This document was created by importing the original Wordperfect 5.1 file into more up-to-date software. As a result of this process, the document has been re-paginated and has undergone some formatting changes, in particular to the tables, figures and equations. Thus, it is no longer identical to the hard-copy of the thesis that was examined in 1990.

Advances in the Measurement of Adult Mortality from Data on Orphanhood

Ian Manfred Timæus

London School of Hygiene and Tropical Medicine

Thesis submitted for the degree of Doctor of Philosophy to the Faculty of Medicine, University of London.

May 1990

ABSTRACT

The premature death of adults is a major, but poorly documented, health problem in developing countries. The inadequacy of registration statistics and difficulty of collecting accurate data directly in surveys mean that indirect methods of analysis, particularly those based on orphanhood, represent an important source of adult mortality estimates.

Assessments of the orphanhood method have expressed concern about the robustness of the procedures used to estimate life table indices from orphanhood data, particularly for males, about under-reporting of orphanhood, particularly among children (the `adoption effect'), and about the ways that complete life tables are derived from indirect estimates.

Investigation of the estimation procedures suggests that they are very robust for female mortality and acceptably so for male mortality. Small increases in accuracy would accrue from use of a regression based method to estimate male mortality, that incorporates a more sophisticated fertility model than the original method. Such a procedure is presented, together with one for female mortality based on consistent assumptions. Existing methods for fitting life tables to indirect estimates are sometimes less satisfactory. An alternative approach is proposed and assessed.

In some countries, orphanhood estimates are seriously biased by the adoption effect. Such errors can be reduced by techniques that analyze data on orphanhood in adulthood. Two such methods are developed and tested. The first estimates mortality from period data on orphanhood after age 20; the second uses data on orphanhood since first marriage. The methods are sensitive to age exaggeration, but data on young adults are a promising source of recent estimates of adult mortality.

Finally, procedures are presented for analyzing data on orphanhood prior to marriage. In countries where adults report this information accurately, it can be used to measure adult mortality up to 35 years before the data were collected.

CONTENTS

TABLES

FIGURES

ACKNOWLEDGEMENTS

Much of chapter 1 of this thesis has been published in Timæus and Graham (1989). While the current author drafted most of the material used here, its revision was a joint exercise. Chapter 3 draws on Timæus (1987a, 1988, and 1990) and on unpublished research conducted in collaboration with John Blacker and Roquelio Fernández. The methods presented in chapter 5 further develop those described in Timæus (1986). Some aspects of this earlier research were funded by the UK Overseas Development Administration, the International Statistical Institute and the World Bank.

Most of the thesis was written at the University of Pennsylvania while on an eight month fellowship awarded by The Population Council. Without their financial assistance and the excellent research facilities provided for me by the Population Studies Center, it would probably never have been completed. I am grateful to Sam Preston, for his comments on chapter 4 and checking the crucial equations in chapters 5 and 6; to Doug Ewbank, especially for figure 7.1; and to all the staff and students of the Population Studies Center for helping to make my stay in Philadelphia so enjoyable and productive.

My title recalls those of papers by Bill Brass and John Blacker. This is indeed the sincerest form of flattery. The value of their advice and support over the last decade has been enormous. I also wish to thank all the friends and colleagues who have worked on, and discussed, adult mortality with me over the years. They include Kamal Ahmad, Pauline Airey, George Alter, Christine Callum, Jane Corbett, Nigel Crook, Richard Feachem, Roquelio Fernández, Richard Hayes, Ken and Althea Hill, Joseph Kamara, Alan Lopez, Carolyn Makinson, Abate Mammo, Alec Mercer, Kath Moser, Jackson Mukiza-Gapere, John Paget, Jane Pryer, Laura Rodrigues, Basia Zaba and Susan Zimicki. Collaboration with Wendy Graham has been particularly rewarding.

Kath Kiernan's well-timed advice helped to persuade me to register. Sarah Macfarlane and David Smith kindly provided me with the Ugandan data. Finally, I am most grateful to my supervisor, Allan Hill, both for his good strategic advice and for his comments on my drafts.

Special personal thanks to AS; to CT and ED, for keeping me in touch; PM and GL, for their weekends; and to Odd, Joy and Rajah.

