare, Zimbabwe, table A1.2, p.9 and p. 177. For the 1982 census see United Nations (1988) *Demographic Yearbook*, table 7, pp. 252-253. Intercensal deaths by age from: http://demog.berkeley.edu/wilmoth/mortality.

Procedure

Columns 1-3. Input age distributions from the two censuses as shown in Table II.8.

Column 4. Input the reported intercensal deaths as shown in column 4.

Column 5. Divide census numbers in column 2 by k_1 to adjust for relative under enumeration. *Note:* This step is necessary only if k_1 1. In this case, $k_1 = 1$ so values remain unchanged.

Column 6. Divide census numbers in column 3 by k_2 to adjust for relative under enumeration. *Note:* This step is necessary only if k_2 1. In this case $k_2 = 0.9735$.

Column 7. Divide the reported deaths in column 4 by c to adjust for under-reporting of deaths. In this case $c = 0.443$.

Column 8. Calculate the number of person years lived in each age group during the intercensal period as the length of the period times the geometric mean of the adjusted numbers in the age group at the beginning and end of the period.

Column 9. Calculate the age-specific death rates by dividing adjusted deaths in column 7 by adjusted person years lived in column 8.

TABLE II.9. GENERAL GROWTH BALANCE METHOD APPLIED TO ZIMBABWE, FEMALES, 1982-1992: LIFE-TABLE BASED ON DEATH RATES ADJUSTED FOR UNDER-REGISTRATION

Age group *(1)*	*Age specific death rate* $_5m_x$ *(2)*	*Age* x *(3)*	*Probability of dying at age x* $_5q_x$ *(4)*	*Survivors at age x* l_x/l_5 *(5)*	*Person years lived between age x and x+5* $_5L_x/l_5$ *(6)*	*Total person years expected to be lived above age x* T_x/l_5 *(7)*	*Life expectancy at age x* e_x *(8)*
0-4	0.012066	0	NA	NA	NA	NA	NA
5- 9	0.000795	5	0.003983	1.000000	4 .9900	64.1371	64.1
10-14	0.000730	10	0.003657	0.996017	4.9710	59.1471	59.4
15-19	0.001515	15	0.007604	0.992375	4.9430	54.1761	54.6
20-24	0.002564	20	0.012903	0.984829	4.8924	49.2331	50.0
25-29	0.003666	25	0.018500	0.972122	4.8157	44.3407	45.6
30-34	0.004388	30	0.022183	0.954138	4.7178	39.5251	41.4
35-39	0.004983	35	0.025229	0.932972	4.6060	34.8073	37.3
40-44	0.006271	40	0.031854	0.909434	4.4747	30.2013	33.2
45-49	0.006803	45	0.034604	0.880465	4.3262	25.7265	29.2
50-54	0.009454	50	0.048414	0.849997	4.1471	21.4004	25.2
55-59	0.011822	55	0.060910	0.808845	3.9211	17.2532	21.3
60-64	0.018614	60	0.097612	0.759579	3.6125	13.3322	17.6
65-69	0.021138	65	0.111587	0.685434	3.2360	9.7197	14.2
70-74	0.026584	70	0.142383	0.608949	2.8280	6.4837	10.6
75+	0.030717	75	1.000000	0.522245	NA	3.6557	7.00

Source: Age specific death rates from column 9 of Table II.8.

Procedure

Columns 1-2. Record ages and adjusted age-specific death rates for ages 5-9 and older age groups. In this case data are from column 9 of table II.8.

Columns 3-4. Compute life table $_5q_x$ values using the formula

$$_5q_x = 5\,_5m_x /[1 - 2.5\,_5m_x]$$

Column 5. Compute l_x/l_5 values by noting that $l_5/l_5=1$ and using the formula

$$l_{x+5} = l_x(1-{}_5q_x),$$

Column 6. Compute $_5L_x/l_5$ where :

$$_5L_x/l_5 = 2.5(l_x/l_5 + l_{x+5}/l_5)$$

Column 7. Based on a preliminary estimate of e_0 of 65 years, put $e_{75} = 7$ years. Then compute T_{75}/l_5 as $e_{75}(l_{75}/l_5)$. Now compute T_x/l_5 using the formula

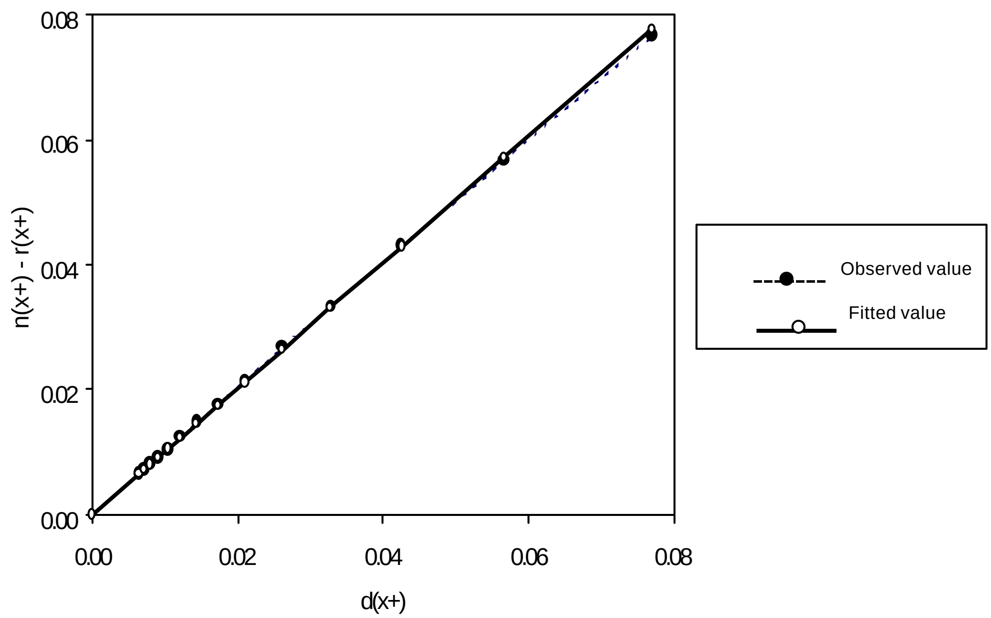
$$T_{x-5}/l_5 = T_x/l_5 + {}_5L_x/l_5,$$

Column 8. Compute e_x for $x = 5, 10, ..., 70$ using the formula

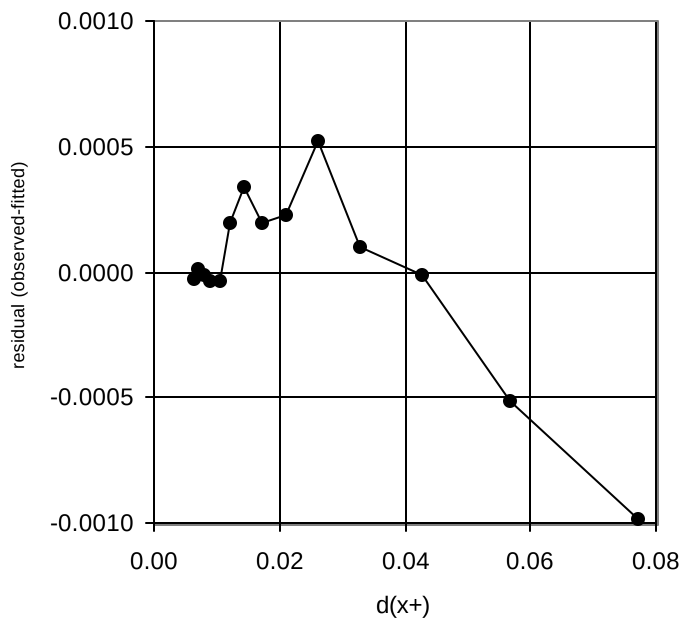
$$e_x = (T_x/l_5)/(l_x/l_5)$$

Figure II.1. General growth balance method applied to Japan, females, 1960-1970

A. Data points and fitted line

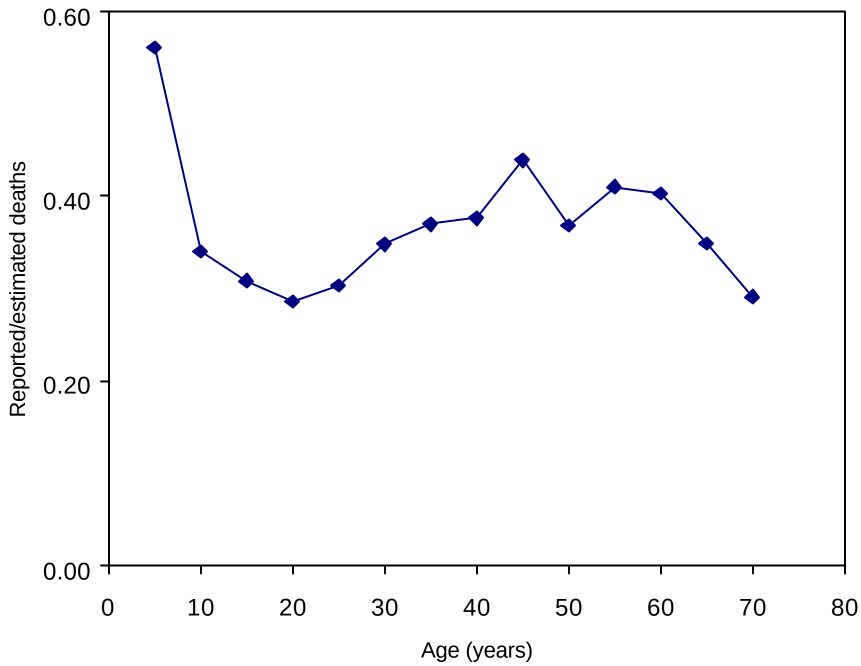


B. Residuals



Source: Table II.2, columns 13 and 14.

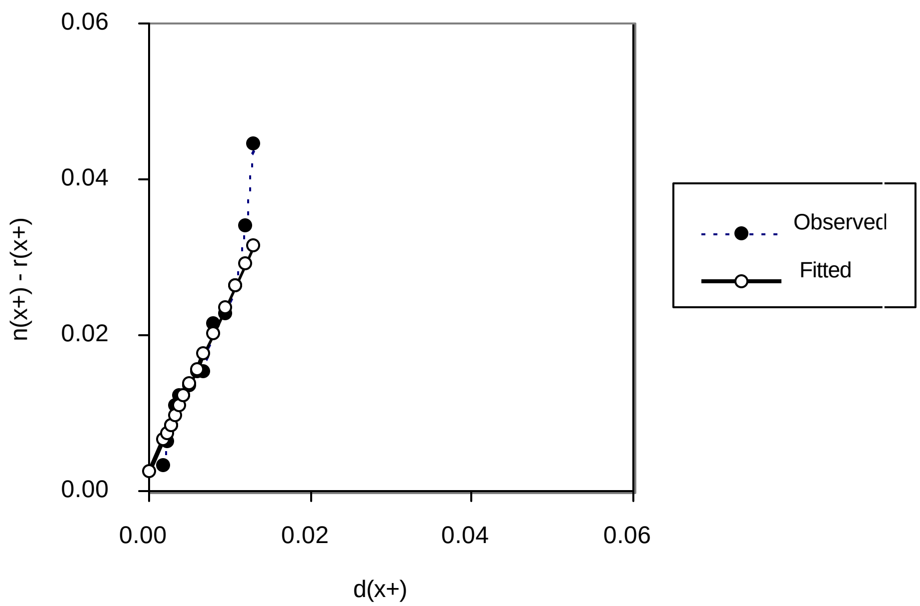
Figure II.2. Simple growth balance method applied to Zimbabwe, females, 1982-1992: plot of ratios indicating completeness of death reporting



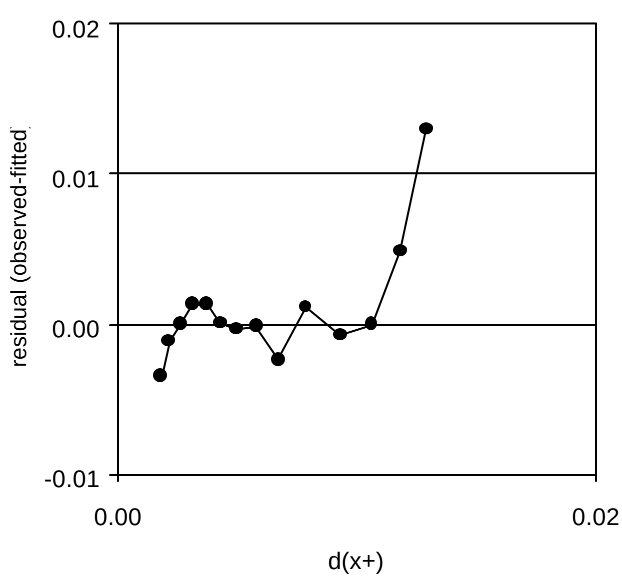
Source : Table II.4, column 11.

Figure II.3. General growth balance method applied to Zimbabwe, females, 1982-1992: scatter plot, fitted line and residuals

A. Data points and fitted line



B. Residuals



Source: Columns 13 and 14 of table II.6

III. THE EXTINCT GENERATIONS METHOD

As with the growth balance methods described in the previous chapter, the extinct generations method estimates adult mortality from two census age distributions and the distribution of intercensal deaths. It takes the same data as the growth balance methods of the preceding chapter and it assumes that migration is negligible and that any under-reporting of deaths is uniform above a certain specified age. In other respects, however, the extinct generations method is quite different from growth balance methods and it may give substantially different results if input data are not perfectly accurate, and/or if the assumptions of the method are violated. The extinct generations method, therefore, indirectly provides a test of whether the data are accurate and whether the assumptions are valid.

A. Stationary population case

Although the idea of the method is simple, the most general implementation involves moderately complicated formulas. It is useful to begin with the simple case of a stationary population, for which the simplicity of the ideas is evident. A stationary population is one that is closed to migration and which experiences constant mortality risks and numbers of births over time. Since everyone dies eventually, the number of persons aged *x* in a population at any given time *t* equals the number of deaths experienced by this cohort from time *t* forward. Therefore,

$$N(x,t) = \int_0^\infty D(x+y,t+y)dy \qquad (1)$$

where $N(x,t)$ denotes the number of people aged *x* at time *t* and $D(x,t)$ denotes the number of deaths at exact age *x* at time *t*.

In a stationary population the number of deaths that will occur at time $t+y$ to the cohort aged *x* at time *t* equals the number of deaths at time *t* to persons aged $x+y$, *i.e.*,

$$D(x+y,t+y) = D(x+y,t) \qquad (2)$$

Substituting (2) into formula (1) yields

$$N(x,t) = \int_0^\infty D(x+y,t)dy \qquad (3)$$

The integral on the right is simply the number of deaths to persons aged *x* and over at time *t*. In application, the integral on the right represents deaths to persons aged *x* and over during a given year, or other time period, and $N(x,t)$ represents the number of persons reaching exact age *x* during this time period.

The idea of the method is to compare an estimate of $N(x,t)$ derived from a census age distribution (denoted as $N^*(x,t)$), with $N(x,t)$ estimated from reported deaths (denoted as $N^d(x,t)$).

$$N^d(x,t) = \int_0^\infty D^*(x+y,t)dy \qquad (4)$$

where D^* represents the number of reported deaths. If deaths are incompletely reported, $N^d(x,t)$ will be smaller than $N^*(x,t)$ by an amount reflecting the extent of under-reporting. The extent of under-reporting can be expressed as a ratio:

$$c(x) = N^d(x,t)/N^*(x,t) \qquad (5)$$

If both the age distribution and the deaths were perfectly reported, and if the population were indeed stationary, these ratios would be equal to one. If the age distribution is accurately reported and deaths are under-reported, but by the same fraction at every age, these ratios will be equal to the completeness of death reporting.

B. Stable population case

The formulas generalise easily to the case of a stable population, which is a population that experiences constant risks of mortality and exponentially increasing births, and that is closed to migration.

In a stable population, the number of persons at every age grows exponentially, and since mortality risks are constant, deaths at any age grow exponentially as well. For a stable population, therefore, deaths at time $t+y$ to the cohort of persons aged *x* at time *t* may be expressed as

$$D(x+y,t+y) = D(x+y,t)e^{ry} \qquad (6)$$

where r is the stable growth rate. Substituting the right hand side of formula (6) for the right hand side of formula (3) gives

$$N(x,t) = \int_0^\infty D(x+y,t)e^{ry}dy. \quad (7)$$

As in the stationary population case, the values of $N^*(x,t)$ (from a census age distribution) and $N^d_{(x,t)}$ (from reported deaths) can be compared using

$$N^d(x,t) = \int_0^\infty D^*(x+y,t)e^{ry}dy \quad (8)$$

and the ratios

$$c(x) = N^d(x,t)/N^*(x,t) \quad (9)$$

can be computed to assess the relative completeness of reporting deaths.

C. CLOSED POPULATION WITH CONSTANT MORTALITY

The generalisation to a closed population subject to constant mortality is more difficult. If mortality risks are constant, deaths at age x grow at the same rate as the population at age x. The stable population formula

$$N(x,t+y) = N(x,t)e^{ry} \quad (10)$$

generalises to

$$N(x,t+y) = N(x,t)\exp\{\int_0^y r(x,t+z)dz\} \quad (11)$$

where $r(x,t)$ denotes the growth rate of the population aged x at time t. Note that the exponential term on the right simplifies to e^{ry} if the growth rate is constant over time.

If mortality risks are constant and there is no migration, formula (11) implies the corresponding relationship for deaths at any age. Therefore

$$D(x,t+y) = D(x+y,t)\exp\{\int_0^y r(x+y,t+z)dz\} \quad (12)$$

Substituting the right hand side of this formula in formula (3) gives

$$N(x,t) = \int_0^\infty D(x+y,t)\exp[\int_0^y r(x+y,t+z)dz]dy. \quad (13)$$

This formulation is not immediately useful, however, because the future growth rates $r(x+y,t+z)$ of the population aged $x+y$ will not be known. If mortality risks are constant, however, then

$$r(x+y,t+z) = r(x+y-z,t) \quad (14)$$

so that

$$\int_0^y r(x+y,t+z)dz = \int_0^y r(x+y-z,t)dz$$

$$= \int_0^y r(x+z,t)dz \quad (15)$$

Substitution in (13) yields

$$N(x,t) = \int_0^\infty D(x+y,t)\exp[\int_0^y r(x+z,t)dz]dy \quad (16)$$

This expression allows the age specific growth rates under the inner integral to be approximated by intercensal age-specific growth rates.

As before, take $N^*(x,t)$ from census age data, calculate the corresponding numbers of persons reaching age x implied by reported deaths, as follows:

$$N^d(x,t) = \int_0^\infty D^*(x+y,t)\exp[\int_0^y r(x+z,t)dz]dy, \quad (17)$$

and then calculate the ratios

$$c(x) = N^d(x,t)/N^*(x,t) \quad (18)$$

If the age distribution and deaths are both perfectly reported, and if the population is indeed closed, these ratios will be equal to one. If the age distribution is correctly reported and the population is closed to migration, but deaths are under-reported uniformly over all ages, the ratios will be constant and be equal to the fraction of deaths that are reported. Variation in the $c(x)$ values with x indicates some departure from these assumptions.

In practice, of course, age distributions are always subject to some degree of error. There will always be some departure from uniformly under-reported deaths. There may also be some degree of migration, although levels may be difficult to determine because of data limitations. The assumption of uniform under-reporting of deaths with age is particularly likely to break down for infant and child deaths. It is therefore customary, when applying this method, always to consider only the population aged *5* (or some higher age) and over.

D. APPLICATION TO INTERCENSAL DEATHS

The formulas of the preceding sections all refer to a particular time t. In application, however, data will be given for an intercensal time period, generally five to ten years. In application, $N(x,t)$, $r(x,t)$ and $D^*(x+y,t)$ are replaced by $N(x)$, $r(x)$, and $D^*(x)$, where $N(x)$ denotes the number of persons reaching exact age x during the intercensal period, $r(x)$ denotes the growth rate of the population aged x during the intercensal period, and $D^*(x)$ the number of deaths at exact age x during the intercensal period.

The number of persons reaching exact age x during the intercensal period is estimated as:

$$N(x) = t0.2[P_1(x-5,5)P_2(x,5)]^{0.5} \qquad (19)$$

in a manner similar to formula (4) of chapter III. The number of persons reaching exact age x implied by the number of intercensal deaths is calculated using formula (16), written now without the time variable t, as

$$N^d(x) = I_0{}^4 D^*(x+y)exp[I_0{}^y r(x+z)dz]dy \qquad (20)$$

To obtain a numerical approximation for use with five-year age group data put x to x-5 in formula (20) and partition the interval of integration to yield the sum of two terms,

$$I_0{}^5 D^*(x!5+y)exp[I_0{}^y r(x!5+z)dz]dy \qquad (20a)$$

and

$$I_5{}^4 D^*(x!5+y)exp[I_0{}^y r(x!5+z)dz]dy. \qquad (20b)$$

Formula (20a) may be approximated by

$$D(x-5,5)exp\{2.5r(x-5,5)\}, \qquad (21a)$$

where $D(x,5)$ denotes the number of intercensal deaths between age x and age $x+5$ and $r(x,5)$ denotes the intercensal growth rate for the same age group. Formula (20b) may be approximated by

$$N(x)exp[5r(x-5,5)], \qquad (21b)$$

and therefore

$$N(x-5) = N(x)exp[5r(x-5,5)] + D(x-5,5)exp[2.5r(x-5,5)] \qquad (22)$$

To calculate $N(x)$ first estimate an initial value of $N(x)$ for the largest possible multiple of five allowed by available age data and then apply formula (22) to obtain the values for younger ages.

To estimate the initial value of $N(x,t)$ for an old age x, Bennett and Horiuchi (1981) propose the formula

$$N(x) = D(x+)\{exp[r(x+)e_{(x)}] - [(r(x+)e_{(x)}]^2/6\} \qquad (23)$$

where $D(x+)$ denotes reported intercensal deaths over age x, $r(x+)$ denotes the intercensal growth rate of the population aged x and over, and $e_{(x)}$ the expectation of life at age x. They propose that $e_{(x)}$ be taken from a model life table with a suitable level of mortality. They note that although in some cases a value of x may be somewhat arbitrary, the resulting estimates of completeness will not be significantly affected.

E. APPLICATION TO JAPAN, FEMALES, 1960-1970

Table III.1 applies the extinct generations method to data for females enumerated in Japan's 1960 and 1970 censuses. The known expectation of life at age 75 (8.25 years), is used in formula (20). The completeness of registration, as indicated by the median of the $c(x)$ ratios over all ages, is 0.9776. This suggests an under-registration of deaths of 2.24 per cent.

An application of the extinct generations method to the synthetic data given in annex table II.5 yields an adjustment factor for deaths of 1.0004, suggesting that the precision of the method in ideal circumstances is sufficiently high to estimate under-registration of this magnitude. The extinct generations estimate of mortality for Japan is substantially higher than the simple and general growth balance methods of the last chapter, however, suggesting that either there is some inaccuracy in the input data, aside from slight under-registration of deaths, or that the assumptions of the method are violated to some degree.

The results of applying the simple growth balance methods in the last chapter indicated that there was a slight underenumeration in the 1970 census relative to the 1960 census and that this resulted in an underestimate of completeness of registration. Using again the synthetic data in annex table II.5, but reducing the age distribution at the second census by 0.07 per cent results in a deaths adjustment factor of 0.975, close to that in table III.1. It may be inferred that a very small underenumeration in the 1970 census, relative to the 1960 census, could create the appearance of more than 2 per cent under-registration of deaths in the intercensal period even if deaths are completely reported. The result of the extinct generations method should therefore not necessarily be interpreted to mean that deaths in Japan during this period were underenumerated by the indicated magnitude.

Close scrutiny of the ratios in column 9 of table III.1 shows that they vary somewhat erratically with a slight downward trend from ages 5 to 45 years, and then rise sharply from ages 45 to 70 years. Applying the method to the synthetic data of annex table II-1 shows the same rise in $c(x)$ values with increasing age. The pattern results from a slight imprecision of the numbers of persons reaching exact age x during the intercensal period as estimated by formula (19). Where the age distribution is approximately linear, this formula gives a very good result. At older ages, however, the curvilinearity of the survival schedule results in a corresponding curvilinearity of the age distribution. Numbers of survivors at each age reduce rapidly at these ages. As a result, the formula underestimates the number of persons reaching each exact age, with the effect increasing with age. The magnitudes involved, about 1.5 per cent, are small however, and would be dwarfed by other errors in many applications.

F. Application to Zimbabwe, females, 1982-1992

In table III.2 the extinct generations method is applied to data for Zimbabwe females for 1982-1992. The estimated completeness of death registration for the intercensal period is 27.6 per cent. This is much lower than the 36 to 44 per cent given by the growth balance methods. However, the ratios in column 9 of table III.2, (plotted in figure III.1), generally fall within a much narrower range that the ratios for the simple growth balance method shown in figure II.2. This suggests that the extinct generations method yields better result.

Table III.3 shows the life table calculated from deaths and age-specific death rates, adjusted for the level of under-registration estimated in columns 10 and 11 of table III.2. Because the extinct generations method estimates the completeness of death registration to be lower than that estimated by growth balance methods, the resultant life expectancies are also lower. The expectation of life at age 5, for example, is 57.6 years, as compared with 61.3 years estimated by the simple growth balance method and 64.1 years estimated by the general growth balance method.

TABLE III.1. THE EXTINCT GENERATIONS METHOD APPLIED TO JAPAN, FEMALES, 1960-1970

Age group	Age (x)	Census population 1960 [a] $P_1(x,5)$	Census population 1970 [b] $P_2(x,5)$	Intercensal deaths $D(x,5)$	Age specific growth rate $r(x,5)$	Number reaching age x as estimated from deaths $N^d(x)$	Number reaching age x as estimated from age distribution $N^*(x)$	Ratio(c(x)) $N^d(x)/N^*(x)$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
0-4	0	3,831,870	4,292,503	184,456	0.011352	8,314,213	NA	NA
5- 9	5	4,502,304	3,988,292	18,690	-0.012123	7,676,158	7,818,597	0.9818
10-14	10	5,397,061	3,852,101	14,762	-0.033724	8,136,558	8,329,065	0.9769
15-19	15	4,630,775	4,492,096	24,849	-0.003040	9,614,921	9,847,663	0.9764
20-24	20	4,193,184	5,347,327	39,171	0.024314	9,737,170	9,952,340	0.9784
25-29	25	4,114,704	4,571,868	45,996	0.010535	8,585,701	8,756,868	0.9805
30-34	30	3,770,907	4,190,340	52,681	0.010547	8,100,333	8,304,700	0.9754
35-39	35	3,274,822	4,085,338	63,353	0.022114	7,632,934	7,849,950	0.9724
40-44	40	2,744,786	3,674,127	76,826	0.029161	6,773,998	6,937,467	0.9764
45-49	45	2,559,755	3,198,934	99,895	0.022291	5,783,512	5,926,344	0.9759
50-54	50	2,160,716	2,648,360	135,676	0.020350	5,079,065	5,207,361	0.9754
55-59	55	1,839,025	2,382,691	176,369	0.025899	4,458,744	4,537,981	0.9825
60-64	60	1,494,043	1,970,485	233,002	0.027679	3,751,859	3,807,241	0.9855
65-69	65	1,133,409	1,584,699	314,309	0.033516	3,049,519	3,077,407	0.9909
70-74	70	870,238	1,172,155	404,578	0.029783	2,289,954	2,305,238	0.9934
75+	75	1,023,300	1,350,960	1,279,281	0.027778	1,597,571	NA	NA
Total		47,540,899	52,802,276	3,163,894			Median	0.9776
							0.5 *interquartile range	0.0032
							Percentage	0.3

Source: Population age distribution for 1960 and 1970 from: Japan Statistical Association (1987) *Historical Statistics of Japan*, volume 1, tables 2-9, pp. 66-83.
[a] Reference date : 1 October 1960.
[b] Reference date : 1 October 1970.

Procedure

Columns 1-5. Enter input data, cumulated census age distributions and average annual intercensal deaths as shown.

Column 6. Compute the age-specific growth rates using $[ln(P_1(x,5)/P_2(x,5))]/t$, where t is the length of the intercensal period and *ln* denotes natural logarithm.

Column 7. Interpolate the value of e_{75}, the expectation of life at age 75, from civil registration data for 1960, 1965 and 1970 and compute the value of the last entry in column 7, $N(75)$, using the formula

$$N^d(75) = D(75+)\{exp[r(75+)e_{75}] - [(r(75+)e_{75}]^2/6 \qquad (23)$$

Then compute the values of $N^d(70)$, $N^d(65)$, ..., from the formula

$$N^d(x-5) = N^d(x)exp[5r(x,5)] + + D(x-5,5)exp[2.5r(x,5)] \qquad (22)$$

where $r(x,5)$ denotes the growth rate for the age interval x to $x+5$.

Column 8. Compute the average number of persons in the x to $x+4$ age group during the intercensal period using the formula

$$N^*(x) = t0.2[P_1(x-5,5)P_2(x,5)]^{0.5} \qquad (19)$$

Column 9. Compute the ratios of the $N^d(x)$ values in column 7 to the $N^*(x)$ values in column 8.

TABLE III.2. THE EXTINCT GENERATIONS METHOD APPLIED TO ZIMBABWE, FEMALES, 1982-1992

Age group	*Age x*	*Census population 1982*[a]	*1992*[b]	*Intercensal deaths*	*Age specific growth rate*	*Number reaching age x from deaths as estimated*	*Number reaching age x as estimated from age distribution*	*Ratio c(x)*	*Adjusted deaths*	*Adjusted death rate*
	x	$P_1(x,5)$	$P_2(x,5)$	$D(x,5)$	$r(x,5)$	$N^*(x)$	$N(x)$	$N^d(x)/N^*(x)$		
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*
0-4	0	666,513	798,430	39,520	0.018059	520,643	NA	NA	143,114	0.01962
5- 9	5	620,383	835,296	2,570	0.029745	437,916	1,492,294	0.293	9,307	0.00129
10-14	10	519,647	734,331	2,024	0.034581	375,013	1,349,913	0.278	7,328	0.00119
15-19	15	413,331	634,658	3,484	0.042884	313,612	1,148,561	0.273	12,616	0.00246
20-24	20	364,837	524,836	5,038	0.036364	249,958	931,517	0.268	18,246	0.00417
25-29	25	281,551	377,773	5,368	0.029398	203,803	742,497	0.274	19,439	0.00596
30-34	30	207,121	327,407	5,130	0.045790	170,956	607,229	0.282	18,579	0.00713
35-39	35	170,467	260,436	4,714	0.042382	131,398	464,507	0.283	17,071	0.00810
40-44	40	139,774	190,152	4,591	0.030780	102,066	360,081	0.283	16,624	0.01020
45-49	45	110,583	143,928	3,853	0.026355	83,256	283,672	0.293	13,954	0.01106
50-54	50	91,039	147,839	4,925	0.048484	69,370	255,722	0.271	17,833	0.01537
55-59	55	60,906	87,023	3,864	0.035684	50,074	178,017	0.281	13,994	0.01922
60-64	60	65,374	84,499	6,212	0.025662	38,357	143,478	0.267	22,494	0.03027
65-69	65	38,928	51,075	4,232	0.027158	27,912	115,568	0.242	15,325	0.03437
70-74	70	30,553	62,691	5,224	0.071876	20,414	98,802	0.207	18,917	0.04322
75+	75	46,842	68,635	7,820	0.038202	9,887	NA	NA	28,318	0.04994
Total		3,827,849	5,329,009	108,569					393,159	
							Median	0.276		
							0.5 * interquartile range	0.007		
							Percentage	2.4		

Source: Population age distribution for 1982 and 1992 from:http://www.census.gov/ipc/www/idbprint.html. See also, for the 1992 census: Central Statistical Office (n.d.) *Census 1992: Zimbabwe National Report*, Harare, Zimbabwe, table A1.2, p.9 and p. 177. For the 1982 census see United Nations (1988) *Demographic Yearbook* , table 7, pp. 252-253. Intercensal deaths by age from: http://demog.berkeley.edu/wilmoth/mortality.

[a] Reference date : 18 August 1982.

[b] Reference date : 18 August 1992.

Procedure

Columns 1-5. Enter input data, cumulated census age distributions and average annual intercensal deaths as shown.

Column 6. Compute the age-specific growth rates using $[ln(P_1(x,5)/P_2(x,5))]/t$, where t is the length of the intercensal period and *ln* denotes natural logarithm.

Column 7. Interpolate the value of e_{75}, the expectation of life at age 75, from vital registration data for 1960, 1965 and 1970 and compute the value of the last entry in column 7, $N(75)$, using the formula

$$N^d(75) = D(75+)\{exp[r(75+)e_{75}] - [(r(75+)e_{75}]^2/6 \quad (23)$$

Then compute the values of $N^d(70)$, $N^d(65)$, ..., from the formula

$$N^d(x\text{-}5) = N^d(x)exp[5r(x,5)] + D(x\text{-}5,5)exp[2.5r(x,5)] \quad (22)$$

where $r(x,5)$ denotes the growth rate for the age interval x to $x+5$.

Column 8. Compute the average number of persons in the x to $x+4$ age group during the intercensal period using the formula

$$N^*(x) = t0.2[P_1(x\text{-}5,5)P_2(x,5)]^{0.5} \quad (19)$$

Column 9. Compute the ratios of the $N^d(x)$ values in column 7 to the $N^*(x)$ values in column 8.

Columns 10 and 11: Compute adjusted deaths by dividing the reported number of intercensal deaths (column 5) by the median $c(x)$ ratio of 0.276. Calculate the adjusted death rate and enter it in column 11.

TABLE III.3. THE EXTINCT GENERATIONS METHOD APPLIED TO ZIMBABWE, FEMALES, 1982-1992: LIFE-TABLE BASED ON REGISTERED DEATHS ADJUSTED FOR UNDER-REGISTRATION

				Conditional life table functions			
Age group	*Adjusted death rate*	*Age* x	*Probability of dying between* age_x *and* $_x$+5 $_5q_x$	*Probability of survival to* age_x l_x/l_5	*Person years lived between* age_x *and* $_x$+5 $_5l_x/l_5$	*Total person years lived above age x* T_x/l_5	*Expectation of life at age x* e_x
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
0-4	NA	0	NA	NA	NA	NA	NA
5- 9	0.00129	5	0.00649	1.0000	4.9838	57.6481	57.6
10-14	0.00119	10	0.00595	0.9935	4.9528	52.6643	53.0
15-19	0.00246	15	0.01239	0.9876	4.9074	47.7115	48.3
20-24	0.00417	20	0.02107	0.9754	4.8255	42.8041	43.9
25-29	0.00596	25	0.03025	0.9548	4.7019	37.9787	39.8
30-34	0.00713	30	0.03632	0.9259	4.5456	33.2768	35.9
35-39	0.00810	35	0.04135	0.8923	4.3693	28.7312	32.2
40-44	0.01020	40	0.05232	0.8554	4.1652	24.3619	28.5
45-49	0.01106	45	0.05688	0.8107	3.9380	20.1968	24.9
50-54	0.01537	50	0.07993	0.7645	3.6700	16.2588	21.3
55-59	0.01922	55	0.10096	0.7034	3.3396	12.5888	17.9
60-64	0.03027	60	0.16371	0.6324	2.9032	9.2492	14.6
65-69	0.03437	65	0.18800	0.5289	2.3958	6.3459	12.0
70-74	0.04322	70	0.24231	0.4295	1.8871	3.9501	9.2
75+	NA	75	1.00000	0.3254	NA	2.0630	6.34

Source: Adjusted death rates from table III.2, column 11.

Procedure

Columns 1-2. Record ages, and age-specific death rates for ages 5-9 and older from column 11 of table III.2.

Columns 3-4. Compute life table $_5q_x$ values for ages 0, 5, 10… using the formula

$$_5q_x = 5\,_5m_x /[1 - 2.5\,_5m_x]$$

Column 5. Compute l_x/l_5 values by noting that $l_5/l_5=1$ and using the formula

$$l_{x+5} = l_x(1-{_5q_x}),$$

Column 6. Compute $_5L_x/l_5$ where :

$$_5L_x/l_5 = 2.5(l_x/l_5 + l_{x+5}/l_5)$$

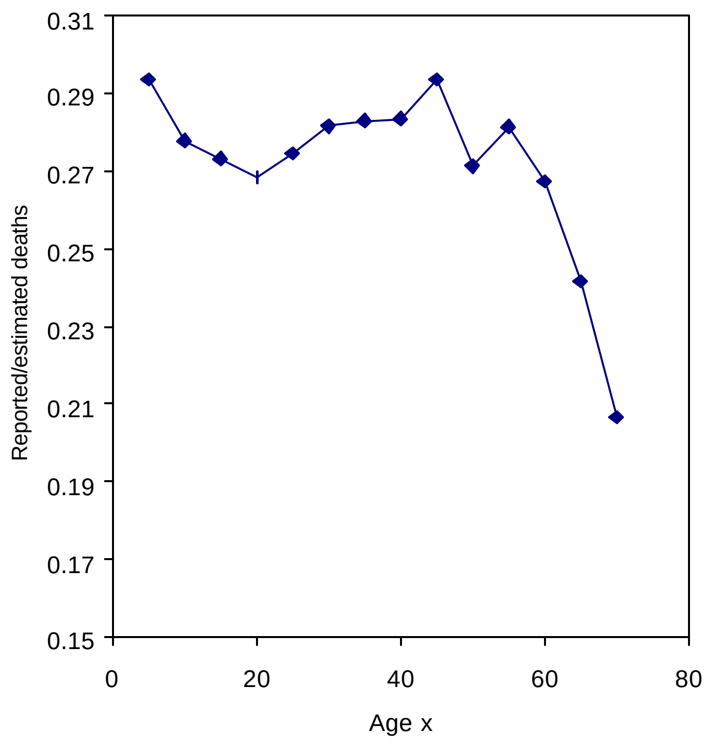
Column 7. Based on a preliminary estimate of e_0 of 57.5 years, put $e_{75} = 6.5$ years. Then compute T_{75}/l_5 as $e_{75}(l_{75}/l_5)$. Now compute T_x/l_5 using the formula

$$T_{x-5}/l_5 = T_x/l_5 + {_5L_x}/l_5,$$

Column 8. Compute e_x for $x = 5, 10, ..., 70$ using the formula

$$e_x = (T_x/l_5)/(l_x/l_5)$$

Figure III.1. The extinct generations method applied to Zimbabwe, females, 1982-1992: plot of estimates of age completeness ratios



Source: Column 9 of table III.2.

IV. ESTIMATES DERIVED FROM INFORMATION ON SURVIVAL OF PARENTS

This chapter and the next one deal with a set of methods that estimate adult mortality using information from a census or survey on the survival of relatives of respondents. This chapter presents methods based on information on the survival status of mothers and fathers. The next chapter presents methods based on information on the survival of brothers and sisters.

The methods discussed in this and the next chapter are different from those considered in earlier chapters in two important respects. First, they do not assume a population closed to migration, and they are therefore applicable to the populations of subnational geographical units, to populations of urban and rural areas, and other populations not closed to migration. This is a strong advantage. However, estimates derived from information on parental survivorship require data that are far less widely available than census age distributions and data on intercensal deaths. This relative scarcity of data is a severe practical disadvantage. However, this disadvantage can be reversed by the inclusion of the necessary questions in future population censuses and surveys.

A. Data required

Parental survivorship methods rely on the simple questions: “Is your mother living?” and “Is your father living?” From such data the proportion of persons in any given age group whose mother or father is surviving can be obtained.

To estimate adult female mortality, the proportions of persons, in five-year age groups, whose mother is surviving and an estimate of the mean age of these mothers at the time of their children’s birth are required. Proportions of persons with mother surviving will usually be calculated from a table showing persons classified in five-year age groups and by the survivorship of their mothers. The mean age of mothers at the time of their children’s birth is most often calculated from data on births in the 12 months preceding the census or survey.

Similarly, to estimate adult male mortality, the proportions of persons in five-year age groups, whose father is surviving, and an estimate of the mean age of these fathers at the time of their children’s conception are required. Conception is the pertinent event for survival of fathers because a father may die between the conception and birth of his child. The proportions of persons with father surviving are calculated from a table showing persons classified by age and by the survivorship status of their fathers. The average age of childbearing for men will usually be obtained by adding an estimate of the average age difference between spouses and the average length of the gestation period to the average age of childbearing for women.

B. Approach and assumptions

Suppose it is known that a group of persons aged x at a particular time t, all had mothers who were aged y when the persons in question were born. The proportion of these mothers who are surviving at time t estimates the life table survival probability l_{y+x}/l_y for the cohort of women born at time $t-(y+x)$. If all women had exactly one surviving child and if there were no data reporting errors, this estimate would be accurate. However, the mortality experience of women who have no surviving children will not be represented at all, and women with more than one surviving child will be over-represented in proportion to the number of their surviving children. The main assumption of the method is that errors incurred in this way will not be very serious.

C. Estimates from maternal survivorship

A plausible approach to estimation would be to use the proportion of persons in a given age group who have surviving mothers to estimate the conditional survival probability l_{M+x}/L_M, where M denotes the mean age of the mothers at the time of the birth of the persons in question, and x denotes the mid-point of the age group. These estimates will not be convenient, however, because M will vary from one application to another.

The approach therefore, is to choose a convenient age, y, near the mean age at childbearing and a convenient age, x, near the mid-point of the age group, and different for each age group. The conditional survival probability, l_{y+x}/l_y, can then be expressed as a linear function of the mean age of mothers and the proportion of persons in the age group with surviving mothers using a regression approach so that:

$$l_{25+x}/l_{25} = a_0(x) + a_1(x)M + a_2(x)S(x-5,5) \qquad (1)$$

where M denotes the mean age of mothers at the birth of their children and $S(x!5,5)$ denotes the proportion of persons aged $x!5$ to x whose mother is surviving. Values for $a_0(x)$, $a_1(x)$, and $a_2(x)$ are obtained by regression on a set of model values of the three variables l_{25+x}/l_{25}, M, and $S(x-5,5)$. The procedure is described in detail in, and the coefficients used here are taken from, Timæus (1992).

D. ESTIMATES FROM PATERNAL SURVIVORSHIP

Estimation of adult male mortality from data on paternal survivorship proceeds in much the same way. Survival probabilities are conditional on reaching age 35, rather than age 25 (because husbands tend to be older than their wives) and proportions with father surviving are taken from two successive age groups rather than a single age group. The equation is

$$l_{35+x}/l_{35} = a_0(x) + a_1(x)M +$$

$$+ a_2(x)S(x-5,5) + a_3(x)S(x,5), \qquad (2)$$

where M denotes the mean age of the fathers at the birth of their children, $S(x!5,5)$ denotes the proportion of persons aged $x!5$ to x whose father is surviving, and $S(x,5)$ the proportion aged x to $x+5$ whose father is surviving. For further discussion see Timaeus (1992).

E. MATERNAL SURVIVORSHIP METHOD APPLIED TO ZIMBABWE, 1992 CENSUS

Table IV.1 illustrates the estimation of adult female mortality using maternal survivorship data from the 1992 census of Zimbabwe. The application incorporates three main elements. The first element, discussed in this section, is the derivation of estimated survivorship probabilities $l_{25+x}/l_{25,}$ using formula (1) of the preceding section. The second two elements, discussed in the following two sections, are model life table translations of the estimated survivorship probabilities and the derivation of time locations or reference dates for these estimates.

The numbers of children reported born to women in the 12 months preceding the 1992 census are shown in table IV.2. The age group labels here refer to age at the time of the census. To calculate the mean age of mothers at the birth of their children, however, it is appropriate to use an age estimate which is one half year less than their age at the time of the census. The age intervals, 15-19, 20-24, 25-29 … 45-49, can therefore be adjusted to become 14.5-19.4, 19.5-24.4, 25.5-29.5, and so on, with midpoints of 17, 22, 27… and 47. The mean age of mothers at the birth of their children would thus be:

17H0.1432 + 22H0.3167 + ... + 47H0.0126

or 26.7 years.

F. MODEL LIFE TABLE FAMILY TRANSLATION

Although the estimated survivorship probabilities shown in column 9 of table IV.1 are the final result of the orphanhood method as originally developed, it is useful to use a model life table family to convert the conditional survivorship estimates to a common statistic. There is no hard and fast rule for what the common statistic should be, and it might be varied from one application to another. A useful default is ${}_{35}q_{30}$, that is, the conditional probability of dying by age 65 given survival to age 30. This corresponds reasonably closely to the age range of the estimates yielded by the estimates from both maternal and paternal survivorship methods. Table IV.1 shows the ${}_{35}q_{30}$ values in column 10. The translation procedure, described in annex II, is facilitated by the table of model life table values shown in table IV.3.

If the data were perfectly accurate and the assumptions of the method perfectly valid, and if mortality levels had not changed during the period in question, and if the true age pattern of mortality in the population conformed to the model life table used, the ${}_{35}q_{30}$ values in column 10 of table IV.1 would be the same for each age. The observed variation in the values is modest, ranging from 0.1735 to 0.2195, suggesting that the data are reasonably accurate though not perfect.

G. TIME LOCATION OF THE ESTIMATES

The survival probabilities in column 9 of table IV.1 refer to different time periods. For persons aged 5-9 years the interval over which the mothers survived begins 5-10 years prior to the census, but for persons aged 45-49 years it begins 45-50 years prior to the census. The estimate of l_{35}/l_{25} from the 5-9 age group therefore represents an average of mortality risks during the 10 years prior to the census, whereas the estimate of l_{75}/l_{25} from the 45-49 age group represents an average of mortality risks over the 50 years prior to the census.

These differences in the reference period of the estimates mean that the proportions of persons whose mothers are alive contain information about the trend, as well as the level of mortality. If mortality has declined substantially over the half century preceding the census or survey, the estimate of l_{75}/l_{25} from the 45-49 age group will represent a higher average level of mortality than the estimate of l_{35}/l_{25} from the 5-9 age group. Without the model life table translation to a common statistic, there is no way of exploiting this trend information. With the translation of both l_{35}/l_{25} and l_{75}/l_{25} to ${}_{35}q_{30}$, however, a change in mortality level may be revealed.

The trend information inherent in the data may be exploited by deriving the relation between the cohort survivorship statistics l_{25+x}/l_{25} shown in column 9 of table IV.1 and the corresponding period statistics at various times prior to the census or survey. If mortality risks have been declining in prior decades, the conditional survival probability l_{25+x}/l_{25} in the period life tables for each time in the past will have been declining. It is intuitively clear that the cohort survival probability over any given time period will lie somewhere between the high period values of the more distant past and the low period values of the recent past. It is plausible, therefore, that there is some time *t,* prior to the census or survey, such that the cohort survivorship estimates in column 9 equals the corresponding period survivorship at time *t*. This point in time is referred to as the "time location" of the estimate. If mortality risks have changed approximately linearly, it is possible to estimate this time location reasonably accurately. The theory on which the time location calculation is based, presented in Brass and Bamgboye (1981), is beyond the scope of this manual, but it is useful to present a brief, heuristic explanation.

If the life table survivorship function l_x is linear over the relevant portion of the age span, the deaths of the mothers of persons aged *N* at the time of the census or survey will be uniformly distributed over the preceding *N* years. It can then be demonstrated that the time location for the corresponding survivorship is the mid-point of this period, *N/2* years prior to census or survey. The survivorship function l_x is indeed approximately linear if mortality levels are high and *x* (age) is not too high. For lower levels of mortality and at older ages, however, there is a sharp downward curvature of the survivorship function. This implies that deaths of mothers during the years prior to the census or survey are disproportionately concentrated in the later portion of the period resulting in a time location estimate that is closer to the survey date than *N/2*.

The time location can therefore be written as

$$T(N) = (N/2)(1 - C(N)), \qquad (3)$$

where *T(N)* is the time location of the estimate for the age group with the midpoint N, and *C(N)* is a correction factor for that age group. Brass and Bamgboye (1981) showed that this correction factor may be calculated as

$$C(N) = ln(S(N))/3 + f(N+M) + \\ + 0.0037(27-M) \qquad (4)$$

where *S(N)* denotes the proportion of persons aged *N* whose mothers are surviving, *M* denotes the mean age of these mothers at the time the persons in question were born and *f(N+M)* is standard function of age whose value can be interpolated between the values given in table IV.4.

The differences between the estimation equations for maternal and paternal survivorship imply slight differences in the application of formula (4). For survivorship of mothers, the conditional survivorship l_{25+x}/l_{25} is estimated from the proportion of persons aged *x*!*5* to *x* whose mother is surviving, *S(x*!*5,5),* therefore *S(N)* in (4) is taken to be *S(x*!*5,5)* and *N* is taken to be the midpoint of the age group, *x-2.5*. The *M* in (1) is the mean age of the mothers of the respondents at the time the respondents were born, *i.e., N* years ago.

For survivorship of fathers, however, l_{35+x}/l_{35} is estimated from the proportions of persons in the age groups *x*!*5* to *x* and *x* to *x+5* whose father is surviving, i.e. *S(x-5,5)* and *S(x,5)*, respectively. In this case *S(N)* is taken as the average of the proportions with fathers surviving in the two age groups and *N* is taken as the mid-point of the two age groups plus the gestation period, *x+0.75*. The *M* in (2) is the average age of the fathers of the respondents at the time of the respondents' birth. For the purpose of equation (3), however, *M* must be taken as the mean age of the fathers of the respondents at the time of the respondents' conception. The average age of fathers at birth can be denoted by M_1 and the average age of fathers at conception denoted by $M_2 = M_1 {}^{-} 0.75$.

H. Time location analysis for maternal survival: Zimbabwe, 1992 census

Figure IV.1 plots the translated ${}_{35}q_{30}$ values against their estimated time locations. In the best of circumstances it is possible to estimate mortality trends by the application of this procedure. In some applications, however, errors in the data and/or departures of actual from assumed conditions overwhelm the trend indication. The conclusions drawn may then refer, not only to mortality trends, but to errors in the reported proportions of surviving mothers or fathers and/or the invalidity of the assumptions.

In the case shown in figure IV.1 it is immediately apparent that the trend indications are somewhat unexpected. It is highly unlikely that adult female mortality risks in Zimbabwe rose in the early 1980s. The subsequent decline is plausible, however, as is the slight increase in mortality risks in the late 1980s, a trend which might be due to an increase in deaths due to the human immunodeficiency virus (HIV) and acquired immunodeficiency syndrome (AIDS). It should be noted, however, that inherent limitations in trend analysis, as discussed below, make the attribution of these trends to any specific factor rather tenuous.

I. Inherent limitations of trend analysis

The estimation of adult mortality from data on survival of parents allows the estimation of long term trends in the level of mortality, but not of short term changes. "Long term" here means roughly 10-50 years, and "short term" less than 10 years.

Short-term fluctuations in these estimates, especially sharp movements over a few years, necessarily represent errors in reporting. This is because the conditional survival probabilities estimated from different age groups of respondents average period mortality experience over relatively long periods of time, roughly 10-50 years. It follows that they cannot contain any information on short term fluctuations in the level of mortality. This can be illustrated by considering a hypothetical example in which mortality fluctuates between arbitrarily chosen high and low values from one year to the next. The average level of mortality to which the mothers of persons in any age group were subject will be essentially the same as if mortality were constant. Year to year fluctuations are lost in the proportions of surviving parents because every parent is exposed to high and low levels for approximately equal periods. The same would be true if mortality alternated between high and low levels every two or three years. The putative trends indicated by time location are valid only on the assumption that the level of mortality has been declining reasonably smoothly over a long period of time.

Sharp fluctuations in level such as those shown in figure IV.1 probably represent differences in reporting errors between age groups, not changes in level of mortality. The practical lesson here is that interpretation of the plot is not simply a matter of reading the putative trend, but of deciding which features reflect changes in mortality and which reflect problems with the data or the method.

Inaccurate reporting of parental survivorship status is an important source of erroneous trends in the data. Reports of parental survivorship for children will often be given by the head of household or another adult in the household in which the child is enumerated. In some countries a significant proportion of these adults will be adoptive parents who may report the child's parent as surviving if the child's adoptive parent is living. This will bias the reported number and proportion of surviving parents upwards.

As persons become older, the chance that their adoptive as well as their biological parent is dead will increase. For persons whose biological and adoptive mothers are both dead, for example, the report on survivorship of mother will be correct even if the respondent is reporting on the adoptive rather than the biological mother. This implies that the "adoption bias" is likely to be most pronounced for younger children whose adoptive parent is more likely to be alive, and to decline with older persons who are more likely to have lost biological and adoptive parents.

Adoption bias is likely to result in lower expectations of life from older age groups relative to younger ones, and may suggest an increase in the expectation of life in the years preceding the census or survey. This phenomenon might explain some or all of the apparent decline in survival probabilities indicated in figure IV.1.

J. Paternal survivorship method applied to Zimbabwe, 1992 census

Table IV.5 shows the estimation of adult male mortality from paternal survivorship data. As is often the case, the calculation of M for males is problematic. The 1992 census marital status tabulations show the mean age of married men to be 42.5 years and the mean age of married women to be 35.3 years, for a difference of 7.2 years. If medians rather than means are used, the figures are, respectively, 37.0, 30.1, and 6.9 years. Other pertinent data are not readily available. In the event, assume a sex difference of 7 years. Adding this to the M = 26.7 years calculated for females in section E gives $M_1 = 33.7$ years for males.

The estimation equations for the survival probabilities are different for males, as already noted, and there are slight differences in the time location calculation, but otherwise the procedure is the same as for maternal survival.

The ${}_{35}q_{30}$ values in column 11 are obtained by interpolation in table IV.6, which has the same format as table IV.3 except for values being conditional on survival to age 35 rather than to age 25. The median of these ${}_{35}q_{30}$ values is 0.331, compared with 0.206 for females (table IV.1), suggesting a much higher level of male adult than female adult mortality.

K. Two-census method

Estimates of adult mortality based on information of parental survivorship can also be derived from data on two censuses or surveys. If data are available for two censuses or surveys five or ten years apart, the synthetic cohort procedure proposed by Zlotnik and Hill (1981) may be applied to obtain an estimate that refers to the intervening period. Let $S_1(x,5)$ denote the proportion of persons aged x to $x+5$ whose mother is surviving at the first point in time, $x = 5, 10, ...,$ and let $S_2(x,5)$ denote the same statistic for the second point in time. In this section it is assumed that the time interval is exactly five years. Assuming a time interval between the censuses or surveys to be exactly five years, the proportions of persons with mother surviving for an hypothetical cohort can be constructed based on changes in proportions with mother surviving between the two censuses. The proportion aged 5-9 with mother surviving in this hypothetical cohort, for example, will be the average of $S_1(5,5)$ and $S_2(5,5)$,

$$S^*(5,5) = [S_1(5,5) + S_2(5,5)]/2. \qquad (5)$$

The proportion with mother surviving in subsequent age groups is

$$S^*(x,5) = [S_2(x,5)/S_1(x\text{-}5,5)]S^*(x\text{-}5,5), \qquad (6)$$

$x = 10, 15, ...$ The ratios $S_2(x,5)/S_1(x\text{-}5,5)$ are analogous to census survival ratios. They represent the change in proportion with mother surviving in the actual cohort aged x to $x+5$ at the first census, and therefore reflect mortality conditions during the intercensal period. The estimation procedure described in preceding sections is applied to the $S^*(x,5)$ values calculated from formulas (5) and (6) exactly as if they were proportions with mothers surviving from a single census or survey.

When the interval between the surveys or censuses is other than five years, an adaptation of the intercensal survival method (chapter I, section D) may be used. In place of the ratios used in (6) above, it is necessary to compute the synthetic ratios

$$R(x,5) = \frac{S(x+5,5)exp[2.5r(x+5,5}{S(x,5)exp[-2.5r(x,5)]} \qquad (7)$$

where

$$S(x,5) = [S_1(x,5) + S_2(x,5)]/2 \qquad (8)$$

and

$$r(x,5) = ln[S_2(x,5)/S_1(x,5)]/t \qquad (9)$$

where t is the length of the intercensal interval. The proportions with surviving mothers for the hypothetical cohort are then calculated using (5) and

$$S^*(x,5) = R(x-5,5)S^*(x-5,5), \qquad (10)$$

$x = 10,15,$ .... Table IV.7 shows the application of the two census method to maternal survival data from the 1982 and 1992 censuses of Zimbabwe. The median probability of dying between 30 and 65 years for females is 0.192, compared with 0.206 obtained from the single census method results shown in table IV.1.

L. Other parental survival methods

Variants of parental survival methods have been developed for use with data on survival of parents at times other than the time of the census or survey. Questions on parental survival may be supplemented by obtaining information on the date of death for persons who report their parent to be deceased. Alternatively, respondents may be asked whether their mother or father was surviving at the time of some particular past event, such as the respondent's 20th birthday or the respondent's marriage. Data of this kind are more likely to be available from surveys than from censuses, but surveys on which suitable questions may be included are very common.

Timæus (1991a) presents a method using the proportion of mothers (fathers) deceased among those respondents whose mother (father) survived to the time the respondent reached age 20. Timæus (1991b) presents a method using the proportions of mothers (fathers) surviving among those respondents whose mother (father) was surviving at the time of the respondent's first marriage. These methods should be applied whenever the requisite data are available.

TABLE IV.1. ESTIMATION OF ADULT FEMALE MORTALITY FROM SURVIVAL OF MOTHERS: ZIMBABWE, 1992 CENSUS

Age group	*Respondents with mother alive*	*Respondents with mother dead*	*Proportion with mothers alive*	*Regression coefficients*				*Estimated conditional survival probability*	*Adult probability of death*	*Parameters for estimating time location of deaths*					
										Age	*Adjusted age*	*Standard function*	*Estimated correction*	*Years back*	*Time*
			S(x-5,5)	a_0	a_1	a_2	*x*	l_{25+x}/l_{25}	${}_{35}q_{30}$	*N*	*N+M*	*f(N+M)*	*C(N)*	*(T(N))*	
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*	*(12)*	*(13)*	*(14)*	*(15)*	*(16)*
5-9	1,627,756	25,192	0.9848	-0.2894	0.00125	1.2559	10	0.9808	0.1763	7.5	34.2	0.090	0.0860	3.4	1989.2
10-14	1,416,594	39,168	0.9731	-0.1718	0.00222	1.1123	15	0.9699	0.1735	12.5	39.2	0.100	0.0920	5.7	1987.0
15-19	1,186,594	60,472	0.9515	-0.1513	0.00372	1.0525	20	0.9495	0.1946	17.5	44.2	0.132	0.1165	7.7	1984.9
20-24	909,322	79,161	0.9199	-0.1808	0.00586	1.0267	25	0.9201	0.2145	22.5	49.2	0.184	0.1573	9.5	1983.1
25-29	626,307	84,615	0.8810	-0.2511	0.00885	1.0219	30	0.8855	0.2195	27.5	54.2	0.248	0.2069	10.9	1981.7
30-34	503,140	101,979	0.8315	-0.3644	0.01287	1.0380	35	0.8423	0.2158	32.5	59.2	0.324	0.2636	12.0	1980.7
35-39	373,521	114,172	0.7659	-0.5181	0.01795	1.0753	40	0.7847	0.2064	37.5	64.2	0.415	0.3272	12.6	1980.0
40-44	240,189	122,345	0.6625	-0.6880	0.02342	1.1276	45	0.6843	0.2077	42.5	69.2	0.523	0.3869	13.0	1979.6
45-49	160,838	126,634	0.5595	-0.8054	0.02721	1.1678	50	0.5745	0.1844	47.5	74.2	0.656	0.4635	12.7	1979.9
50-54	119,261	159,565	0.4277	NA	NA	NA	55	NA	NA	NA	NA	NA	NA	NA	NA
								Median	0.2064						
								0.5 * interquartile range	0.0151						
								Percentage	7.3						

Source: Numbers of persons with "mother alive" and "mother deceased" from unpublished tables of the 1992 census of Zimbabwe. Coefficients in columns 5-7 from: Timaeus, Ian M. (1992) Estimation of Adult Mortality from Paternal Orphanhood: A Reassessment and a New Approach, *Population Bulletin of the United Nations*, No. 33. pp. 47-63 and table 2, page 56.

Procedures

Columns 1-3. Record the distribution of respondents in five year age groups, by whether their mothers are alive or dead.

Column 4. Compute the proportion of persons with mother living and enter that in column 4. Persons for whom mother's survival status is not given should be excluded.

Columns 5-9. Using the coefficients in columns 5-7 and the formula

$$l_{25+x}/l_{25} = a_0(x) + a_1(x)M + a_2(x)S(x\text{-}5,5), \qquad (1)$$

calculate the estimated conditional survival probabilities l_{25+x}/l_{25}. Enter the value of x in column 8 and the survivorship probability in column 9. The value of M is calculated from data on births during the 12 months prior to the census. See text for further discussion.

Column 10. Calculate the conditional probability of dying before age 65 years, given survival to age 30 years, corresponding to the given value of l_{25+x}/l_{25} by interpolation in table IV.3.

Columns 11-13. Record the midpoint of the age group, denoted by N, the value of $N+M$, and interpolate in table IV.4 to obtain the value of $f(N+M)$.

Column 14. Compute the value of $C(N)$ using the formula

$$C(N) = ln(S(x\text{-}5,5))/3 + f(N+M) + 0.0037(27\text{-}M) \qquad (4)$$

Columns 15-16. Compute the "years back" value using the formula

$$T(N) = (N/2)(1 - C(N)), \qquad (3)$$

and the date to which the estimate pertains "time" by subtracting this value from the date of the census or survey.

TABLE IV.2. NUMBER AND PROPORTION OF CHILDREN BORN IN THE LAST TWELVE MONTHS, BY AGE GROUP OF MOTHER, ZIMBABWE, 1992

Age group of mother	*Number of births*	*Proportion born in last 12 months*
15-19	51,532	0.1432
20-24	113,965	0.3167
25-29	77,393	0.2150
30-34	58,693	0.1631
35-39	37,559	0.1044
40-44	15,224	0.0423
45-49	4,520	0.0126
TOTAL	359,886	1.0000

Source: Zimbabwe, Central Statistical Office (1994). *Census 1992: Zimbabwe National Report*, Appendix table A8.1, page 200, Harare.

TABLE IV.3. ESTIMATION OF ADULT FEMALE MORTALITY FROM SURVIVAL OF MOTHERS:
CONDITIONAL SURVIVAL PROBABILITIES l_{25+x}/l_{25} FOR MODEL LIFE TABLE TRANSLATION

Column	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*Column*
l(35)/l(25)	0.8304	0.8391	0.8479	0.8567	0.8655	0.8743	0.8830	0.8917	0.9001	l(35)/l(25)
l(40)/l(25)	0.7510	0.7627	0.7745	0.7865	0.7987	0.8108	0.8231	0.8353	0.8475	l(40)/l(25)
l(45)/l(25)	0.6703	0.6842	0.6984	0.7130	0.7278	0.7429	0.7582	0.7737	0.7893	l(45)/l(25)
l(50)/l(25)	0.5848	0.6001	0.6160	0.6325	0.6494	0.6669	0.6848	0.7031	0.7218	l(50)/l(25)
l(55)/l(25)	0.4935	0.5094	0.5260	0.5434	0.5617	0.5807	0.6005	0.6210	0.6422	l(55)/l(25)
l(60)/l(25)	0.3968	0.4120	0.4283	0.4455	0.4638	0.4831	0.5036	0.5252	0.5478	l(60)/l(25)
l(65)/l(25)	0.2977	0.3112	0.3256	0.3412	0.3579	0.3760	0.3954	0.4162	0.4385	l(65)/l(25)
l(70)/l(25)	0.2026	0.2130	0.2244	0.2368	0.2504	0.2653	0.2815	0.2993	0.3188	l(70)/l(25)
l(75)/l(25)	0.1197	0.1266	0.1341	0.1425	0.1517	0.1620	0.1734	0.1861	0.2003	l(75)/l(25)
e_0	20.00	22.50	25.00	27.50	30.00	32.50	35.00	37.50	40.00	e_0
e_5	40.07	41.12	42.20	43.31	44.45	45.62	46.83	48.08	49.36	e_5
${}_{35}q_{15}$	0.5048	0.4862	0.4669	0.4469	0.4263	0.4051	0.3833	0.3610	0.3383	${}_{35}q_{15}$
${}_{35}q_{30}$	0.6733	0.6604	0.6466	0.6316	0.6156	0.5983	0.5797	0.5597	0.5383	${}_{35}q_{30}$

Column	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*Column*
l(35)/l(25)	0.9001	0.9085	0.9166	0.9246	0.9323	0.9397	0.9467	0.9535	0.9598	l(35)/l(25)
l(40)/l(25)	0.8475	0.8595	0.8714	0.8831	0.8945	0.9056	0.9164	0.9266	0.9364	l(40)/l(25)
l(45)/l(25)	0.7893	0.8049	0.8204	0.8359	0.8511	0.8661	0.8808	0.8949	0.9085	l(45)/l(25)
l(50)/l(25)	0.7218	0.7408	0.7599	0.7792	0.7984	0.8176	0.8365	0.8551	0.8731	l(50)/l(25)
l(55)/l(25)	0.6422	0.6641	0.6865	0.7094	0.7326	0.7561	0.7797	0.8032	0.8263	l(55)/l(25)
l(60)/l(25)	0.5478	0.5716	0.5964	0.6223	0.6491	0.6767	0.7049	0.7336	0.7625	l(60)/l(25)
l(65)/l(25)	0.4385	0.4624	0.4879	0.5150	0.5438	0.5743	0.6063	0.6397	0.6743	l(65)/l(25)
l(70)/l(25)	0.3188	0.3401	0.3634	0.3889	0.4167	0.4470	0.4799	0.5155	0.5536	l(70)/l(25)
l(75)/l(25)	0.2003	0.2162	0.2341	0.2541	0.2766	0.3020	0.3307	0.3629	0.3990	l(75)/l(25)
e_0	40.00	42.50	45.00	47.50	50.00	52.50	55.00	57.50	60.00	e_0
e_5	49.36	50.68	52.03	53.43	54.86	56.34	57.86	59.43	61.05	e_5
${}_{35}q_{15}$	0.3383	0.3152	0.2919	0.2685	0.2451	0.2218	0.1988	0.1763	0.1544	${}_{35}q_{15}$
${}_{35}q_{30}$	0.5383	0.5154	0.4910	0.4650	0.4374	0.4082	0.3775	0.3455	0.3123	${}_{35}q_{30}$

Column	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*Column*
l(35)/l(25)	0.9598	0.9658	0.9713	0.9764	0.9809	0.9850	0.9886	0.9917	0.9942	l(35)/l(25)
l(40)/l(25)	0.9364	0.9456	0.9542	0.9622	0.9695	0.9760	0.9817	0.9866	0.9907	l(40)/l(25)
l(45)/l(25)	0.9085	0.9214	0.9336	0.9449	0.9554	0.9648	0.9731	0.9802	0.9862	l(45)/l(25)
l(50)/l(25)	0.8731	0.8904	0.9069	0.9224	0.9368	0.9499	0.9616	0.9717	0.9802	l(50)/l(25)
l(55)/l(25)	0.8263	0.8489	0.8708	0.8916	0.9111	0.9292	0.9454	0.9596	0.9717	l(55)/l(25)
l(60)/l(25)	0.7625	0.7914	0.8198	0.8474	0.8738	0.8985	0.9212	0.9414	0.9586	l(60)/l(25)
l(65)/l(25)	0.6743	0.7097	0.7457	0.7816	0.8169	0.8509	0.8829	0.9119	0.9372	l(65)/l(25)
l(70)/l(25)	0.5536	0.5943	0.6373	0.6820	0.7278	0.7738	0.8187	0.8611	0.8995	l(70)/l(25)
l(75)/l(25)	0.3990	0.4396	0.4847	0.5345	0.5887	0.6468	0.7074	0.7685	0.8273	l(75)/l(25)
e_0	60.00	62.50	65.00	67.50	70.00	72.50	75.00	77.50	80.00	e_0
e_5	61.05	62.71	64.42	66.19	68.02	69.91	71.87	73.89	75.99	e_5
${}_{35}q_{15}$	0.1544	0.1333	0.1132	0.0943	0.0768	0.0609	0.0467	0.0344	0.0240	${}_{35}q_{15}$
${}_{35}q_{30}$	0.3123	0.2783	0.2438	0.2094	0.1755	0.1429	0.1123	0.0845	0.0602	${}_{35}q_{30}$

NOTE: Calculated using the Brass General model life table family with parameter β=1. See Annex II for details.

TABLE IV.4. INTERPOLATION TABLE FOR ESTIMATING TIME LOCATION OF ESTIMATES DERIVED FROM INFORMATION ON SURVIVAL OF MOTHERS AND FATHERS

Adjusted age	*Standard function*	*Adjusted age*	*Standard function*	*Adjusted age*	*Standard function*	*Adjusted age*	*Standard function*
N+M	*f(N+M)*	*N+M*	*f(N+M)*	*N+M*	*f(N+M)*	*N+M*	*f(N+M)*
26	0.090	39	0.099	52	0.218	65	0.431
27	0.090	40	0.104	53	0.231	66	0.452
28	0.090	41	0.109	54	0.245	67	0.473
29	0.090	42	0.115	55	0.259	68	0.495
30	0.090	43	0.122	56	0.274	69	0.518
31	0.090	44	0.130	57	0.289	70	0.542
32	0.090	45	0.139	58	0.305	71	0.568
33	0.090	46	0.149	59	0.321	72	0.595
34	0.090	47	0.160	60	0.338	73	0.622
35	0.091	48	0.171	61	0.356	74	0.650
36	0.092	49	0.182	62	0.374	75	0.678
37	0.093	50	0.193	63	0.392		
38	0.095	51	0.205	64	0.411		

Source: Brass W. and E. A. Bamgboye (1981). *The time location of reports of survivorship estimates for maternal and paternal orphanhood and the ever-widowed.* Working Paper No. 81-11. London School of Hygiene and Tropical Medicine, Center for Population Studies, annex p. 12. Reproduced in United Nations (1983). Manual X: Indirect Techniques for Demographic Estimation (United Nations publication, Sales No. E.83.XIII.2), table 88, p.104.

TABLE IV.5. ESTIMATION OF ADULT MALE MORTALITY FROM SURVIVAL OF FATHERS: ZIMBABWE, 1992 CENSUS

				Regression coefficients							Parameters for estimating time location of deaths					
Age group	*Respondents with father alive*	*Respondents with father dead*	*Proportion with fathers alive*						*Estimated conditional survival probability*	*Adult probability of death*	*Age*	*Adjusted age*	*Standard function*	*Estimated correction*	*Years back*	*Time*
			$S(x\text{-}5,5)$	a_0	a_1	a_2	a_3	x	l_{x+35}/l_{35}	${}_{35}q_{30}$	N	$N+M$	$f(N+M)$	$C(N)$	$(T(N))$	
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*	*(12)*	*(13)*	*(14)*	*(15)*	*(16)*	*(17)*
5-9	1,566,032	77,879	0.9526	-0.8251	0.00261	2.7269	-0.9953	10	0.9476	0.3074	10.75	44.45	0.134	0.093	4.9	1987.8
10-14	1,326,940	119,763	0.9172	-0.4013	0.00576	1.5602	-0.3522	15	0.9220	0.2780	15.75	49.45	0.187	0.133	6.8	1985.8
15-19	1,064,392	177,806	0.8569	-0.3329	0.01031	0.6656	0.3419	20	0.8522	0.3279	20.75	54.45	0.251	0.175	8.6	1984.1
20-24	771,021	215,161	0.7818	-0.4726	0.01559	0.2161	0.7896	25	0.7728	0.3410	25.75	59.45	0.329	0.222	10.0	1982.6
25-29	495,263	214,363	0.6979	-0.7056	0.02076	0.1997	0.9066	30	0.6753	0.3408	30.75	64.45	0.420	0.275	11.1	1981.5
30-34	361,210	243,084	0.5977	-0.9153	0.02493	0.3484	0.8631	35	0.5521	0.3349	35.75	69.45	0.529	0.332	11.9	1980.7
35-39	236,524	250,619	0.4855	-0.9950	0.02635	0.4269	0.8263	40	0.3960	0.3309	40.75	74.45	0.663	0.397	12.3	1980.3
40-44	129,639	232,622	0.3579	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
									Median	0.3309					10.0	1982.6
									0.5 * interquartile range	0.0101						
									Percentage	3.1						

Source: Numbers of persons with "father alive" and "father deceased" from unpublished tables of the 1992 census of Zimbabwe. Coefficients in columns 5-8 from: Timaeus, Ian M. (1992) Estimation of Adult Mortality from Paternal Orphanhood: A Reassessment and a New Approach, *Population Bulletin of the United Nations*, No. 33. pp. 47-63 and table 2, page 56.

Procedure

Columns 1-3. Record the distribution of respondents in five year age groups, by whether their fathers are alive or dead.

Column 4. Compute the proportions of person with father living. Persons for whom father's survival status is not given should be excluded.

Columns 5-10. Using the coefficients in columns 5-8 and the formula

$$l_{35+x}/l_{35} = a_0(x) + a_1(x)M + a_2(x)S(x\text{-}5,5) + a_3(x)S(x,5), \qquad (2)$$

calculate the estimated conditional survival probabilities l_{35+x}/l_{35}. Enter the value of x in column 9 and the survivorship probability in column 10. The value of M for males may be taken as the value for females (Table IV.1) plus an estimate of the sex difference (7 years in this case), plus the length of the gestation period, 0.75 years. See text for further explanation.

Column 11. Calculate the conditional probability of dying by age 65 years, given survival to age 30 years corresponding to the given value of l_{35+x}/l_{35} by interpolation in table IV.6.

Column 12. Record the value of the exposure period N. Because the paternal survivorship estimates utilize information from two successive age groups, the midpoint of the age group is the age dividing these two age groups. Also, because fathers are alive at the conception of their children, but not necessarily at their birth, the period of exposure exceeds that for maternal survivorship by 0.75 year.

Columns 13-14. Record the value of $N+M$ and $f(N+M)$ as interpolated from Table IV.4.

Column 15. Compute the value of C(N) using the formula

$$C(N) = ln(S(x\text{-}5,5))/3 + f(N+M) + 0.0037(27\text{-}M) \qquad (4)$$

Columns 16-17. Compute the years back value using the formula

$$T(N) = (N/2)(1 - C(N)), \qquad (3)$$

and the date to which the estimate pertains, "time", by subtracting this value from the date of the census or survey.

TABLE IV.6. ESTIMATION OF ADULT MALE MORTALITY FROM SURVIVAL OF FATHERS: CONDITIONAL SURVIVAL PROBABILITIES l_{35+x}/l_{35} FOR MODEL LIFE TABLE TRANSLATION

e_0	*20.0*	*22.5*	*25.0*	*27.5*	*30.0*	*32.5*	*35.0*	*37.5*	*40.0*	*e_0*
Column	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*Column*
l(45)/l(35)	0.8073	0.8154	0.8237	0.8322	0.8409	0.8497	0.8587	0.8677	0.8768	l(45)/l(35)
l(50)/l(35)	0.7043	0.7152	0.7265	0.7383	0.7503	0.7628	0.7755	0.7886	0.8019	l(50)/l(35)
l(55)/l(35)	0.5944	0.6071	0.6204	0.6343	0.6489	0.6642	0.6800	0.6964	0.7134	l(55)/l(35)
l(60)/l(35)	0.4778	0.4911	0.5051	0.5200	0.5358	0.5526	0.5703	0.5890	0.6086	l(60)/l(35)
l(65)/l(35)	0.3586	0.3709	0.3840	0.3983	0.4136	0.4300	0.4478	0.4668	0.4871	l(65)/l(35)
l(70)/l(35)	0.2440	0.2539	0.2647	0.2765	0.2893	0.3034	0.3188	0.3357	0.3541	l(70)/l(35)
l(75)/l(35)	0.1442	0.1508	0.1582	0.1663	0.1753	0.1853	0.1964	0.2087	0.2226	l(75)/l(35)
e_5	40.07	41.12	42.20	43.31	44.45	45.62	46.83	48.08	49.36	e_5
${}_{35}q_{15}$	0.5048	0.4862	0.4669	0.4469	0.4263	0.4051	0.3833	0.3610	0.3383	${}_{35}q_{15}$
${}_{35}q_{30}$	0.6733	0.6604	0.6466	0.6316	0.6156	0.5983	0.5797	0.5597	0.5383	${}_{35}q_{30}$

e_0	*40.0*	*42.5*	*45.0*	*47.5*	*50.0*	*52.5*	*55.0*	*57.5*	*60.0*	*e_0*
Column	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*Column*
l(45)/l(35)	0.8768	0.8859	0.8950	0.9041	0.9130	0.9218	0.9303	0.9386	0.9465	l(45)/l(35)
l(50)/l(35)	0.8019	0.8154	0.8290	0.8427	0.8565	0.8701	0.8836	0.8968	0.9096	l(50)/l(35)
l(55)/l(35)	0.7134	0.7309	0.7489	0.7672	0.7859	0.8047	0.8236	0.8424	0.8609	l(55)/l(35)
l(60)/l(35)	0.6086	0.6292	0.6507	0.6731	0.6962	0.7201	0.7446	0.7694	0.7945	l(60)/l(35)
l(65)/l(35)	0.4871	0.5089	0.5322	0.5570	0.5834	0.6112	0.6404	0.6709	0.7025	l(65)/l(35)
l(70)/l(35)	0.3541	0.3743	0.3965	0.4206	0.4470	0.4757	0.5069	0.5406	0.5768	l(70)/l(35)
l(75)/l(35)	0.2226	0.2380	0.2554	0.2748	0.2968	0.3214	0.3493	0.3806	0.4157	l(75)/l(35)
e_5	49.36	50.68	52.03	53.43	54.86	56.34	57.86	59.43	61.05	e_5
${}_{35}q_{15}$	0.3383	0.3152	0.2919	0.2685	0.2451	0.2218	0.1988	0.1763	0.1544	${}_{35}q_{15}$
${}_{35}q_{30}$	0.5383	0.5154	0.4910	0.4650	0.4374	0.4082	0.3775	0.3455	0.3123	${}_{35}q_{30}$

e_0	*60*	*62.5*	*65*	*67.5*	*70*	*72.5*	*75*	*77.5*	*80*	*e_0*
Column	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*Column*
l(45)/l(35)	0.9465	0.9541	0.9612	0.9678	0.9739	0.9794	0.9843	0.9885	0.9920	l(45)/l(35)
l(50)/l(35)	0.9096	0.9219	0.9337	0.9448	0.9550	0.9643	0.9726	0.9799	0.9859	l(50)/l(35)
l(55)/l(35)	0.8609	0.8790	0.8965	0.9132	0.9288	0.9433	0.9563	0.9677	0.9773	l(55)/l(35)
l(60)/l(35)	0.7945	0.8194	0.8440	0.8679	0.8907	0.9122	0.9318	0.9492	0.9642	l(60)/l(35)
l(65)/l(35)	0.7025	0.7349	0.7677	0.8005	0.8328	0.8638	0.8930	0.9195	0.9427	l(65)/l(35)
l(70)/l(35)	0.5768	0.6154	0.6561	0.6985	0.7419	0.7855	0.8281	0.8683	0.9047	l(70)/l(35)
l(75)/l(35)	0.4157	0.4551	0.4990	0.5474	0.6002	0.6566	0.7155	0.7749	0.8321	l(75)/l(35)
e_5	61.05	62.71	64.42	66.19	68.02	69.91	71.87	73.89	75.99	e_5
${}_{35}q_{15}$	0.1544	0.1333	0.1132	0.0943	0.0768	0.0609	0.0467	0.0344	0.0240	${}_{35}q_{15}$
${}_{35}q_{30}$	0.3123	0.2783	0.2438	0.2094	0.1755	0.1429	0.1123	0.0845	0.0602	${}_{35}q_{30}$

NOTE: Calculated using the Brass General model life table family with parameter β=1. See Annex II for details.

TABLE IV.7. TWO CENSUS METHOD FOR ESTIMATING ADULT FEMALE MORTALITY FROM INFORMATION ON SURVIVAL OF MOTHERS: ZIMBABWE, 1982-1992

Age group	*Proportion of respondents with mother alive*		*Growth rate*	*Intercensal average*	*Synthetic ratio*	*Adjusted proportion with mothers alive*	*Regression coefficients*				*Conditional probability of survival*	*Probability of dying between age 30 and 65 years*
	1982 census	*1992 census*					a_0	a_1	a_2	x	l_{25+x}/l_{25}	${}_{35}q_{30}$
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*	*(12)*	*(13)*
5-9	0.9840	0.9848	0.000081	0.9844	0.9888	0.9844	-0.2894	0.00125	1.2559	10	0.9803	0.1800
10-14	0.9734	0.9731	-0.000031	0.9733	0.9783	0.9734	-0.1718	0.00222	1.1123	15	0.9702	0.1720
15-19	0.9547	0.9515	-0.000336	0.9531	0.9638	0.9523	-0.1513	0.00372	1.0525	20	0.9503	0.1920
20-24	0.9177	0.9199	0.000239	0.9188	0.9564	0.9178	-0.1808	0.00586	1.0267	25	0.9179	0.2194
25-29	0.8697	0.8810	0.001291	0.8754	0.9453	0.8778	-0.2511	0.00885	1.0219	30	0.8822	0.2250
30-34	0.8046	0.8315	0.003289	0.8181	0.9286	0.8298	-0.3644	0.01287	1.0380	35	0.8405	0.2180
35-39	0.7157	0.7659	0.006779	0.7408	0.8919	0.7705	-0.5181	0.01795	1.0753	40	0.7897	0.2016
40-44	0.6102	0.6625	0.008223	0.6364	0.8804	0.6872	-0.6880	0.02342	1.1276	45	0.7122	0.1871
45-49	0.5165	0.5595	0.007997	0.5380	0.7999	0.6050	-0.8054	0.02721	1.1678	50	0.6277	0.1536
50-54	0.4038	0.4277	0.005750	0.4158	NA	0.4840	NA	NA	NA	NA	NA	NA
											Median	0.1920
											0.5 * interquartile range	0.0190
											Error	4.9

Source: Column 2 from Table IV.1 (column 4) and column 3 from: Zimbabwe, Central Statistical Office (1985). 1982 Population Census: Harare, Zimbabwe.

Procedure

Columns 1-3. Record the proportions of persons in each age group with mother surviving at the first and second census.

Column 4. Compute the age-specific growth rate of the proportion surviving using the formula

$$r(x,5) = \ln[S_2(x,5)/S_1(x,5)]/t \qquad (9)$$

where $S_i(x,5)$ denotes the proportion of persons in the x to $x+5$ age group at the I-th census with mother surviving and t denotes the length of the intercensal interval.

Column 5. Compute the average proportion $S(x,5)$ of persons with mother surviving between the two censuses using the formula

$$S(x,5) = [S_1(x,5) \text{ and } S_2(x,5)]/2 \qquad (8)$$

Column 6. Compute the synthetic ratios

$$R(x,5) = \frac{S(x+5,5)\exp[2.5r(x+5,5)]}{S(x,5)\exp[-2.5r(x,5)]} \qquad (7)$$

Column 7. Enter the value of the intercensal average in column 5 for age 5-9 and compute the adjusted proportions of persons with mother surviving for subsequent age groups using the formula

$$S^*(x,5) = R(x-5,5)S^*(x-5,5), \qquad (10)$$

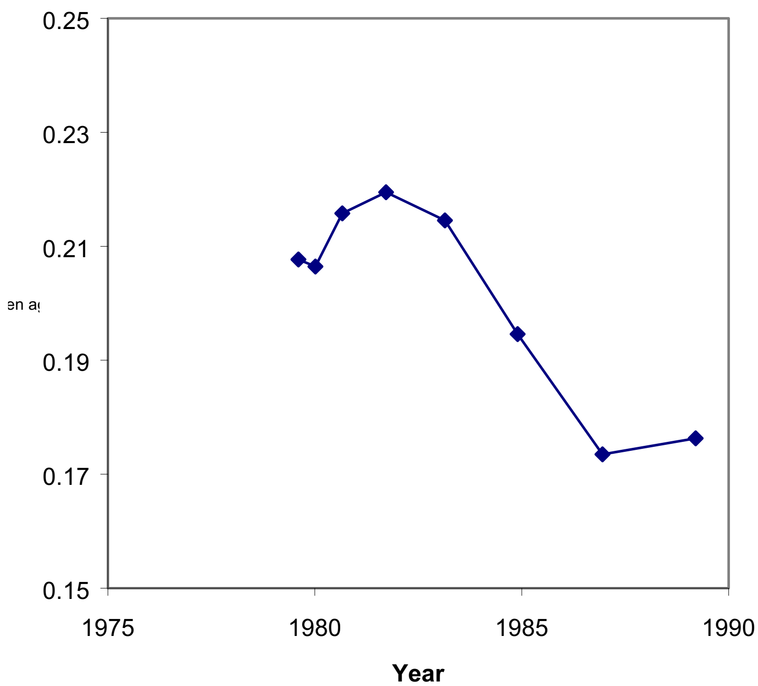
Columns 8-12. Using the coefficients in columns 8-10 and the formula

$$l_{25+x}/l_{25} = a_0(x) + a_1(x)M + a_2(x)S(x-5,5), \qquad (1)$$

calculate the estimated conditional survival probabilities l_{25+x}/l_{25}. Enter the value of x in column 11 and the survivorship probability in column 12. The value of M is calculated from data on births during the 12 months prior to the census. See text for further discussion.

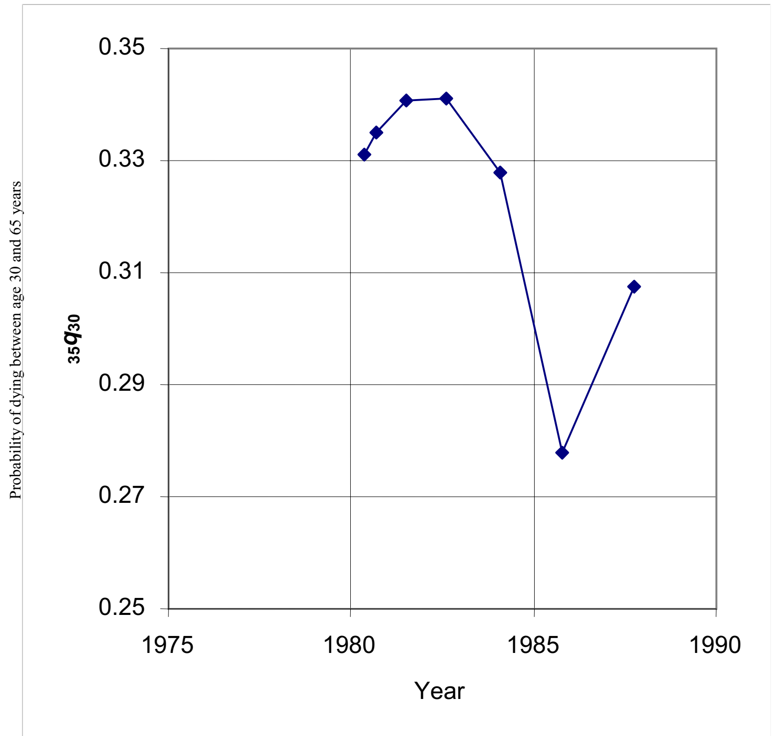
Column 13. Calculate the conditional probability of dying before age 65 years, given survival to age 30 years, corresponding to the given value of l_{25+x}/l_{25} by interpolation in table IV.5.

Figure IV.1. Estimation of adult female mortality from survival of mothers, Zimbabwe, 1992 census: estimated dated probabilities of dying between 30 and 65 years



Source: Columns 16 and 10 of table IV.1.

Figure IV.2. Estimation of adult male survival from survival of fathers, Zimbabwe, 1992 census: estimated dated probabilities of dying between 30 and 65 years



Source: Columns 17 and 11 of table IV.3

V. ESTIMATES DERIVED FROM INFORMATION ON SURVIVAL OF SIBLINGS

The idea of using information on the survival of siblings to estimate mortality arose from the consideration that, on average, the ages of siblings are likely to be very close to the age of a respondent. The proportion of a respondent's siblings who are still alive would, therefore, be a good estimator of survival to the age of the respondent. Although the approach had methodological appeal because the relationship between the proportion surviving and probability of survival was extremely strong, practical problems were encountered in the application of the method. First, field experience with the approach suggested that it was difficult to make clear to interviewers that the respondent was not to be included among his or her siblings. Second, siblings who died before or shortly after the birth of the respondent were likely to be omitted by the respondent.

Interest in the sibling survivorship approach was revived by the proposal that information on the survival of the sisters of a respondent could provide a basis for measuring maternal mortality. Graham and others (1989) showed that if adult female respondents are asked how many of their sisters (born of the same mother) survived to the age of 15, and how many of them died thereafter, and if it can also be ascertained whether siblings who died were pregnant at the time of death or had been pregnant during the 6 to 8 weeks before death, the proportions of sisters who had died of maternal causes could be converted into estimates of the maternal mortality rate. Limiting the consideration of siblings to only those who survived to age 15 years is intended to prevent the omission of siblings who died while still young and who could therefore have been forgotten by the respondent.

Although the "sisterhood method", as it became known, focussed only on maternal mortality, its development stimulated the collection of data on the survivorship of sisters in a wide variety of settings and led to the development of a maternal mortality module for inclusion in the Demographic and Health Surveys. This module was based on a full sibling history, that is, asking a respondent for the name, sex, date of birth, survival status and, if dead, age at death for each sibling born of the same mother.

The availability of sibling survivorship data permit the calculation of estimates of adult mortality using standard life table methods. They also allow the indirect estimation of adult mortality from proportions surviving of brothers (to estimate male mortality) or sisters (to estimate female mortality) by age of respondent. Application of the sibling method requires that information on sibling survival be available for each respondent aged 15 years and over (or aged 15 to 49 years). These data, categorized by five-year age group of respondent, represent the basic inputs of the method.

A. ASSUMPTIONS AND PROCEDURES OF THE SIBLING SURVIVAL METHOD

As with all indirect methods, the sibling method estimates average mortality over an extended period in the past. If mortality trends have been reasonably regular over that period, it will be possible to arrive at an approximate reference date for each estimate. The method also assumes that the age pattern of mortality is similar to those of model life tables, which are required for the estimation. It also assumes that the correlation between the mortality experienced by siblings is not strong, and that most respondents have some siblings (the method would not work well in a country with a long history of low fertility where the proportion of persons without siblings is high).

Assuming that the siblings of a respondent aged x were, on average, also born x years ago, the proportion surviving among these siblings should approximate the probability of surviving to age x, l_x/l_0. The same argument applies if consideration is limited to siblings who survive to age 15. In this case, for respondents aged x, the proportion of siblings surviving among those who had already survived to age 15 should approximate l_x/l_{15}.

Timaeus and others (1997) have calculated the relationship between the proportions of surviving siblings and life table probabilities of surviving from age 15. These model relationships turn out to be very strong and are effectively the same for males and females. For both males and females, the relationship can be expressed as:

$$l_x/l_{15} \;=\; a(x) \;+\; b(x)S(x\text{-}5,5) \qquad (1)$$

where $S(x\text{-}5,5)$ is the proportion of brothers (or sisters) who, having survived to age 15, are still alive among those reported by respondents aged $(x\text{-}5,x)$.

B. Application to males enumerated in the 1994 Demographic and Health Survey, Zimbabwe

Table V.1 illustrates how the estimation of adult male mortality from the survival of brothers, as reported in the 1994 Zimbabwe Demographic and Health Survey (DHS), is carried out. It should be noted that the data used as input are derived from a full sibling history (that is, from recording the survival status of all siblings of each respondent). Tabulation is, however, limited to those siblings who survived to age 15. The data have been expanded using the sampling weights corresponding to the respondents. In principle, the basic data could also have been derived from simpler questions on numbers of surviving brothers, numbers of surviving sisters, and numbers of brothers and sisters who survived to age 15. However, no examples with data gathered in that way could be found. Details of the calculation are provided in the notes to the table.

C. Using model life tables to assess results

The estimated survival probabilities shown in column 8 of table V.1 should decline with age, since the estimates based on older respondents' reports imply greater exposure to mortality risks. It is, however, difficult to judge whether the estimates decline sufficiently from one age to the next. To make this assessment, conversion to a common index of mortality, as was done in the previous chapter, is necessary. This provides a convenient way of making the estimates comparable, both with each other and with estimates from other sources. Conversions have been made in column 13 of table V.1 to a common statistic, in this case ${}_{35}q_{15}$, which is the conditional probability of survival to age 50, given survival to age 15. The translation is facilitated by table V.2, which shows life table estimates of conditional probabilities of survival and implied life expectancy estimates for given values of l_x/l_{15}.

The translated ${}_{35}q_{15}$ values in table V.1 range from 0.0609, as estimated for respondents aged 45-49, to 0.2303 for respondents aged 20-24. This suggests strongly that adult male mortality has increased sharply over time. As in the case of parental survival discussed in chapter IV, the siblings of older persons have been exposed to the risk of dying over a period extending into the more distant past than the siblings of younger persons. If mortality had been falling during the years prior to the survey, the mortality risks experienced by the siblings of older respondents would have been higher than those experienced by the siblings of younger respondents. Although the pattern observed here suggests that mortality has been rising, this could also be due to errors in the data.

If the change in mortality has been approximately linear over time, it is possible to estimate time locations for the estimates, just as for the estimates derived from information on the survival of mothers and fathers (chapter IV). Timaeus and others (1991c) provide a simplification of the procedure of Brass and Bamgboye (1981) for estimating the time location of sibling survival estimates. The time reference of each estimate, (measured as the number of years before the survey – $T(x)$), is given by

$$T(x) = c(x) - d(x)ln(S(x-5,5) \qquad (2)$$

where $c(x)$ and $d(x)$ are the coefficients shown in columns 9 and 10 of table V.1.

D. Assessing mortality trends

The time references calculated using equation (2) are shown in columns 11 and 12 of table V.1. They indicate that the mortality estimates obtained refer to periods much closer to the survey date than the reference periods of estimates based on the survival of parents which was discussed in chapter IV. In this example, the value of ${}_{35}q_{15}$, based on respondents aged 20-24, applies to 1991.4 or roughly 3 years before the survey.

The mortality estimates plotted in figure V.1 show a consistent increase in adult male mortality risks in Zimbabwe from the early 1980s to the early 1990s. The leftmost point in the series, derived from the 45-49 age group, is an outlier, and can be ignored. The remaining points show a substantial increase in the probability of adult death from 0.15 to about 0.23 in less than 10 years.

It is important to note that the estimation equations (1) and (2) are derived on the assumption that the underlying age pattern of mortality does not change. An increase in deaths due to the prevalence of HIV/AIDS in Zimbabwe from the late 1980s invalidates this assumption because AIDS deaths are concentrated in adult ages, whereas non-AIDS deaths are concentrated in very young and very old ages. The analysis of synthetic data given in Timæus and others (1998) suggests that the errors incurred by a rise in AIDS deaths are modest, generally 5 per cent or less.

E. Application to females enumerated in the 1994 Demographic and health survey, Zimbabwe

Table V.3 shows the estimation of female survivorship from data on survival of sisters, as obtained from the 1994 Zimbabwe Demographic and Health Survey (DHS). The calculations are the same as those in the case of male survivorship.

Figure V.2 shows the estimates for females to be similar to those shown in figure V.1 for males. Both sets of estimates show similar patterns. For females (figure V.2), the second point in the series, derived from the 40-44 age group of respondents, is somewhat anomalous, but the remaining points display a fairly regular upward trend from a ${}_{35}q_{15}$ of 0.11 at the beginning of the series to just under 0.21 at the end. Although this increase in mortality is not as high as that noted for males, it is still a substantial increase, which may be attributable to the same factors underlying the increase in male mortality.

TABLE V.1. ADULT MALE MORTALITY ESTIMATED FROM SURVIVAL OF BROTHERS: ZIMBABWE, 1994 DHS

Respondent's age group	*Number of male siblings*		*Age*	*Proportion of brothers alive*	*Estimation coefficients*		*Estimated*	*Time location coefficients*		*Years back*	*Date of estimate*	*Implied*
	Alive	*Dead*	*x*	*S(x-5,5)*	*a(x)*	*b(x)*	l_x/l_{15}	*c(x)*	*d(x)*	*T(x)*		${}_{35}q_{15}$
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*	*(12)*	*(13)*
15-19	2,045	99	NA	0.9538	NA	NA	NA	NA	NA	NA	NA	NA
20-24	2,592	140	25	0.9488	-0.0003	1.0011	0.9495	3.23	1.12	3.29	1991.4	0.2303
25-29	2,213	147	30	0.9377	-0.1546	1.1560	0.9294	5.46	1.95	5.59	1989.1	0.2073
30-34	2,293	171	35	0.9306	-0.1645	1.1660	0.9206	7.52	2.78	7.72	1987.0	0.1724
35-39	1,736	170	40	0.9108	-0.1388	1.1406	0.9001	9.38	3.62	9.72	1985.0	0.1650
40-44	1,293	154	45	0.8936	-0.1140	1.1168	0.8839	11.00	4.45	11.50	1983.2	0.1495
45-49	935	159	50	0.8547	-0.1018	1.1066	0.8440	12.32	5.28	13.15	1981.6	0.0609
											Median	0.1687
											0.5 * interquartile range	0.0226
											Per cent	13.4

Source: Unpublished data from the 1994 Zimbabwe Demographic and Health Survey. Coefficients *a(x)* and *b(x)* in columns 6 and 7 and *c(x)* and *d(x)* in columns 9 and 10 from Timaeus and others (1991c), table 3.

Procedure

Columns 1-3: Enter the age distribibution and number of brothers of persons in each age group who survived to age 15 and who were still alive at the time of interview, and the number of brothers of persons in each age group who survived to age 15 but were deceased at the time of interview.

Column 4: Enter the upper limit of age group *x*.

Column 5: Compute the proportion of brothers who survived to age 15 and who were still alive at the time of interview, i.e. for each age group, number in column 2 divided by sum of numbers in columns 2 and 3.

Columns 6-8: Using the coefficients in columns 6 and 7 and the formula

$$l_x/l_{15} \;=\; a(x) \;+\; b(x)S(x\text{-}5,5) \qquad (1)$$

calculate l_x/l_{15} for $x = 25, 30, ..., 50$.

Columns 9-11: Using the coefficients in columns 9 and 10 and the formula

$$T(x) \;=\; c(x) \;-\; d(x)S(x\text{-}5,5) \qquad (2)$$

calculate "years back", *T(x),* for $x = 25, 30, ..., 50$.

Column 12: Calculate the time to which each estimate refers by subtracting the years back from the reference time of the census or survey.

Column 13: Calculate the value of ${}_{35}q_{15}$ corresponding to each value of l_x/l_{25} using table V.2. The interpolation procedure is described in annex II, section E.

TABLE V.2. ADULT MALE MORTALITY ESTIMATED FROM SURVIVAL OF BROTHERS:
CONDITIONAL SURVIVAL PROBABILITIES l_x/l_{15} FOR MODEL LIFE TABLE TRANSLATION

e_0	*20.00*	*22.50*	*25.00*	*27.50*	*30.00*	*32.50*	*35.00*	*37.50*	*40.00*	*e_0*
Column	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*Column*
l(25)/l(15)	0.8467	0.8561	0.8654	0.8745	0.8834	0.8921	0.9006	0.9088	0.9168	*l(25)/l(15)*
l(30)/l(15)	0.7716	0.7845	0.7973	0.8100	0.8226	0.8350	0.8472	0.8591	0.8708	*l(30)/l(15)*
l(35)/l(15)	0.7030	0.7183	0.7337	0.7492	0.7646	0.7800	0.7952	0.8103	0.8252	*l(35)/l(15)*
l(40)/l(15)	0.6359	0.6529	0.6703	0.6878	0.7055	0.7234	0.7413	0.7591	0.7769	*l(40)/l(15)*
l(45)/l(15)	0.5676	0.5857	0.6044	0.6235	0.6430	0.6628	0.6829	0.7032	0.7236	*l(45)/l(15)*
l(50)/l(15)	0.4952	0.5138	0.5331	0.5531	0.5737	0.5949	0.6167	0.6390	0.6617	*l(50)/l(15)*
e_5	40.07	41.12	42.20	43.31	44.45	45.62	46.83	48.08	49.36	e_5
${}_{35}q_{15}$	0.5048	0.4862	0.4669	0.4469	0.4263	0.4051	0.3833	0.3610	0.3383	${}_{35}q_{15}$
${}_{35}q_{30}$	0.6733	0.6604	0.6466	0.6316	0.6156	0.5983	0.5797	0.5597	0.5383	${}_{35}q_{30}$

e_0		*42.50*	*45.00*	*47.50*	*50.00*	*52.50*	*55.00*	*57.50*	*60.00*	*e_0*
Column		*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*Column*
l(25)/l(15)		0.9244	0.9318	0.9388	0.9455	0.9518	0.9578	0.9634	0.9686	*l(25)/l(15)*
l(30)/l(15)		0.8821	0.8931	0.9037	0.9139	0.9236	0.9328	0.9415	0.9497	*l(30)/l(15)*
l(35)/l(15)		0.8398	0.8541	0.8680	0.8814	0.8944	0.9068	0.9186	0.9297	*l(35)/l(15)*
l(40)/l(15)		0.7946	0.8120	0.8291	0.8458	0.8620	0.8777	0.8927	0.9070	*l(40)/l(15)*
l(45)/l(15)		0.7440	0.7645	0.7847	0.8047	0.8244	0.8436	0.8621	0.8800	*l(45)/l(15)*
l(50)/l(15)		0.6848	0.7081	0.7315	0.7549	0.7782	0.8012	0.8237	0.8456	*l(50)/l(15)*
e_5		50.68	52.03	53.43	54.86	56.34	57.86	59.43	61.05	e_5
${}_{35}q_{15}$		0.3152	0.2919	0.2685	0.2451	0.2218	0.1988	0.1763	0.1544	${}_{35}q_{15}$
${}_{35}q_{30}$		0.5154	0.4910	0.4650	0.4374	0.4082	0.3775	0.3455	0.3123	${}_{35}q_{30}$

e0		*62.50*	*65.00*	*67.50*	*70.00*	*72.50*	*75.00*	*77.50*	*80.00*	*e0*
Column		*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*Column*
l(25)/l(15)		0.9734	0.9778	0.9818	0.9854	0.9886	0.9914	0.9937	0.9956	*l(25)/l(15)*
l(30)/l(15)		0.9573	0.9643	0.9707	0.9764	0.9815	0.9860	0.9898	0.9929	*l(30)/l(15)*
l(35)/l(15)		0.9401	0.9498	0.9586	0.9667	0.9738	0.9801	0.9855	0.9899	*l(35)/l(15)*
l(40)/l(15)		0.9205	0.9331	0.9447	0.9553	0.9649	0.9732	0.9804	0.9864	*l(40)/l(15)*
l(45)/l(15)		0.8969	0.9129	0.9278	0.9414	0.9538	0.9647	0.9741	0.9819	*l(45)/l(15)*
l(50)/l(15)		0.8667	0.8868	0.9057	0.9232	0.9391	0.9533	0.9656	0.9760	*l(50)/l(15)*
e_5		62.71	64.42	66.19	68.02	69.91	71.87	73.89	75.99	e_5
${}_{35}q_{15}$		0.1333	0.1132	0.0943	0.0768	0.0609	0.0467	0.0344	0.0240	${}_{35}q_{15}$
${}_{35}q_{30}$		0.2783	0.2438	0.2094	0.1755	0.1429	0.1123	0.0845	0.0602	${}_{35}q_{30}$

NOTE: Calculated using the Brass General model life table family with parameter β=1. See annex II for details.

TABLE V.3. ADULT FEMALE MORTALITY ESTIMATED FROM SURVIVAL OF SISTERS: ZIMBABWE, 1994 DHS

Respondent's age group	*Number of female siblings*		*Age*	*Proportion of sisters alive*	*Estimation coefficients*		*Estimated*	*Time location coefficients*		*Years back*	*Date of estimate*	*Implied*
	Alive	*Dead*	*(x)*	*S(x-5,5)*	*a(x)*	*b(x)*	l_x/l_{15}	*c(x)*	*d(x)*	*T(x)*		${}_{35}q_{15}$
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*	*(12)*	*(13)*
15-19	2,017	108	NA	0.9492	NA	NA	NA	NA	NA	NA	NA	NA
20-24	2,591	122	25	0.9550	-0.0003	1.0011	0.9558	3.23	1.12	3.28	1991.4	0.2065
25-29	2,236	125	30	0.9471	-0.1546	1.1560	0.9402	5.46	1.95	5.57	1989.2	0.1797
30-34	2,344	141	35	0.9433	-0.1645	1.1660	0.9353	7.52	2.78	7.68	1987.0	0.1430
35-39	1,849	131	40	0.9338	-0.1388	1.1406	0.9263	9.38	3.62	9.63	1985.1	0.1240
40-44	1,429	155	45	0.9021	-0.1140	1.1168	0.8935	11.00	4.45	11.46	1983.3	0.1375
45-49	983	117	50	0.8936	-0.1018	1.1106	0.8907	12.32	5.28	12.91	1981.8	0.1093
											Median	0.1402
											0.5 * interquartile range	0.0216
											Per cent	15.4

Source: Unpublished data from the 1994 Zimbabwe Demographic and Health Survey. Coefficients in columns 6-7 and 9-10 from Timaeus and others (1991c), table 3.

Procedure

Columns 1-3: Enter the number and age distribution of sisters of persons in each age group who survived to age 15 and who were still alive at the time of interview, and the number of sisters of persons in each age group who survived to age 15 but were deceased at the time of interview.

Column 4: Enter the upper limit of age group *x*.

Column 5: Compute the proportion of sisters who survived to age 15 and who were still alive at the time of interview, i.e. for each age group, the number in column 2 divided by sum of numbers in columns 2 and 3.

Columns 6-8: Using the coefficients in columns 6 and 7 and the formula

$$l_x/l_{15} = a(x) + b(x)S(x-5,5) \quad (1)$$

calculate l_x/l_{15} for $x = 25, 30, \ldots, 50$.

Columns 9-11: Using the coefficients in columns 9 and 10 and the formula

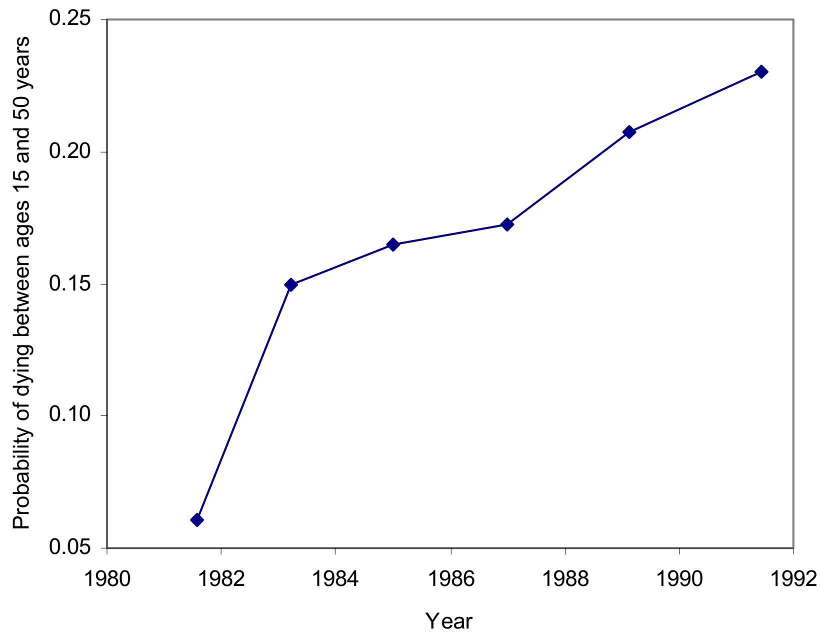
$$T(x) = c(x) - d(x)S(x-5,5) \quad (2)$$

calculate "years back", $T(x)$, for $x = 25, 30, \ldots, 50$.

Column 12: Calculate the time to which each estimate refers by subtracting the "years back" from the reference time of the census or survey.

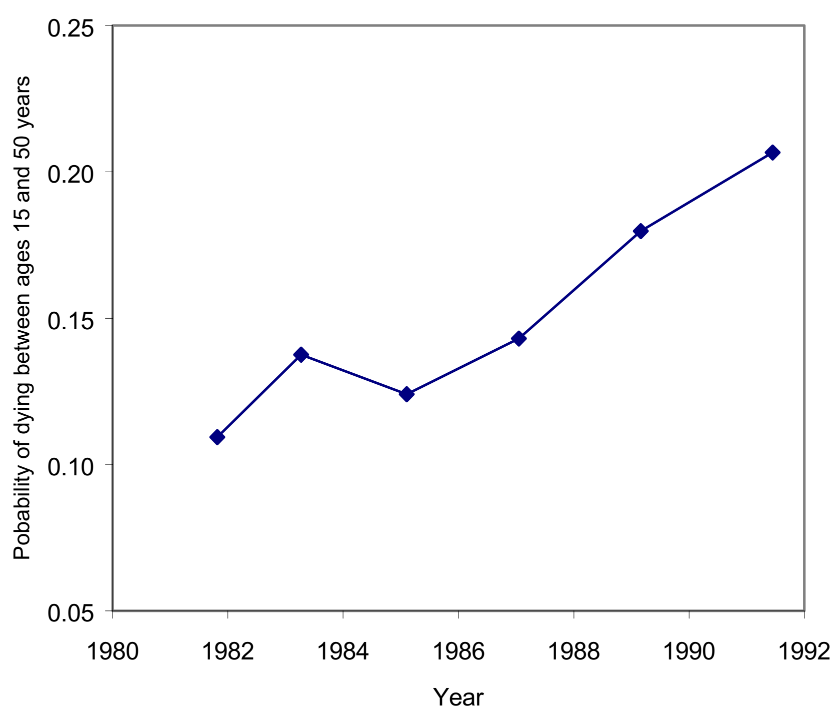
Column 13: Calculate the value of ${}_{35}q_{15}$ corresponding to each value of l_x/l_{25} using table V.2. The interpolation procedure is described in annex II, section E.

Figure V.1. Dated male probability of dying between ages 15 and 50 as estimated from information on survival of brothers, Zimbabwe, 1994 DHS



Source: Columns 12 and 13 of table V.1.

Figure V.2. Dated female probability of dying between ages 15 and 50 as estimated from information on survival of sisters, Zimbabwe, 1994 DHS



Source: Columns 12 and 13 of table V.3.

ANNEX I
Practical matters

Application of the methods described in this manual involves various practical considerations that are learned by experience by anyone who applies them sufficiently often. These matters are often not formally taught, however, and no convenient reference is available. This appendix discusses a number of these practical issues.

A. PRIMARY SOURCES

Primary sources should be used for assembling required data insofar as possible. This is important partly because secondary sources may contain errors, but also because primary sources often contain information on context that is usually important and sometimes essential to appropriate interpretation of the results. Much of the work of getting useful estimates by application of the methods described here involves assessing likely errors in the input data. To do this it is essential to know far more about the data than would be required in many other contexts. It is important to know not just what the data purport to represent, but the source from which they derive and the way in which they have been generated.

This rule applies even to data so evidently transparent and standardized as population age distributions. In the application of any of the three intercensal deaths methods, for example, it is essential to know whether the age distributions derive from a population census, *i.e.*, a complete enumeration of the population, or from a sample. This is because the results of the growth balance method depend critically on the age distributions representing the size of the population as well as its age composition. If both age distributions are from censuses, the only issue is the relative completeness of enumeration. If one age distribution is from a survey, however, it is essential to know how the sample counts were inflated to population totals.

B. ASSEMBLING SOURCE MATERIALS

Insofar as possible, all pertinent sources of data for a given country should be assembled before embarking on an analysis. Application of the estimation methods described here will often result in questions that can be answered, to the extent that they can be answered, only by consulting statistical information ranging far beyond their nominal required input.

Census survival and intercensal deaths methods, for example, assume a population closed to migration. Often this assumption will be doubtful, and it will be important to ascertain what evidence is available on the level, direction and sex and age patterns of migration. In this connection one will want to know, for example, whether the available censuses included questions on place of birth, and if so, whether tabulations are available to suggest how important immigration might be.

Censuses of countries that receive international migrants from the country under study may sometimes be used to obtain information on emigration, for example. Complete and accurate statistics on international migration that would allow for formal statistical correction are almost never available, but available information will often assist in interpreting problematic results and arriving at better conclusions.

C. CHECKING INPUT DATA

The first step in applying any method is to record the necessary input data. When calculations are carried out using a hand calculator, this first step usually consists of transcribing numbers and labels from the source to a worksheet of some kind. When using a computer spreadsheet or other program, it may consist either of keying in data from a source or "copying" and "pasting" from one computer file to another.

In either case, input data should always be checked before proceeding to the next step. This may seem so obvious as to be not worth mentioning, but as with many other elementary disciplines, there is a constant temptation to get through the tedious initial steps quickly and get on with the more interesting work. Those who do not learn to resist this temptation from admonition will eventually learn it by more or less painful experience.

A very effective and widely applicable check consists of transcribing or keying both a set of numbers and their sum from the source, summing the entered numbers, and checking whether the calculated sum equals the entered sum. This is sometimes called a "sum check," and it is very effective at catching simple keying and transcribing errors. Sum checks should always be carried out when applicable and any discrepancies immediately rectified.

It is important to remember that data consists not merely of numbers, but also of words that lend meaning to the numbers. Thus the number

10,401,767

is not data, but merely a number. It becomes data only if it is suitably labelled, as, for example, the number of persons enumerated by the 1992 population census of Zimbabwe.

It follows that checking for errors in data means checking the accuracy of labels as well as checking the accuracy of numbers. The numbers of males and females in each age group may be correct as numbers, but interchanging the "male" and "female" labels renders all the numbers, considered as data, wrong. Errors of this kind are easier to make than the inexperienced might suppose, in part because the work is tedious and intellectually trivial, so that attention may wander. The risk of errors of this kind is probably greater when using computer spreadsheets or other programs because great masses of numbers may be moved from one place to another with very little effort.

Input data often includes, in addition to what we are likely to think of as "the data proper," various supplementary information, such as the dates of population censuses or the values of expectation of life at some old age required to compute life tables. These inputs also must be checked.

D. Not stated values

The data to which the methods are applied will very often include numbers of cases for whom information is not stated. Not stated values should be prorated, *i.e.*, eliminated by distributing them among the stated cases in the same proportions as the stated cases. They should never be incorporated into the open-ended age group.

E. Open-ended age groups

It will sometimes be desirable to lower the open ended age group provided in the data, to reduce the impact of age exaggeration, because the old age detail is considered unimportant, or simply because an available computer program will not accommodate the available open ended group. The older age groups of a source age distribution consisting of five-year ages groups to 95-99 and a concluding 100+ age group may, for example, be merged to a 75+ age group.

F. Significant figures and rounding

Counts of persons and events deriving from censuses, surveys and vital registration should in general be given in full detail, even though this often entails carrying large numbers of insignificant figures. Rounding to the nearest thousand to reduce the number of insignificant figures will all too often lead to difficulty because some numbers such as the numbers of persons in very old age groups and the numbers of deaths in young adult age groups end up with too few significant figures. It also complicates the use of sum checks, which are easier to apply if one does not have to decide in each case whether a discrepancy could be accounted for by rounding errors.

Prorating not stated cases will give fractional numbers of persons in the categories prorated to. When working manually, results will of course be rounded to the nearest whole number. When working with spreadsheets, however, some ten or more digits after the decimal place will be carried by default. It makes no sense to input this information to the methods, and doing so will make it slightly more difficult to check results. When working with spreadsheets, therefore, prorated numbers should be explicitly rounded to the nearest whole number using an appropriate spreadsheet function. This will, of course, result in slight discrepancies between the rounded terms and their total.

G. Number of places after decimal point

All the methods require the calculation of proportions and/or ratios and so require a decision on how many places after the decimal place should be carried.

Several general rules are applicable. First, no more digits should be carried than is justified by the precision of the values calculated. Other things being equal, too many digits are a misleading nuisance and distraction.

Second, it is better to err on the side of one too many than one too few places. Carrying too few places results in information loss, which is more serious than any consequence of carrying too many places. It should be borne in mind that Asignificance≅ depends on context. Four places after the decimal may be well over the precision that can expected of the source data, but still useful for comparisons internal to the method.

A third general rule, to be followed if it does not entail serious violation of the first two, is not to vary the number of digits after the decimal any more than necessary.

Fourth, identify those circumstances in which the number of places is particularly important. When using population growth rates to calculate person years lived, for example, it may be necessary to maintain six places after the decimal to have a sufficient number of significant figures.

H. Importance of manual calculation

Computers are increasingly available nearly everywhere in which work of the kind described in this manual is done, and computers should certainly be used for doing much of it, where they are available. Precisely because this is the case, it is important to emphasise the value of manual calculation, which in this context means working with pencil, paper and a hand calculator. If a prepared program or spreadsheet is used to apply a method, all that is required to produce the initial output is to enter the input data. Doing this teaches one nothing about how the method works.

Creating a spreadsheet or a computer program to implement a method requires some understanding of the method. The required understanding is abstract, however, divorced from the details of any particular data set, and the necessity of figuring out how best to program the method distracts attention from the method itself.

Manual application of a method has the great virtue of focusing attention not only on the details of the method itself, but on the details of the particular data set to which it is applied. The relatively slow pace of the work combined with the routine nature of keying in numbers and recording results allows and encourages the mind to focus in depth on exactly what is being done and what the results are at each stage of the process. This leads to a deeper understanding of both of the method and of the particular data to which it is applied, than any use of a computer for the same purpose.

A good rule of thumb is that one should not use a computer to apply a method until one has applied it by hand at least several times, preferably several times on different sets of data. Once learned, none of the methods described in this manual takes much more than an hour to carry out with a hand calculator. When applying a method for the first time, however, one may expect to spend perhaps three times this long. The learning process that reduces execution time is very valuable. It is easy to read the description of a method and suppose that one understands it, but an application to actual data nearly always reveals some lack of understanding.

I. Use of computer spreadsheet programs

Once a method has been learned by applying it manually to several sets of data, the case for using computer spreadsheets is very strong. Computers have become nearly universally available and are therefore familiar to nearly everyone likely to be involved in work of this nature.

Spreadsheets are ideal tools for data entry, checking and pre-processing, saving much time and tedium. They generally include powerful built-in functions for plotting, equation solving, and numerical minimization. The plotting functions, in particular, enable one to produce plots with vastly less effort, indeed with almost no effort, than would be required to produce plots manually.

A further advantage that will become increasingly important in the future, and that is important in many contexts already, is that by incorporating data and results in digital form, spreadsheets make it possible to store and transmit results far more efficiently than is possible with results on paper.

Many of these advantages may be realized with other kinds of computer software. The advantage of spreadsheets is their combination of considerable power and exceptionally broad availability, which means that nearly everyone involved in work of this kind is likely to have them and know how to use them.

J. Documentation

The importance of documenting work as it proceeds can hardly be over emphasised, not so much because it is important, which ought to be too obvious to require explicit mention, but because the temptation to avoid or defer it are so strong. The twin purposes of documentation are quality control and efficiency. Knowing where data came from or how calculations were made is necessary to check whether the data and the calculations are correct. Large quantities of time may be wasted searching for data sources, or trying to figure out how some simple calculation was done, when it would have taken only a few minutes to document at the time the data was retrieved or the calculation made.

It is good practice to record the source of data before the data itself, making it less likely that the source will be omitted. It follows that source notes are better placed at the top than the bottom of worksheets, whether paper worksheets or computer spreadsheets. Source notes should indicate full detail, including page and table numbers as well as bibliographic information.

K. Calculations with dates

Calculations with dates are facilitated by determining the fraction of a year represented by the date the data pertain to. This is done by adding the number of days in the months preceding the census or survey, to the date of the month in question and dividing by 365. The reference date of the Japanese censuses since 1950, for example, is October 1, which translates into

$$(31+28+31+30+31+30+31+31+30+1)/365$$

or 274/365 = 0.751. Thus the time of the 1960 census in decimal form 1990.751. Precision to a single place after the decimal will suffice for most practical work. It is recommended that three places after the decimal be routinely recorded, however, because this allows recovery of a date from its decimal equivalent. This may be seen in the table below, which shows all dates and their decimal equivalents.

ANNEX TABLE I-1. TRANSLATION TABLE FOR DECIMAL FORMS OF DATES

Day\Month	***Jan***	***Feb***	***Mar***	***Apr***	***May***	***Jun***	***Jul***	***Aug***	***Sep***	***Oct***	***Nov***	***Dec***
1	0.003	0.088	0.164	0.249	0.332	0.416	0.499	0.584	0.668	0.751	0.836	0.918
2	0.005	0.090	0.167	0.252	0.334	0.419	0.501	0.586	0.671	0.753	0.838	0.921
3	0.008	0.093	0.170	0.255	0.337	0.422	0.504	0.589	0.674	0.756	0.841	0.923
4	0.011	0.096	0.173	0.258	0.340	0.425	0.507	0.592	0.677	0.759	0.844	0.926
5	0.014	0.099	0.175	0.260	0.342	0.427	0.510	0.595	0.679	0.762	0.847	0.929
6	0.016	0.101	0.178	0.263	0.345	0.430	0.512	0.597	0.682	0.764	0.849	0.932
7	0.019	0.104	0.181	0.266	0.348	0.433	0.515	0.600	0.685	0.767	0.852	0.934
8	0.022	0.107	0.184	0.268	0.351	0.436	0.518	0.603	0.688	0.770	0.855	0.937
9	0.025	0.110	0.186	0.271	0.353	0.438	0.521	0.605	0.690	0.773	0.858	0.940
10	0.027	0.112	0.189	0.274	0.356	0.441	0.523	0.608	0.693	0.775	0.860	0.942
11	0.030	0.115	0.192	0.277	0.359	0.444	0.526	0.611	0.696	0.778	0.863	0.945
12	0.033	0.118	0.195	0.279	0.362	0.447	0.529	0.614	0.699	0.781	0.866	0.948
13	0.036	0.121	0.197	0.282	0.364	0.449	0.532	0.616	0.701	0.784	0.868	0.951
14	0.038	0.123	0.200	0.285	0.367	0.452	0.534	0.619	0.704	0.786	0.871	0.953
15	0.041	0.126	0.203	0.288	0.370	0.455	0.537	0.622	0.707	0.789	0.874	0.956
16	0.044	0.129	0.205	0.290	0.373	0.458	0.540	0.625	0.710	0.792	0.877	0.959
17	0.047	0.132	0.208	0.293	0.375	0.460	0.542	0.627	0.712	0.795	0.879	0.962
18	0.049	0.134	0.211	0.296	0.378	0.463	0.545	0.630	0.715	0.797	0.882	0.964
19	0.052	0.137	0.214	0.299	0.381	0.466	0.548	0.633	0.718	0.800	0.885	0.967
20	0.055	0.140	0.216	0.301	0.384	0.468	0.551	0.636	0.721	0.803	0.888	0.970
21	0.058	0.142	0.219	0.304	0.386	0.471	0.553	0.638	0.723	0.805	0.890	0.973
22	0.060	0.145	0.222	0.307	0.389	0.474	0.556	0.641	0.726	0.808	0.893	0.975
23	0.063	0.148	0.225	0.310	0.392	0.477	0.559	0.644	0.729	0.811	0.896	0.978
24	0.066	0.151	0.227	0.312	0.395	0.479	0.562	0.647	0.732	0.814	0.899	0.981
25	0.068	0.153	0.230	0.315	0.397	0.482	0.564	0.649	0.734	0.816	0.901	0.984
26	0.071	0.156	0.233	0.318	0.400	0.485	0.567	0.652	0.737	0.819	0.904	0.986
27	0.074	0.159	0.236	0.321	0.403	0.488	0.570	0.655	0.740	0.822	0.907	0.989
28	0.077	0.162	0.238	0.323	0.405	0.490	0.573	0.658	0.742	0.825	0.910	0.992
29	0.079	NA	0.241	0.326	0.408	0.493	0.575	0.660	0.745	0.827	0.912	0.995
30	0.082	NA	0.244	0.329	0.411	0.496	0.578	0.663	0.748	0.830	0.915	0.997
31	0.085	NA	0.247	NA	0.414	NA	0.581	0.666	NA	0.833	NA	1.000
Day/Month	***Jan***	***Feb***	***Mar***	***Apr***	***May***	***Jun***	***Jul***	***Aug***	***Sep***	***Oct***	***Nov***	***Dec***

ANNEX II
The use of model life tables

A number of methods discussed in this manual refer to the use of model life tables as tools in the mortality estimation process, or as aids in assessing the reliability or accuracy of data. This appendix discusses the utility of model life tables in adult mortality estimation, explains the rationale for employing them and describes and illustrates pertinent concepts in their application.

A. AGE PATTERNS OF MORTALITY

Although mortality risks vary widely between populations and within the same population over time, the age pattern of human mortality is strongly patterned. The simplest and most general feature is that higher (or lower) mortality risks over any age interval tend to be associated with higher (or lower) risks over all other intervals.

Consider annex figure II.1, which shows conditional probabilities of dying (${}_nq_x$) derived from life tables for Trinidad and Tobago males for the periods 1920-1922, 1945-1947, and 1959-1961. The male expectation of life at birth for Trinidad and Tobago increased from 37.6 years in 1920-1922 to 62.4 years in 1959-1961. It is clear that all age groups benefited from the decline in mortality over time, thus shifting the entire probability of dying function downward with declining mortality. The pattern of mortality decline in Trinidad and Tobago is an example of a pattern of mortality change noted across populations. This tendency for mortality change to be consistent across ages implies that given the value of one statistic, such as e_5, it is possible to derive a reasonably good estimate of another statistic, such as e_{30}. This possibility of "translating" one life table statistic from another is very useful in the indirect estimation of mortality and in analysing the results of various mortality estimation procedures.

While it would be possible to derive *ad hoc* relationships between life table parameters each time they are needed, a simpler and more systematic approach is to use model life tables. Model life tables provide a full life table for a series of mortality levels and they are based on data from observed populations with a variety of mortality experiences.

B. MODEL LIFE-TABLE FAMILIES

Various approaches have been used to express in analytical or tabular form, the variety of frequently observed age and sex patterns of mortality. The first set of model life tables was developed by the Population Division of the United Nations Secretariat in the 1950s. The United Nations model life tables were based on a collection of 158 tables for each sex. The tables allow the estimation of other life table parameters from a single index, such as ${}_1q_0$. One way of displaying the information in life tables is to list tables one after the other. This is the mode of presentation used, for example, in the United Nations *Model Life Tables for Developing Countries* (1982). A set of life-tables and associated stable populations prepared by the Office of Population Research at Princeton University (Coale and Demeny, 1966), have also been widely used because they offer four families of life tables, each of which is based on a regional pattern of mortality.

Annex figure II.2 shows the relationship between the expectation of life at age 5 (e_5) and at age 30 (e_{30}) for the 72 male and female life tables used in the construction of the United Nations *Model Life Tables for Developing Countries* (United Nations, 1982, annex 5, pp. 285-351). Despite the considerable diversity of the national populations represented, the points for observed combinations of e_5 and e_{30} values fall closely along a simple, slightly curved line.

It is important to note that the relationship between e_5 and e_{30} is very close because both statistics refer to post-childhood mortality. The relationship between one statistic pertaining to the childhood years (0-4 years) and another pertaining to post-childhood years is likely to be weaker. Annex figure II.3 shows that the relation between e_5 and ${}_1q_0$ for the same life tables referenced in annex figure II.2. While there is clearly a strong relationship, the points are far more scattered than in annex figure II.2.

An important shortcoming of model life tables is that their accuracy depends on the data that generated them. They also often represent the experience of a limited range of possible human experience. Brass and colleagues (1968), and later Carrier and Hobcraft (1971), have derived life tables based on a logit transformation of corresponding life table probabilities.

The Brass model life table family is defined by a simple mathematical transformation involving two parameters, α and β, and a "standard" set of $logit(l_x)$ values, where l_x is a standard reference schedule, for single years of age from 1 through 99. Broadly speaking, the parameter α represents the level of mortality and the parameter β represents the balance between mortality at older ages and mortality at younger ages. A one-parameter model is obtained from this two- parameter model by fixing the value of β, (see Brass (1971) for a general discussion). The standard logit values for the Brass General model are given in Hill and Trussell (1977, p. 316). This table is reproduced in United Nations (1983, p. 19). In this table, however, two digits in the value shown for the logit of l_{63} are transposed. The correct value, as shown in the original source, is 0.3024, not 0.3204, as is evident from the differences of the series. A slightly different version of the standard, lacking single year detail at ages over 50, is given in Brass (1971, p. 77).

The value of l_x corresponding to any given values of α and β is given by

$$l_x(\alpha,\beta) = 1/[1+exp(\alpha+ \beta Y_x)] \qquad (1)$$

where Y_x denotes the standard logit value. These l_x values suffice to calculate values of q_x (the probability of dying at age x) and d_x (the number of deaths at age x). To calculate the number of person years lived at age x (L_x,), the total number of person years lived above age x (T_x) and the life expectancy at age x (e_x), all of which depend on the continuous series of l_x values, further formulas are required. For $x>1$, L_x may be calculated as

$$L_x= 0.5(l_x + l_{x+1}) \qquad (2)$$

The linearity assumption on which this approximation is based is unsatisfactory for calculating the number of person years lived at age 0 years (L_0). Instead, L_0 is calculated using the procedure detailed in Coale and Demeny (1966, p. 20). Specifically, L_0 is calculated as

$$L_0 = k_0 l_0 + (1-k_0)l_1 \qquad (3)$$

where,

$$k_0 = 0.34 \qquad (4a)$$

if $q_0<0.100$ and

$$k_0 = 0.463 + 2.9375q_0 \qquad (4b)$$

if $q_0 \geq 0.100$. Since L_x values are given to age $x=99$, T_x values may be computed directly from the L_x values. Values of e_x are calculated using

$$e_x = T_x/l_x \qquad (5)$$

Annex tables II.1 through II.4 show the values of selected parameters from the Brass General model life table family.

C. Constructing synthetic data: stationary populations

A relatively unusual but important use of model life tables is constructing synthetic data for purposes of testing the performance of estimation procedures under known conditions. Most estimation procedures involve minor interpolations or approximations that can affect the precision of their results. Often the limitations are insignificant, but in some applications it is important to know precisely what they are.

Annex table II.5 shows two hypothetical age distributions ten years apart and intercensal deaths for a stationary population corresponding to the Brass General model life table with expectation of life at birth 72.5 years and with a radix (annual number of births) 100,000 persons. The age distribution at both points in time is given by the ${}_5L_x$ values of the life table, taken from annex table II.2 and corresponding to a life expectancy of 72.5 years. Since the population is stationary, annual deaths over age x equal the life table numbers of survivors at age x. To obtain the number of intercensal deaths for the ten-year period these annual numbers are multiplied by ten.

Applying the simple growth balance method to the data in annex table II.2 yields a deaths adjustment factor of 1.0004. The ratios for ages $x = 5, 10, ...,$ though generally small, show a very distinct pattern: a slight rise from age *5* to *10,* level from age *10* through about *40*, followed by a gradual and then accelerating rise at older ages. This pattern reflects the imperfect estimation of persons reaching exact age x during the intercensal interval from the census age distributions. In the age ranges in which the survivorship curve is nearly linear, the approximation is very good. The survivorship curve slopes down faster at young ages, however, and rises more sharply at older ages. This results in an over estimation of persons reaching exact age *5* and of those reaching the oldest ages. Applying the general growth balance and extinct generations method to the data gives similar results.

D. Constructing synthetic data: stable populations

Constructing synthetic data for stationary populations is relatively easy because of their very simple structure, but the assumption of stationarity is unacceptable for most developing countries. Stable populations, by contrast, provide a good first approximation to the age distribution of population and deaths observed in many developing country populations.

Annex table II.6 shows the calculation of synthetic data for a stable population with an expectation of life at birth of 60 years and a growth rate of 3 per cent per annum. The calculation makes the standard assumption that survivorship proportions calculated for a stationary population may be applied to a stable population. For most purposes this assumption will be more than adequate. If a very high level of precision is required, alternative methods using single years of age or numerical integration on even smaller age intervals may be required. The calculations are explained in the notes to the table.

E. Deriving model life table parameters through interpolation

To find the model life table value of e_5 corresponding to an estimated $e_{30}=40.3$ and to make the calculation transparent and avoid careless errors, it is useful to make a simple table with space for the pertinent quantities and to proceed step by step as shown below.

	Lower	*Given*	*Upper*
e_{30}	39.39	40.35	40.61
Column No.	17	0.7869	18
e_5	61.05	62.36	62.71

Step 1. Label the rows and columns. The first row is for the statistic to be translated, the last row for the statistic translated to. The remaining row and column labels are the same in all cases.

Step 2. Enter the value to be translated, 40.35 years in this example, in the middle, “Given”, column of the first row.

Step 3. Identify the lower and upper bracketing life tables. The lower bracketing life table is the table that has the highest value of e_{30} lower than the given value. The upper bracketing life table is the table that has the lowest value of e_{30} higher than the given value. In annex table II.2 an e_{30} of 40.35 years is bracketed by $e_{30}=39.39$ years in column 17 and $e_{30}=40.61$ in column 18. Enter these e_{30} values in the “Lower” and “Upper” columns of the first row and the column numbers in the “Lower” and “Upper” columns of the second row. For spreadsheet calculation, use a suitable “lookup” function to identify the bracketing columns.

Step 4. Find the values of the statistic to be estimated, e_5 in this example, from the columns identified in the preceding step. The value of e_5 in column 17 is 61.05 years. The value of e_5 in column 18 is 62.71 years. Enter these values in the first and last column of the last row of the table.

Step 5. Interpolate between the first and last entries in the first row. In this example,

$$(40.35 - 39.39)/(40.61 - 39.39) = 0.7869.$$

Enter this interpolation fraction in the centre cell in the table.

Step 6. Compute the desired estimate by adding the interpolation constant multiplied by the difference between the first and last entries in the last row of the table to the value in the first row, *i.e.*, in this example,

$$61.05 + 0.7869(62.36 - 61.05) = 62.36$$

F. Accuracy of translation

When the relationship between various life table statistics is used to assess the accuracy of adult mortality estimates, a fundamental goal is to ascertain the extent to which the estimates derived from data on various population age groups all point to a similar underlying model life table. In chapter two, for example, where model life tables were used to assess estimated expectation of life for Zimbabwe, it was shown that data reported by different age groups implied somewhat different levels of e_5, suggesting some degree of error in data reporting.

The accuracy of model life table-derived indicators of mortality depends on the closeness of the relationship in the reference life tables, and on the extent to which the reference life tables are representative of the mortality experience of the population for which the estimation is carried out. The more representative the family of life tables selected the better the result of the estimation procedure.

ANNEX TABLE II.1. BRASS GENERAL MODEL LIFE TABLE FAMILY VALUES OF EXPECTATION OF LIFE AT AGE X, X = 0, 5, ...,95, FOR TABLES WITH EXPECTATION OF LIFE AT BIRTH OF 20, 22.5, ..., 90 YEARS

e_0	**20.00**	**22.50**	**25.00**	**27.50**	**30.00**	**32.50**	**35.00**	**37.50**	**40.00**	**42.50**	**45.00**	**47.50**	**50.00**	**52.50**	**55.00**	**e_0**
Column	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**Column**
								Expectation of life at age x								
Age (x)																*Age (x)*
0	20.00	22.50	25.00	27.50	30.00	32.50	35.00	37.50	40.00	42.50	45.00	47.50	50.00	52.50	55.00	**0**
5	40.07	41.12	42.20	43.31	44.45	45.62	46.83	48.08	49.36	50.68	52.03	53.43	54.86	56.34	57.86	**5**
10	37.40	38.33	39.28	40.27	41.29	42.35	43.44	44.57	45.74	46.95	48.20	49.49	50.82	52.21	53.63	**10**
15	33.97	34.81	35.69	36.60	37.54	38.52	39.54	40.60	41.69	42.83	44.01	45.23	46.49	47.81	49.17	**15**
20	31.31	32.04	32.81	33.61	34.44	35.31	36.22	37.17	38.15	39.18	40.25	41.37	42.54	43.75	45.02	**20**
25	29.16	29.77	30.41	31.08	31.79	32.53	33.31	34.13	34.99	35.89	36.83	37.82	38.86	39.95	41.10	**25**
30	26.76	27.27	27.80	28.37	28.96	29.59	30.26	30.96	31.70	32.49	33.32	34.19	35.12	36.10	37.13	**30**
35	24.13	24.55	24.99	25.47	25.97	26.50	27.07	27.67	28.31	29.00	29.72	30.50	31.32	32.19	33.13	**35**
40	21.41	21.76	22.12	22.51	22.93	23.38	23.86	24.37	24.92	25.50	26.13	26.81	27.53	28.31	29.14	**40**
45	18.69	18.96	19.26	19.57	19.92	20.28	20.68	21.11	21.57	22.06	22.60	23.18	23.80	24.48	25.21	**45**
50	16.05	16.26	16.49	16.74	17.01	17.31	17.62	17.97	18.34	18.75	19.19	19.68	20.20	20.78	21.41	**50**
55	13.55	13.71	13.88	14.07	14.27	14.50	14.74	15.01	15.30	15.62	15.97	16.36	16.79	17.26	17.78	**55**
60	11.25	11.36	11.47	11.61	11.75	11.91	12.08	12.28	12.49	12.73	12.99	13.29	13.61	13.98	14.39	**60**
65	9.16	9.22	9.30	9.39	9.48	9.58	9.70	9.83	9.97	10.13	10.31	10.52	10.75	11.01	11.31	**65**
70	7.30	7.34	7.38	7.43	7.48	7.54	7.61	7.68	7.77	7.87	7.98	8.10	8.25	8.41	8.60	**70**
75	5.67	5.69	5.71	5.74	5.76	5.79	5.82	5.86	5.90	5.95	6.01	6.07	6.15	6.24	6.34	**75**
80	4.29	4.29	4.30	4.31	4.32	4.33	4.34	4.36	4.37	4.39	4.41	4.44	4.47	4.51	4.55	**80**
85	3.14	3.14	3.14	3.15	3.15	3.15	3.15	3.16	3.16	3.17	3.17	3.18	3.19	3.20	3.21	**85**
90	2.23	2.23	2.23	2.23	2.23	2.24	2.24	2.24	2.24	2.24	2.24	2.24	2.24	2.24	2.24	**90**
95	1.53	1.53	1.53	1.53	1.53	1.53	1.53	1.53	1.53	1.53	1.53	1.53	1.53	1.53	1.53	**95**
Column	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**Column**

ANNEX TABLE II.1. (CONTINUED)

e_0	55.00	57.50	60.00	62.50	65.00	67.50	70.00	72.50	75.00	77.50	80.00	82.50	85.00	87.50	90.00	e_0
Column	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	Column
								Expectation of life at age x								
Age (x)																*Age (x)*
0	55.00	57.50	60.00	62.50	65.00	67.50	70.00	72.50	75.00	77.50	80.00	82.50	85.00	87.50	90.00	**0**
5	57.86	59.43	61.05	62.71	64.42	66.19	68.02	69.91	71.87	73.89	75.99	78.16	80.40	82.72	85.10	**5**
10	53.63	55.11	56.64	58.22	59.86	61.56	63.32	65.15	67.06	69.03	71.09	73.22	75.44	77.74	80.11	**10**
15	49.17	50.59	52.06	53.59	55.17	56.82	58.54	60.33	62.19	64.13	66.16	68.27	70.47	72.76	75.12	**15**
20	45.02	46.34	47.72	49.16	50.66	52.24	53.88	55.60	57.40	59.29	61.27	63.35	65.52	67.78	70.13	**20**
25	41.10	42.30	43.57	44.90	46.30	47.77	49.32	50.96	52.68	54.50	56.42	58.45	60.58	62.82	65.15	**25**
30	37.13	38.23	39.39	40.61	41.91	43.29	44.75	46.31	47.96	49.71	51.57	53.55	55.64	57.85	60.17	**30**
35	33.13	34.12	35.18	36.31	37.52	38.80	40.18	41.65	43.23	44.92	46.72	48.65	50.71	52.89	55.18	**35**
40	29.14	30.03	31.00	32.03	33.14	34.34	35.63	37.02	38.52	40.13	41.88	43.76	45.77	47.93	50.20	**40**
45	25.21	26.01	26.87	27.80	28.82	29.92	31.11	32.42	33.83	35.38	37.06	38.88	40.85	42.97	45.22	**45**
50	21.41	22.10	22.85	23.68	24.59	25.58	26.68	27.88	29.21	30.66	32.27	34.02	35.94	38.02	40.25	**50**
55	17.78	18.36	19.00	19.71	20.50	21.38	22.35	23.45	24.66	26.02	27.53	29.20	31.06	33.09	35.28	**55**
60	14.39	14.85	15.37	15.95	16.61	17.35	18.20	19.15	20.24	21.47	22.87	24.44	26.21	28.17	30.32	**60**
65	11.31	11.65	12.04	12.48	13.00	13.59	14.28	15.08	16.00	17.08	18.33	19.77	21.42	23.30	25.38	**65**
70	8.60	8.83	9.09	9.39	9.76	10.19	10.70	11.31	12.04	12.92	13.98	15.24	16.74	18.49	20.48	**70**
75	6.34	6.46	6.61	6.79	7.00	7.27	7.59	8.00	8.51	9.15	9.96	10.97	12.24	13.79	15.64	**75**
80	4.55	4.60	4.66	4.74	4.84	4.97	5.13	5.33	5.61	5.98	6.49	7.18	8.11	9.35	10.95	**80**
85	3.21	3.22	3.24	3.27	3.29	3.33	3.38	3.45	3.55	3.69	3.90	4.22	4.71	5.47	6.62	**85**
90	2.24	2.25	2.25	2.25	2.26	2.26	2.27	2.28	2.30	2.33	2.37	2.44	2.56	2.80	3.26	**90**
95	1.53	1.53	1.53	1.53	1.53	1.53	1.53	1.53	1.53	1.53	1.54	1.54	1.55	1.57	1.61	**95**
Column	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**Column**

ANNEX TABLE II.2. BRASS GENERAL MODEL LIFE TABLE FAMILY VALUES OF PERSON YEARS LIVED IN FIVE -YEAR AGE GROUPS, X = 0, 5, ...,985, FOR TABLES WITH EXPECTATION OF LIFE AT BIRTH OF 20, 22.5, ..., 90 YEARS

e_0	**20.00**	**22.50**	**25.00**	**27.50**	**30.00**	**32.50**	**35.00**	**37.50**	**40.00**	**42.50**	**45.00**	**47.50**	**50.00**	**52.50**	**55.00**	**e_0**
Column	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**Column**
								Person-years lived in age group								
Age (x)																*Age (x)*
0	2.6087	2.8165	3.0102	3.1908	3.3589	3.5155	3.6613	3.7968	3.9226	4.0393	4.1474	4.2473	4.3395	4.4222	4.4974	**0**
5	2.1003	2.3224	2.5343	2.7360	2.9277	3.1096	3.2816	3.4442	3.5973	3.7413	3.8764	4.0027	4.1204	4.2299	4.3313	**5**
10	2.0025	2.2223	2.4334	2.6357	2.8292	3.0138	3.1895	3.3564	3.5145	3.6639	3.8046	3.9368	4.0605	4.1760	4.2832	**10**
15	1.8941	2.1103	2.3196	2.5217	2.7164	2.9034	3.0827	3.2540	3.4173	3.5724	3.7194	3.8581	3.9886	4.1110	4.2251	**15**
20	1.7353	1.9447	2.1498	2.3500	2.5450	2.7344	2.9177	3.0946	3.2648	3.4280	3.5839	3.7323	3.8730	4.0058	4.1306	**20**
25	1.5808	1.7818	1.9808	2.1773	2.3709	2.5608	2.7466	2.9277	3.1037	3.2741	3.4383	3.5960	3.7467	3.8902	4.0259	**25**
30	1.4418	1.6335	1.8253	2.0168	2.2072	2.3960	2.5825	2.7661	2.9462	3.1221	3.2932	3.4590	3.6188	3.7721	3.9183	**30**
35	1.3096	1.4911	1.6745	1.8594	2.0451	2.2311	2.4167	2.6012	2.7839	2.9641	3.1411	3.3141	3.4823	3.6451	3.8016	**35**
40	1.1774	1.3473	1.5207	1.6973	1.8765	2.0578	2.2406	2.4243	2.6081	2.7912	2.9729	3.1523	3.3285	3.5007	3.6678	**40**
45	1.0401	1.1963	1.3575	1.5234	1.6936	1.8678	2.0455	2.2260	2.4088	2.5931	2.7781	2.9629	3.1465	3.3279	3.5060	**45**
50	0.8937	1.0337	1.1797	1.3318	1.4898	1.6535	1.8225	1.9967	2.1753	2.3580	2.5439	2.7322	2.9219	3.1121	3.3014	**50**
55	0.7377	0.8584	0.9858	1.1201	1.2616	1.4102	1.5660	1.7289	1.8987	2.0751	2.2576	2.4458	2.6387	2.8355	3.0349	**55**
60	0.5750	0.6732	0.7783	0.8906	1.0106	1.1386	1.2750	1.4200	1.5740	1.7370	1.9091	2.0902	2.2799	2.4779	2.6831	**60**
65	0.4132	0.4868	0.5666	0.6530	0.7468	0.8484	0.9586	1.0780	1.2073	1.3471	1.4982	1.6610	1.8361	2.0240	2.2245	**65**
70	0.2647	0.3136	0.3673	0.4262	0.4911	0.5625	0.6412	0.7281	0.8241	0.9304	1.0480	1.1782	1.3223	1.4820	1.6583	**70**
75	0.1435	0.1709	0.2011	0.2348	0.2722	0.3140	0.3608	0.4133	0.4724	0.5391	0.6146	0.7004	0.7982	0.9100	1.0381	**75**
80	0.0608	0.0726	0.0857	0.1005	0.1171	0.1357	0.1568	0.1808	0.2082	0.2396	0.2757	0.3177	0.3666	0.4241	0.4920	**80**
85	0.0177	0.0211	0.0250	0.0294	0.0343	0.0399	0.0462	0.0535	0.0619	0.0715	0.0828	0.0960	0.1117	0.1303	0.1528	**85**
90	0.0029	0.0035	0.0041	0.0048	0.0056	0.0066	0.0076	0.0088	0.0102	0.0118	0.0137	0.0160	0.0186	0.0218	0.0257	**90**
95	0.0002	0.0002	0.0003	0.0003	0.0004	0.0004	0.0005	0.0006	0.0007	0.0008	0.0009	0.0011	0.0013	0.0015	0.0018	**95**
Column	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**Column**

ANNEX TABLE II.2. (CONTINUED)

e_0	55.00	57.50	60.00	62.50	65.00	67.50	70.00	72.50	75.00	77.50	80.00	82.50	85.00	87.50	90.00	e_0
Column	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**Column**
							Person-years lived in age group									
Age (x)																*Age (x)*
0	4.4974	4.5670	4.6310	4.6894	4.7424	4.7900	4.8324	4.8695	4.9015	4.9285	4.9506	4.9679	4.9809	4.9898	4.9953	**0**
5	4.3313	4.4247	4.5102	4.5882	4.6587	4.7220	4.7782	4.8274	4.8697	4.9054	4.9346	4.9576	4.9747	4.9865	4.9938	**5**
10	4.2832	4.3824	4.4735	4.5567	4.6322	4.7001	4.7605	4.8134	4.8591	4.8977	4.9292	4.9541	4.9726	4.9854	4.9933	**10**
15	4.2251	4.3310	4.4287	4.5183	4.5998	4.6732	4.7387	4.7962	4.8460	4.8881	4.9226	4.9497	4.9700	4.9840	4.9927	**15**
20	4.1306	4.2471	4.3552	4.4549	4.5460	4.6285	4.7024	4.7675	4.8241	4.8719	4.9113	4.9424	4.9656	4.9817	4.9916	**20**
25	4.0259	4.1536	4.2728	4.3834	4.4851	4.5776	4.6609	4.7346	4.7988	4.8533	4.8983	4.9339	4.9605	4.9789	4.9903	**25**
30	3.9183	4.0568	4.1870	4.3085	4.4209	4.5237	4.6167	4.6994	4.7717	4.8334	4.8843	4.9247	4.9550	4.9760	4.9890	**30**
35	3.8016	3.9510	4.0925	4.2255	4.3494	4.4634	4.5670	4.6597	4.7410	4.8106	4.8683	4.9142	4.9487	4.9726	4.9874	**35**
40	3.6678	3.8287	3.9825	4.1282	4.2649	4.3916	4.5075	4.6119	4.7039	4.7830	4.8489	4.9014	4.9410	4.9684	4.9855	**40**
45	3.5060	3.6794	3.8469	4.0071	4.1589	4.3009	4.4319	4.5506	4.6561	4.7473	4.8236	4.8847	4.9309	4.9630	4.9830	**45**
50	3.3014	3.4883	3.6712	3.8485	4.0186	4.1795	4.3297	4.4672	4.5904	4.6979	4.7884	4.8614	4.9167	4.9554	4.9795	**50**
55	3.0349	3.2353	3.4350	3.6320	3.8241	4.0091	4.1843	4.3471	4.4949	4.6254	4.7364	4.8266	4.8955	4.9439	4.9742	**55**
60	2.6831	2.8943	3.1098	3.3276	3.5452	3.7596	3.9674	4.1647	4.3476	4.5119	4.6540	4.7711	4.8614	4.9253	4.9655	**60**
65	2.2245	2.4372	2.6614	2.8955	3.1374	3.3842	3.6315	3.8745	4.1069	4.3222	4.5135	4.6746	4.8014	4.8924	4.9501	**65**
70	1.6583	1.8526	2.0657	2.2983	2.5501	2.8197	3.1042	3.3985	3.6953	3.9845	4.2541	4.4911	4.6843	4.8269	4.9192	**70**
75	1.0381	1.1852	1.3540	1.5481	1.7705	2.0246	2.3123	2.6339	2.9861	3.3604	3.7411	4.1051	4.4250	4.6761	4.8460	**75**
80	0.4920	0.5728	0.6696	0.7865	0.9286	1.1023	1.3157	1.5781	1.8994	2.2876	2.7440	3.2548	3.7820	4.2627	4.6304	**80**
85	0.1528	0.1803	0.2140	0.2563	0.3098	0.3787	0.4691	0.5899	0.7541	0.9812	1.2983	1.7394	2.3338	3.0724	3.8534	**85**
90	0.0257	0.0304	0.0364	0.0439	0.0537	0.0665	0.0840	0.1082	0.1432	0.1957	0.2780	0.4136	0.6472	1.0614	1.7762	**90**
95	0.0018	0.0021	0.0025	0.0030	0.0037	0.0046	0.0059	0.0076	0.0102	0.0141	0.0204	0.0315	0.0526	0.0973	0.2037	**95**
Column	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**Column**

ANNEX TABLE II.3. BRASS GENERAL MODEL LIFE TABLE FAMILY VALUES OF LIFE TABLE SURVIVORS, X = 0, 5, ...,95, FOR TABLES WITH EXPECTATION OF LIFE AT BIRTH OF 20, 22.5, ..., 90 YEARS

e_0	**20.00**	**22.50**	**25.00**	**27.50**	**30.00**	**32.50**	**35.00**	**37.50**	**40.00**	**42.50**	**45.00**	**47.50**	**50.00**	**52.50**	**55.00**	**e_0**
Column	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**Column**
								Life table survivors at age x								
Age (x)																*Age (x)*
0	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	**0**
5	43,402	47,872	52,114	56,134	59,937	63,529	66,914	70,099	73,090	75,893	78,513	80,956	83,227	85,333	87,279	**5**
10	40,882	45,300	49,531	53,574	57,431	61,101	64,586	67,888	71,009	73,951	76,717	79,310	81,733	83,991	86,086	**10**
15	39,120	43,488	47,697	51,745	55,627	59,343	62,890	66,267	69,475	72,512	75,380	78,080	80,612	82,980	85,183	**15**
20	36,395	40,662	44,814	48,846	52,749	56,517	60,145	63,627	66,961	70,141	73,165	76,030	78,734	81,278	83,659	**20**
25	33,122	37,230	41,276	45,250	49,141	52,941	56,638	60,224	63,692	67,032	70,237	73,301	76,216	78,981	81,587	**25**
30	30,186	34,115	38,028	41,912	45,757	49,549	53,278	56,931	60,497	63,964	67,322	70,560	73,668	76,638	79,459	**30**
35	27,503	31,239	34,997	38,765	42,533	46,286	50,012	53,699	57,332	60,898	64,382	67,772	71,053	74,215	77,242	**35**
40	24,876	28,395	31,969	35,590	39,247	42,927	46,618	50,306	53,977	57,616	61,206	64,733	68,178	71,528	74,763	**40**
45	22,203	25,472	28,828	32,261	35,766	39,331	42,945	46,596	50,270	53,952	57,625	61,271	64,871	68,408	71,858	**45**
50	19,371	22,343	25,427	28,619	31,914	35,305	38,786	42,346	45,974	49,655	53,375	57,114	60,854	64,575	68,249	**50**
55	16,347	18,964	21,711	24,590	27,601	30,741	34,009	37,398	40,903	44,513	48,216	51,997	55,838	59,719	63,614	**55**
60	13,142	15,340	17,677	20,159	22,791	25,578	28,522	31,627	34,892	38,315	41,892	45,614	49,470	53,444	57,514	**60**
65	9,862	11,585	13,440	15,439	17,590	19,905	22,393	25,065	27,929	30,994	34,267	37,752	41,449	45,358	49,466	**65**
70	6,710	7,931	9,263	10,717	12,306	14,044	15,945	18,026	20,304	22,797	25,525	28,506	31,760	35,306	39,156	**70**
75	3,966	4,712	5,536	6,447	7,455	8,575	9,821	11,209	12,760	14,495	16,441	18,626	21,085	23,856	26,978	**75**
80	1,901	2,268	2,677	3,133	3,643	4,217	4,863	5,594	6,424	7,370	8,454	9,701	11,143	12,820	14,777	**80**
85	661	790	935	1,098	1,281	1,488	1,724	1,993	2,302	2,658	3,072	3,556	4,126	4,804	5,617	**85**
90	138	165	196	230	269	313	364	421	488	565	655	761	887	1,039	1,223	**90**
95	13	15	18	22	25	29	34	40	46	53	62	72	84	98	116	**95**
Column	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**Column**

ANNEX TABLE II.3. (CONTINUED)

e_0	55.00	57.50	60.00	62.50	65.00	67.50	70.00	72.50	75.00	77.50	80.00	82.50	85.00	87.50	90.00	e_0
Column	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**Column**
								Life table survivors at age x								
Age (x)																*Age (x)*
0	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	**0**
5	87,279	89,067	90,702	92,189	93,532	94,735	95,802	96,734	97,537	98,213	98,765	99,200	99,523	99,746	99,883	**5**
10	86,086	88,018	89,792	91,411	92,878	94,195	95,365	96,392	97,276	98,022	98,633	99,113	99,471	99,718	99,871	**10**
15	85,183	87,222	89,100	90,817	92,377	93,780	95,030	96,127	97,075	97,874	98,530	99,046	99,431	99,697	99,861	**15**
20	83,659	85,873	87,921	89,803	91,519	93,068	94,453	95,672	96,727	97,619	98,352	98,930	99,362	99,660	99,844	**20**
25	81,587	84,029	86,302	88,403	90,329	92,077	93,646	95,033	96,237	97,260	98,101	98,766	99,263	99,607	99,820	**25**
30	79,459	82,121	84,616	86,936	89,075	91,028	92,788	94,351	95,713	96,873	97,831	98,589	99,157	99,550	99,793	**30**
35	77,242	80,119	82,835	85,377	87,736	89,901	91,862	93,612	95,144	96,452	97,535	98,395	99,040	99,487	99,765	**35**
40	74,763	77,864	80,814	83,597	86,196	88,597	90,786	92,749	94,475	95,956	97,186	98,166	98,902	99,413	99,730	**40**
45	71,858	75,197	78,404	81,456	84,331	87,008	89,465	91,684	93,646	95,338	96,750	97,878	98,729	99,320	99,687	**45**
50	68,249	71,849	75,346	78,713	81,919	84,934	87,728	90,273	92,541	94,510	96,163	97,490	98,493	99,193	99,629	**50**
55	63,614	67,489	71,312	75,047	78,655	82,096	85,325	88,302	90,984	93,335	95,323	96,931	98,154	99,010	99,544	**55**
60	57,514	61,646	65,809	69,958	74,048	78,024	81,825	85,390	88,654	91,556	94,041	96,072	97,628	98,725	99,412	**60**
65	49,466	53,752	58,190	62,741	67,355	71,967	76,501	80,866	84,963	88,688	91,944	94,648	96,750	98,245	99,189	**65**
70	39,156	43,314	47,781	52,541	57,563	62,795	68,156	73,535	78,789	83,751	88,239	92,080	95,139	97,355	98,771	**70**
75	26,978	30,491	34,438	38,858	43,779	49,211	55,131	61,466	68,076	74,741	81,158	86,970	91,827	95,481	97,879	**75**
80	14,777	17,073	19,778	22,976	26,766	31,261	36,576	42,813	50,022	58,138	66,905	75,802	84,059	90,840	95,586	**80**
85	5,617	6,600	7,801	9,287	11,146	13,501	16,523	20,442	25,568	32,280	40,963	51,811	64,410	77,292	88,141	**85**
90	1,223	1,449	1,730	2,085	2,543	3,145	3,954	5,073	6,669	9,020	12,612	18,276	27,349	41,452	60,723	**90**
95	116	137	164	199	243	303	383	497	663	918	1,331	2,047	3,399	6,206	12,625	**95**
Column	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**Column**

ANNEX TABLE II.4. BRASS GENERAL MODEL LIFE TABLE FAMILY VALUES OF LIFE TABLE DEATHS, X = 0, 5, ..., 95, FOR TABLES WITH EXPECTATION OF LIFE AT BIRTH OF 20, 22.5, ..., 90 YEARS

e_0	**20.00**	**22.50**	**25.00**	**27.50**	**30.00**	**32.50**	**35.00**	**37.50**	**40.00**	**42.50**	**45.00**	**47.50**	**50.00**	**52.50**	**55.00**	**e_0**
Column	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**Column**
								Life table deaths at age x								
Age (x)																*Age (x)*
0	216,957	185,080	159,077	137,477	119,272	103,743	90,368	78,753	68,602	59,681	51,809	44,839	38,653	33,166	28,286	**0**
5	12,001	11,076	10,195	9,357	8,561	7,807	7,094	6,420	5,786	5,191	4,632	4,111	3,624	3,172	2,754	**5**
10	8,796	8,152	7,534	6,942	6,375	5,833	5,318	4,829	4,365	3,926	3,513	3,125	2,762	2,423	2,108	**10**
15	14,388	13,393	12,429	11,497	10,598	9,734	8,905	8,113	7,357	6,638	5,957	5,313	4,707	4,139	3,608	**15**
20	18,860	17,647	16,460	15,301	14,173	13,078	12,019	10,996	10,013	9,070	8,170	7,313	6,501	5,734	5,014	**20**
25	18,575	17,481	16,398	15,329	14,277	13,244	12,234	11,250	10,294	9,369	8,477	7,620	6,802	6,023	5,286	**25**
30	18,610	17,607	16,604	15,603	14,607	13,619	12,644	11,684	10,742	9,822	8,927	8,060	7,225	6,423	5,658	**30**
35	20,062	19,076	18,080	17,075	16,066	15,056	14,047	13,045	12,052	11,073	10,111	9,171	8,256	7,371	6,520	**35**
40	22,696	21,689	20,661	19,614	18,551	17,476	16,392	15,302	14,212	13,125	12,047	10,982	9,936	8,914	7,921	**40**
45	27,228	26,157	25,051	23,913	22,745	21,550	20,330	19,091	17,836	16,570	15,299	14,029	12,767	11,519	10,294	**45**
50	33,844	32,695	31,495	30,246	28,948	27,603	26,214	24,782	23,313	21,810	20,280	18,730	17,168	15,600	14,040	**50**
55	43,438	42,215	40,923	39,560	38,124	36,616	35,035	33,381	31,657	29,865	28,009	26,097	24,134	22,131	20,100	**55**
60	57,041	55,783	54,437	52,997	51,460	49,820	48,072	46,212	44,238	42,147	39,939	37,614	35,178	32,633	29,994	**60**
65	76,266	75,051	73,734	72,306	70,758	69,080	67,262	65,291	63,157	60,848	58,353	55,663	52,770	49,664	46,348	**65**
70	103,709	102,644	101,476	100,193	98,783	97,231	95,520	93,631	91,545	89,238	86,685	83,859	80,730	77,266	73,438	**70**
75	143,838	143,030	142,135	141,141	140,035	138,802	137,422	135,875	134,134	132,170	129,948	127,426	124,555	121,273	117,517	**75**
80	204,165	203,670	203,117	202,499	201,805	201,024	200,141	199,138	197,995	196,685	195,176	193,430	191,396	189,010	186,195	**80**
85	295,995	295,776	295,530	295,254	294,943	294,591	294,190	293,731	293,204	292,595	291,886	291,054	290,072	288,900	287,490	**85**
90	433,598	433,539	433,474	433,399	433,315	433,220	433,111	432,987	432,843	432,676	432,481	432,251	431,978	431,650	431,251	**90**
95	654,833	654,826	654,819	654,810	654,800	654,789	654,776	654,762	654,745	654,725	654,702	654,675	654,643	654,604	654,557	**95**
Column	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**Column**

ANNEX TABLE II.4. (CONTINUED)

e_0	55.00	57.50	60.00	62.50	65.00	67.50	70.00	72.50	75.00	77.50	80.00	82.50	85.00	87.50	90.00	e_0
Column	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**Column**
							Life table deaths at age x									
Age (x)																*Age (x)*
0	28,286	23,940	20,078	16,657	13,638	10,991	8,688	6,706	5,025	3,626	2,494	1,611	958	509	233	**0**
5	2,754	2,369	2,017	1,695	1,405	1,144	913	710	536	389	269	174	104	55	25	**5**
10	2,108	1,816	1,548	1,304	1,082	882	705	549	414	301	208	135	81	43	20	**10**
15	3,608	3,116	2,662	2,245	1,866	1,524	1,218	950	718	522	361	234	140	75	34	**15**
20	5,014	4,342	3,718	3,143	2,618	2,142	1,716	1,340	1,014	738	511	332	198	106	49	**20**
25	5,286	4,593	3,946	3,346	2,794	2,292	1,840	1,440	1,092	796	552	359	214	114	52	**25**
30	5,658	4,934	4,253	3,618	3,030	2,492	2,005	1,572	1,194	872	605	394	235	126	58	**30**
35	6,520	5,708	4,938	4,214	3,541	2,920	2,357	1,852	1,410	1,031	717	467	279	149	69	**35**
40	7,921	6,965	6,051	5,185	4,373	3,619	2,930	2,309	1,762	1,291	899	587	351	188	86	**40**
45	10,294	9,101	7,949	6,845	5,800	4,821	3,919	3,100	2,373	1,744	1,218	796	477	255	117	**45**
50	14,040	12,499	10,990	9,526	8,122	6,791	5,550	4,412	3,393	2,503	1,753	1,149	690	370	170	**50**
55	20,100	18,058	16,021	14,010	12,048	10,157	8,365	6,698	5,184	3,846	2,707	1,781	1,073	576	266	**55**
60	29,994	27,275	24,499	21,690	18,880	16,109	13,420	10,863	8,491	6,356	4,508	2,984	1,807	974	450	**60**
65	46,348	42,826	39,112	35,228	31,208	27,103	22,979	18,922	15,031	11,421	8,207	5,493	3,356	1,820	844	**65**
70	73,438	69,220	64,590	59,534	54,054	48,176	41,961	35,513	28,991	22,613	16,646	11,379	7,070	3,882	1,814	**70**
75	117,517	113,212	108,269	102,595	96,091	88,663	80,244	70,817	60,460	49,407	38,097	27,203	17,554	9,925	4,731	**75**
80	186,195	182,853	178,858	174,048	168,219	161,113	152,416	141,758	128,745	113,039	94,543	73,711	51,953	31,783	16,078	**80**
85	287,490	285,776	283,668	281,041	277,722	273,464	267,911	260,546	250,617	237,047	218,365	192,793	158,800	116,654	71,153	**85**
90	431,251	430,760	430,149	429,374	428,375	427,058	425,279	422,806	419,257	413,967	405,749	392,427	370,042	332,068	270,791	**90**
95	654,557	654,499	654,426	654,333	654,213	654,053	653,836	653,529	653,080	652,392	651,274	649,330	645,643	637,900	619,760	**95**
Column	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**Column**

ANNEX TABLE II.5. STATIONARY POPULATION AGE DISTRIBUTIONS AND INTERCENSAL DEATHS: SYNTHETIC DATA DERIVED FROM BRASS GENERAL MODEL LIFE TABLE FAMILY

Age group	*Stationary population L_x*	*Probability of survival l_x*	*Age distribution time t*	*Age distribution time t+10*	*Cumulative intercensal deaths*	*Intercensal deaths*
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
0-4	4.8695	1.00000	4,869,500	4,869,500	1,000,000	326,600
5-9	4.8274	0.96734	4,827,400	4,827,400	967,340	34,200
10-14	4.8134	0.96392	4,813,400	4,813,400	963,920	26,500
15-19	4.7962	0.96127	4,796,200	4,796,200	961,270	45,500
20-24	4.7675	0.95672	4,767,500	4,767,500	956,720	63,900
25-29	4.7346	0.95033	4,734,600	4,734,600	950,330	68,200
30-34	4.6994	0.94351	4,699,400	4,699,400	943,510	73,900
35-39	4.6597	0.93612	4,659,700	4,659,700	936,120	86,300
40-44	4.6119	0.92749	4,611,900	4,611,900	927,490	106,500
45-49	4.5506	0.91684	4,550,600	4,550,600	916,840	141,100
50-54	4.4672	0.90273	4,467,200	4,467,200	902,730	197,100
55-59	4.3471	0.88302	4,347,100	4,347,100	883,020	291,200
60-64	4.1647	0.85390	4,164,700	4,164,700	853,900	452,400
65-69	3.8745	0.80866	3,874,500	3,874,500	808,660	733,100
70-74	3.3985	0.73535	3,398,500	3,398,500	735,350	1,206,900
75-79	2.6339	0.61466	2,633,900	2,633,900	614,660	1,865,300
80-84	1.5781	0.42813	1,578,100	1,578,100	428,130	2,237,100
85-89	0.5899	0.20442	589,900	589,900	204,420	1,536,900
90-94	0.1082	0.05073	108,200	108,200	50,730	457,600
95-99	0.0076	0.00497	7,600	7,600	4,970	49,700
100+	0.0000	0.00000	0	0	0	0
Total	72.4999	15.0101	72,499,900	72,499,900	NA	10,000,000

Source: Stationary population (column 2) extracted from Annex table II.2, column 22

Procedure

Columns 1 to 3: Enter age, life table ${}_5L_x$ and l_x values (in this case from $e_0 = 72.5$ column of annex table II.2 and II.3).

Columns 4 and 5: Multiply ${}_5L_x$ value for each age group by the radix of the stationary population, defined as the number of persons entering the population each year, to obtain the number of persons in the age group. the radix is taken here to be one million.

Column 6: Calculate the number of persons reaching exact age $x = 0, 5, \ldots$, during the intercensal period as the proportion of the cohort surviving to exact age x.

Column 7: Because the population is stationary, the number reaching exact age x during the period in column 6 equals the number dying at an age greater than or equal to x during the period. Thus deaths in the age group x to $x+5$ may be calculated by subtracting the number of persons dying at age $x+5$ and over from the number dying at age x and over.

.

ANNEX TABLE II.6. STABLE POPULATION AGE DISTRIBUTIONS AND INTERCENSAL DEATHS: SYNTHETIC DATA DERIVED FROM BRASS GENERAL MODEL LIFE TABLE FAMILY

Age group	*Age limit* x	*Stable population* ${}_5L_x$	*Probability of survival* l_x	*Age distribution at time t* $N_1(x,5)$	*Intercensal deaths A* $D(x,A)$	*Persons reaching age x* $N(x)$	*Intercensal deaths B* $D(x,B)$	*Age distribution at time t+5* $N_2(x,5)$	$N_1(x,5)$ - $[D(x,A)+$ $+D(x+5,B)]$	*Total intercensal deaths* $D(x,5)$
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*
0-4	0	4.6310	1.00000	1,700,208	35,208	2,132,757	157,397	1,975,360	NA	192,605
5-9	5	4.5102	0.90702	1,425,210	6,510	1,665,000	9,142	1,655,858	1,655,858	15,652
10-14	10	4.4735	0.89792	1,216,708	5,032	1,418,700	5,087	1,413,613	1,413,613	10,119
15-19	15	4.4287	0.89100	1,036,742	7,643	1,211,676	7,154	1,204,522	1,204,522	14,797
20-24	20	4.3552	0.87921	877,523	8,080	1,029,099	9,563	1,019,536	1,019,536	17,643
25-29	25	4.2728	0.86302	741,001	7,284	869,443	8,523	860,920	860,920	15,807
30-34	30	4.1870	0.84616	624,978	6,754	733,717	7,596	726,121	726,121	14,350
35-39	35	4.0925	0.82835	525,783	6,655	618,224	7,351	610,873	610,873	14,006
40-44	40	3.9825	0.80814	440,382	6,889	519,128	7,477	511,651	511,651	14,366
45-49	45	3.8469	0.78404	366,134	7,576	433,493	8,106	425,387	425,387	15,682
50-54	50	3.6712	0.75346	300,742	8,651	358,558	9,146	349,412	349,412	17,797
55-59	55	3.4350	0.71312	242,197	10,192	292,091	10,698	281,393	281,393	20,890
60-64	60	3.1098	0.65809	188,725	12,156	232,005	12,738	219,267	219,267	24,894
65-69	65	2.6614	0.58190	139,015	14,226	176,569	15,057	161,512	161,512	29,283
70-74	70	2.0657	0.47781	92,870	15,457	124,789	16,889	107,900	107,900	32,346
75-79	75	1.3540	0.34438	52,394	14,128	77,413	16,540	60,873	60,873	30,668
80-84	80	0.6696	0.19778	22,302	9,311	38,266	12,355	25,911	25,911	21,666
85-89	85	0.2140	0.07801	6,135	3,655	12,991	5,863	7,128	7,128	9,518
90-94	90	0.0364	0.01730	898	696	2,480	1,437	1,043	1,043	2,133
95-99	95	0.0025	0.00164	53	53	202	140	62	62	193
100+	100	0.0000	0.00000	0	0	0	0	0	0	NA
Total	NA	59.9999	NA	10,000,000	NA	NA	NA	11,618,342	NA	514,415

Source: Stable population (column 3) extracted from Annex table II.2, column 17.

Procedure

Columns 1 and 4: Enter life table ${}_5L_x$ and l_x values for each age (in this example from $e_0 = 60$ column of annex table II.2 and annex table II.3).

Column 5: Calculate stable population age distribution as

$$N(x,5) = B\exp(-rx)\,{}_5L_x$$

where $N(x,5)$ denotes the number of persons aged x to $x+5$, B denotes the number of births during the five years preceding time t, and r denotes the population growth rate. In this example B has been chosen to give a total population of 10 million, *i.e.*, $B = 1,835,681$.

Column 9: (*This column should be filled in before columns 6-8*). Calculate the number of persons in each five year age group at time t+5 as

$$N_2(x,5) = N_1(x,5)\exp(5r)$$

Column 6: Calculate deaths before exact age $x+5$ to persons aged x to $x+5$ at time t as

$$D(x,A) = N(x,5)[1 - 5l_{x+5} / {}_5L_x]$$

for $x = 0, 5, ...$

Column 7: Calculate the number of persons reaching exact age $x+5$ between times t and t+5 as

$$N(x+5) = N(x,5) - D(x,A)$$

for $x = 0, 5, ...$ Calculate $N(0)$, the number of births between times t and t+5, as $\exp(5r)$ times the number of births in the preceding five year period.

Column 8. Calculate the number of deaths of persons reaching exact age x between times t and t+5 as

$$D(x,B) = N(x) - N(x,5),$$

for $x = 0, 5, ...$

Column 10: Calculate

$$N_1(x,5) - [D(x,A)+D(x+5,B)]$$

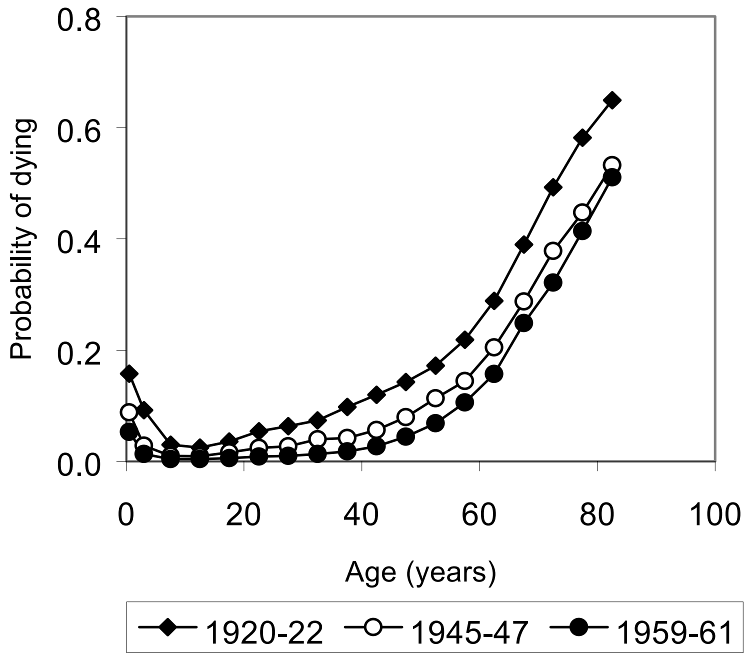
and check that this equals $N_2(x+5,5)$.

Column 11: Calculate the number of deaths of persons aged x to $x+5$ between times t and t+5 as

$$D(x,5) = D(x,A) + D(x,B)$$

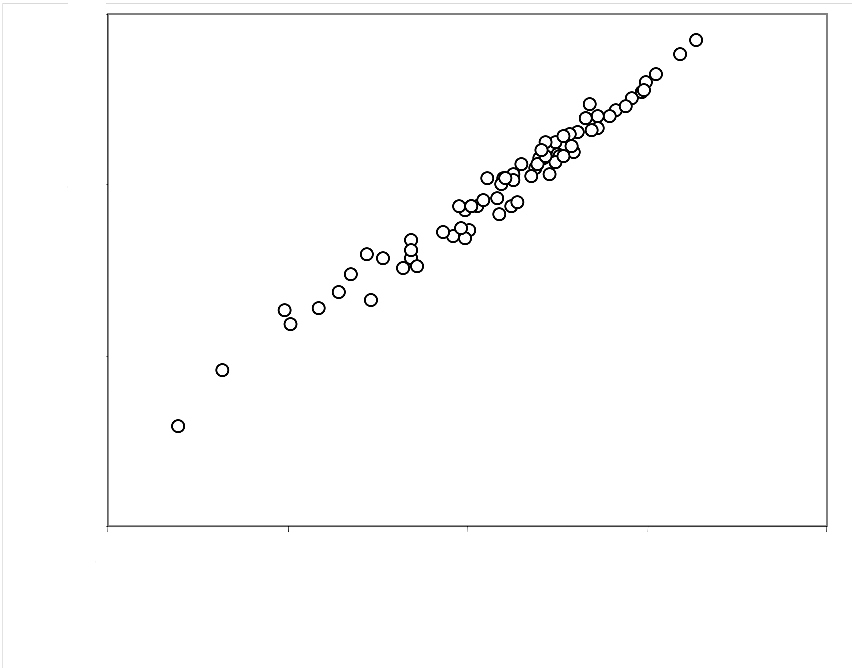
for $x = 0, 5, ...$

Annex Figure II.1. Probabilities of dying, by age, Trinidad and Tobago, males, 1920-1922, 1945-1947, and 1959-1961



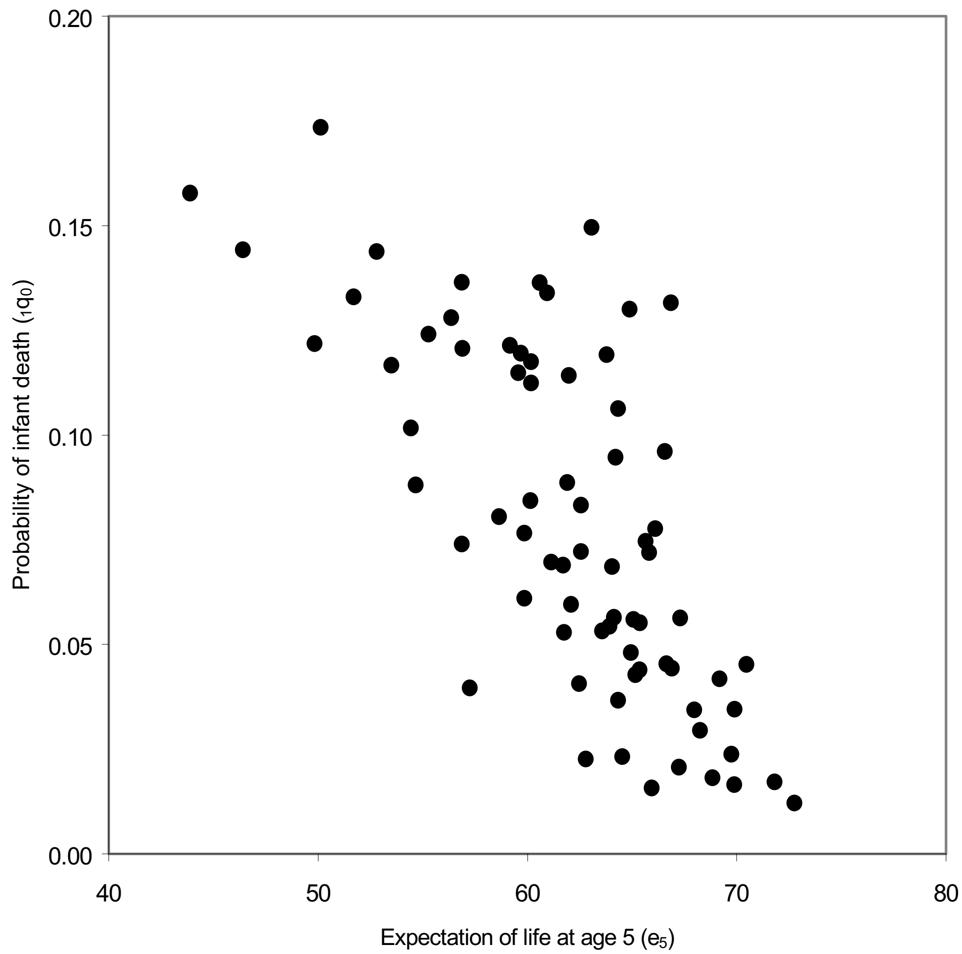
Source: *Model Life Tables for Developing Countries* (United Nations, New York, 1982, annex 5, pp. 346-349).

Annex Figure II.2. Relationship between male and female e_5 and e_{30} for 72 observed life tables



Source: *Model Life Tables for Developing Countries* (United Nations, New York, 1982, annex 5, pp. 285-351).

Annex Figure II.3. Relationship between probability of infant death ($_1q_0$) and the expectation of life at age 5 (e_5) for 72 observed life tables



Source: *Model Life Tables for Developing Countries* (United Nations, New York, 1982, annex 5, pp. 285-351).

ANNEX III
Robust straight line fitting

The least squares method is by far the most familiar approach for fitting straight lines to pairs of points (x_i,y_i) to estimate values for the intercept a and slope b of the equation

$$y_i = a + bx_i. \qquad (1)$$

A disadvantage of least squares fitting, however, is that outlying observations, which are frequently encountered in the demographic applications discussed in this manual, exert an inordinate influence on the estimates of a and b and will often give very poor results.

One approach to this problem is to plot the points to be fitted, identify outlying points, and to exclude them from the calculations. However, while extreme outliers are easy to identify, there will often be points whose status is doubtful.

A simpler approach is to utilize some form of robust fitting that is less sensitive to outliers than least squares. While many such methods are available, most of them are computationally intensive and relatively complicated to implement. Many statistical computing programs incorporate one or more robust line-fitting procedures, but these programs are generally more expensive and less widely available than spreadsheet programs.

Since all other calculations required for the methods discussed here can be implemented using a spreadsheet program only, it is desirable to have a robust method of straight line-fitting that may be simply implemented with a spreadsheet. This annex describes such a method. It was proposed originally by Nair and Srivastava (1942-1944). The modified procedure described here is due to Tukey 1977 and McNeill 1977.

The procedure will be illustrated using the result of the application of the general growth balance method to Zimbabwe females, 1982-1992. The (x,y) points to which a line is to be fitted are shown in columns 2 and 3 of annex table III-1.

The fitted line is determined in two stages. To estimate the slope, the data points are first divided into three groups: those with lower third of x values, those with x values in the middle third, and those with x values in the upper third. In case the number of points is not evenly divisible by three, the number of points in the upper and lower groups will be taken to be the next higher integer from the number of points divided by three. For the data points shown in annex table III-1, the points in the lower third are those with index numbers (column 1) from 1-5 and the points in the upper third are those with index numbers 10-14.

The slope is estimated as the slope of the line connecting the points (x_l,y_l) and (x_u,y_u), where x_l is the median of the x values in the lower group, y_l the median of the corresponding y values, x_u is the median of the x values in the upper group, and y_u the median of the corresponding y values. This calculation is shown with the procedures for annex table III.1. The estimate of the slope b in this case is

$$(0.02646 – 0.00854)/0.01067 – 0.00263) = 2.229.$$

The intercept of the line should be chosen so that the fitted line divides the data points into two equal groups. To determine this value we first calculate the intercept of the line with the estimated slope that passes through each data point, *i.e.*,

$$a_i = y_i - bx_i \qquad (2)$$

for $i = 1, 2,...,14$. These values are shown in column 4 of annex table III-1. The estimated intercept a is calculated as the median of these values,

$$a = median(a_1, a_2, ..., a_n) \qquad (3)$$

Whatever method of fitting is used, it is important to scrutinize the fit to see whether it is reasonable. Generally this should involve making two plots, one of the data points and fitted line, and another of the residuals. These are shown in annex figures III.1 and III.2.

It is useful to have indicators of the possible error of the estimated slope and intercept, as a reflection of the scatter of the observed data points about the fitted line. A useful error indicator for the intercept a is one half the interquartile range of the slopes in column 4, in this case, 0.00094. It is useful also to express this as a per cent of the estimated intercept, *100*0.00094/0.00268 = 35* per cent.

To obtain an error indicator for the slope we compute the slope of the line passing through each data point and the intercept of the fitted line with the *y* axis. These values are shown in column 5 of annex table III-1. Their median will generally be very close to the estimated slope. One half the interquartile range of these values may be used as the error indicator for the slope. Again, it is useful to express this also as a percentage of the estimated slope, *100*0.185/2.227 = 8.3* per cent.

ANNEX TABLE III.1: ROBUST STRAIGHT LINE FITTING

Index	*x-point*	*y-point*	*Intercepts y-bx*	*Slopes*	*y-fitted a+bx*	*Residuals y-(a+bx)*	*Per cent deviation*
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
1	0.00182	0.00325	-0.00082	0.313	0.00675	-0.00350	-107.5
2	0.00217	0.00638	0.00155	1.707	0.00752	-0.00113	-17.8
3	0.00263	0.00854	0.00267	2.224	0.00855	-0.00001	-0.2
4	0.00315	0.01102	0.00400	2.647	0.00971	0.00132	11.9
5	0.00374	0.01231	0.00398	2.576	0.01101	0.00130	10.5
6	0.00432	0.01240	0.00277	2.250	0.01231	0.00009	0.7
7	0.00500	0.01349	0.00236	2.164	0.01382	-0.00032	-2.4
8	0.00583	0.01549	0.00250	2.198	0.01567	-0.00018	-1.2
9	0.00675	0.01535	0.00031	1.878	0.01772	-0.00237	-15.4
10	0.00789	0.02141	0.00383	2.374	0.02026	0.00115	5.4
11	0.00933	0.02276	0.00196	2.151	0.02349	-0.00072	-3.2
12	0.01067	0.02646	0.00269	2.230	0.02645	0.00001	0.0
13	0.01186	0.03398	0.00754	2.639	0.02912	0.00486	14.3
14	0.01294	0.04451	0.01567	3.233	0.03152	0.01299	29.2
		Median	0.00268	2.227			
		0.5*Interquartile Range	0.00094	0.185			
		Per cent	35.1	8.3			

Source: x-point and y-point estimates from table II.7.

Procedure

Columns 1-3: Enter *x* and *y* values of data points to which the line is to be fitted, ordered by *x* value, lowest to highest. The values in this example are from table II-7.

To calculate the slope

Step 1: Identify the data points with *x* values in the lower third and those with *x* values in the upper third. If the number of data points does not divide evenly by three, as in this case, choose the first and last *n* values, where *n* is the next higher integer from the number of data points divided by three. In this case the upper and lower groups of data points are those with index numbers 1-5 and 10-14, respectively.

Step 2: Determine the points (x_l,y_l) and (x_u,y_u), where x_l is the median of the *x* values in the lower group, y_l the median of the corresponding *y* values, x_u is the median of the *x* values in the upper group, and y_u the median of the corresponding *y* values. In this case the point (x_l,y_l) is (0.00263, 0.00854) and the point (x_u,y_u) is (0.01067,0.02646).

Step 3: Calculate the slope as the slope of the line passing through the points (x_l,y_l) and (x_u,y_u), *i.e.*, $(y_u-y_l)/(x_u-x_l)$. In this case, (0.02646 - 0.00854)/(0.01067 - 0.00263) = 2.229, as shown below:

Calculation of slope

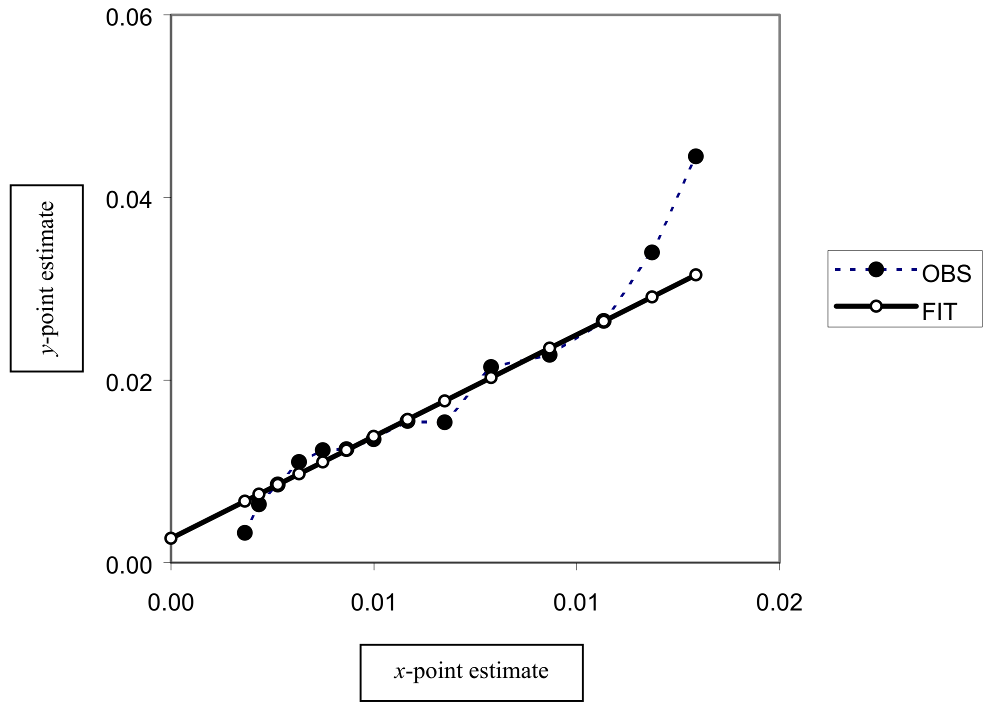
Group of Points	Median x-point	Median y-point	
Lower 3rd	0.00263	0.00854	
Upper 3rd	0.01067	0.02646	
Slope	2.229		
Slope (b) =	2.229	0.185	8.3
Intercept (a) =	0.00268	0.00094	35.1

Column 4: Calculate the intercept of the line with slope $b=2.229$ and passing through each data point, *i.e.*, $a_i = y_i - bx_i$. Calculate the estimated intercept *a* as the median of these values, 0.00268 in this case. As an indicator of the error of *a*, calculate one half the interquartile range of these values, 0.00094 in this case. As an indicator of relative error, calculate one half the interquartile range as a percentage of *a*, $100*0.00094/0.00286 = 8.3$ in this case.

Column 5: To obtain an indicator of the error of the estimated slope, calculate the slope b_i of the line connecting the intercept of the fitting line, $(0,a)$, with each data point (x_i,y_i). The median of these b_i values will generally be very close to the estimated slope. The error indicator for the slope is one half the interquartile range of these values.

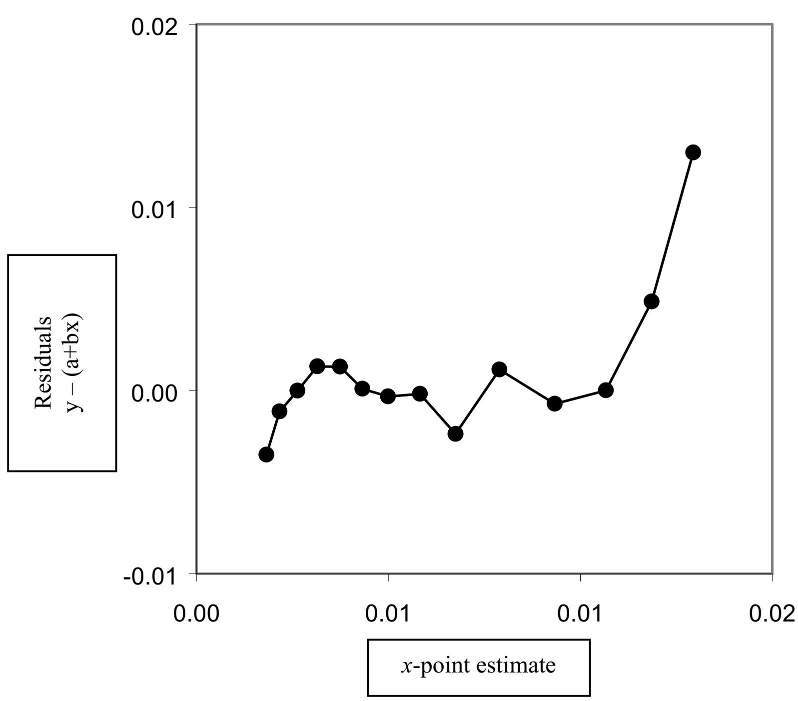
Columns 6-8: Calculate the fitted *y* values for each data point, $\hat{y}_i = a+bx_i$ (column 6), the residuals, *i.e.*, the deviations of the fitted from the observed values, $y_i - \hat{y}_i$ (column 7) and the residuals as a percentage of the observed values (column 8).

Annex figure III.1. Robust fitting: observed points and fitted line



Source: Annex table III.1.

Annex figure III.2. Robust fitting: residuals of fitted line



Source: Annex table III.1.

REFERENCES

Bennett, N. G., and Shiro Horiuchi (1981). Estimating the completeness of death registration in a closed population. *Population Studies* (London), vol. 47, No. 2, pp. 207-221.

Brass, W. and others (1968). *The Demography of Tropical Africa,* Princeton, New Jersey: Princeton University Press.

Brass, W. (1971). On the scale of mortality. In *Biological Aspects of Demography,* W. Brass, ed. London: Taylor & Francis Ltd., pp. 69-110.

__________ (1985). Advances in methods for estimation of fertility and mortality from limited and defective data. An Occasional Publication of the Centre for Population Studies, London School of Hygiene and Tropical Medicine, University of London.

__________, and E. A. Bamgboye (1981). The time location of reports of survivorship estimates for maternal and paternal orphanhood and the ever-widowed. Working Paper No. 81-1. London School of Hygiene and Tropical Medicine, Centre for Population Studies.

Carrier, Norman, and John Hobcraft (1971). *Demographic Estimation for Developing Societies*. London, Population Investigation Committee, London School of Economics, Houghton Street, Aldwych.

Coale, Ansley J., and Paul Demeny (1966). *Regional Model Life Tables and Stable Populations.* Princeton, New Jersey: Princeton University Press.

Graham, W., W. Brass and R. Snow (1989). Estimating maternal mortality: The sisterhood method. *Studies in Family Planning* (New York), vol. 20, pp. 125-135.

Hill, K. (1987). Estimating Census and Death Registration Completeness. *Asian and Pacific Population Forum* (Honolulu, Hawaii), vol. 1, No. 3, pp. 8-13.

__________ and J. Trussell (1977). Further developments in indirect mortality estimation. *Population Studies* (London), vol. 31, No. 2, pp. 313-333.

Japan Statistical Association (1987). *Historical Statistics of Japan*, vol. I of V. Tokyo: Japan Statistical Association.

Nair, K. R., and M. P. Srivastava (1942-44). On a simple method of curve fitting. *Sankya*, vol. 6, pp. 121-132.

McNeil, Donald R. (1977). *Interactive Data Analysis: A Practical Primer*. New York: John Wiley & Sons.

Preston, S. H., and N. G. Bennett (1983). A census-based method for estimating adult mortality. *Population Studies* (London), vol. 37, No. 1, pp. 91-104.

Preston, S. H., and A. J. Coale (1981). Age structure, growth, attrition, and accession: A new synthesis. *Population Index* (Princeton, New Jersey), vol. 48, No. 2, pp. 217-259.

Timæus, Ian M. (1991a). Measurement of adult mortality in less developed countries: A comparative Review. *Population Index* (Princeton, New Jersey), vol. 57, No. 4, pp. 552-568.

__________ (1991b). Estimation of mortality from orphanhood in adulthood. *Demography* (Washington, D. C.), vol. 28, No. 2, pp. 213-227.

__________ (1991c). Estimation of adult mortality from orphanhood before and since marriage. *Population Studies* (London), vol. 45, No. 3, pp. 455-472.

__________ (1992). Estimation of adult mortality from paternal orphanhood: A reassessment and a new approach. *Population Bulletin of the United Nations* (New York), No. 33, pp. 47-63.

__________, B. Zaba and M. Ali (2000). Estimation of Adult Mortality from Data on Adult Siblings. In *Brass Tacks: Essays in Medical Demography: A Tribute to the Memory of William Brass,* Basia Zaba and John Blacker eds., London. Athlone Press.

Tukey, John (1977). *Exploratory Data Analysis.* Reading, Massachusetts: Addison-Wesley Publishing Company.

UNAIDS/WHO (1998). *Report on the Global HIV/AIDS Epidemic, June 1998*. Joint United Nations Program on HIV/AIDS and World Health Organization, UNAIDS/98.10 - WHO/EMC/VIR/98.2 - WHO/ASD/98.2.

United Nations (1967). *Manual IV: Methods of Estimating Bsic Demographic Measures from Incomplete Data.* Sales No. S.67.XIII.2.

__________ (1982). *Model Life Tables for Developing Countries*. Sales No. E.81.XIII.7.

__________ (1983). *Manual X: Indirect Techniques for Demographic Estimation*. Sales No. E.83.XIII.2.

Zimbabwe, Central Statistical Office (1985). *Zimbabwe 1982 Population Census: Main Demographic Features of the Population of Zimbabwe: An Advance Report Based on a Ten Percent Sample.* Harare: Central Statistical Office.

Zimbabwe, Central Statistical Office (1994). *Census 1992: Zimbabwe National Report*. Harare: Central Statistical Office.

__________ (1995). *Zimbabwe Demographic and Health Survey 1994*. Harare: Central Statistical Office.

Zlotnik, Hania, and Kenneth Hill (1981). The use of hypothetical cohorts in estimating demographic parameters under conditions of changing fertility and mortality. *Demography* (Washington, D. C.), vol. 18, No. 1, pp. 103-122.