IAN TIMÆUS

Chapter 1

SOURCES OF INFORMATION ON ADULT MORTALITY

1.1 Background

Deaths among adults in developing countries represent a large and neglected problem. Comparisons with low mortality countries suggest that about 8 million avoidable deaths occur between ages 15 and 60 each year, representing 21 *per cent* of avoidable years of life lost before age 85 in developing countries (Murray and Feachem, 1990). From the individual's point of view, while mortality rates are higher in childhood, the length of the adult life span means that the likelihood of dying prematurely in adulthood is greater than the probability of dying in childhood in most countries.

A great deal has now been discovered about child mortality, about its causes and underlying determinants and about ways to intervene cheaply and effectively to prevent death in childhood. Far less is known about adult mortality. In many countries even the overall level of adult mortality is unknown and, in much of sub-Saharan Africa in particular, we are almost completely ignorant about the main causes of death in early adulthood and middle age, about the distribution of adult deaths within national populations and about how adult mortality is influenced by and affects the economy, society and development activities.

Adult ill-health is a problem of increasing importance. The decline in fertility that is occurring in much of the developing world has a beneficial effect on population growth but is resulting in demographic ageing. As the proportion of adults in developing country populations rises so will the relative quantitative importance of adult morbidity and mortality. In addition, the experience of developed countries indicates that the absolute magnitude of some adult health problems may rise in the coming decades. Perhaps the most worrying trend is the increase in smoking in many developing countries, but motor vehicle accidents, occupationally related disease, environmental pollution in urban areas and AIDS are all major causes for concern.

The adult ages encompass the main productive years of life. Though little quantitative evidence of this exists at present, it seems likely that adult deaths have major adverse economic consequences. Because those aged 15-59 years are responsible for caring for old people and dependent children, the impact of an adult death on other household members can be catastrophic. The costs associated with terminal illness and funeral expenses, combined with loss of the domestic labour and income-generating activities of the deceased, may lead to asset sales, increases in debt or withdrawals of children from school that are difficult to reverse. Fragmentary evidence suggests that young children who lose their mother are at a high risk of dying themselves. Thus deaths of adults may have important implications for the distribution of income and welfare between socio-economic groups, for the social support of dependents and for overall production.

Unless adult health issues are explicitly addressed alongside those of children in the formulation of policy, it is likely that cost-effective investments to improve the health of adults and to minimize the consequences of adult disease and mortality will be neglected to the detriment of the welfare of adults and their dependents. Implicit and unconsidered resource allocation decisions may be made, such as a concentration of health care expenditure on expensive, tertiary care of adults suffering from chronic disease, to the detriment of more effective forms of intervention.

The level of mortality is clearly an incomplete and distorted indicator of the health of a population. It is, however, of peculiar importance. Inequality in the face of death is perhaps the most appalling of all social disparities and trends in mortality are perhaps the most general pointer to major changes in human welfare. Demographic research has repeatedly established that the relative levels of child and adult mortality vary markedly between populations (eg. Ledermann and Breas, 1959). Such differences can extend to neighbouring populations and to different ethnic groups living in the same area (Blacker *et al*., 1985). It is dangerous to make inferences about the level of adult mortality by extrapolating from information on child deaths. Only genuine measures of adult mortality can serve as a basis for planning.

More plentiful and accurate population-based measures of mortality in adulthood could contribute to its amelioration in several ways. Firstly, such information can be used to quantify the extent of the public health problem represented by adult mortality, making it possible to assess its priority in the allocation of resources. Secondly, better information could be used to clarify the determinants of adult mortality and thereby to identify points for intervention. Thirdly, quantitative measures are needed in order to monitor and evaluate the effectiveness of programmes intended to reduce adult mortality. Mortality rates by age, sex and cause for sub-groups of the population are required to calculate most indices of need or of programme impact. Unfortunately, most developing countries do not possess the resources required to collect and process such detailed data using the methods currently available.

It is worth mentioning briefly at the outset some of the general issues that make the study of adult mortality more complex than that of child mortality. In broad terms, adult mortality rates are an order of magnitude lower than those of children. Adult deaths are relatively rare events. To obtain reliable measures of adult mortality requires data either on a large sample of people or on events occurring during a long reference period. A second issue that arises is that it is difficult to identify an appropriate informant who can provide reliable information about deceased adults. Accurate data on child mortality can usually be collected from the children's parents and, in particular, their mothers. In addition, the characteristics of parents are among the more important determinants of the risk of dying in childhood. Since there is no single, universally suitable informant to provide data about adult deaths, problems of under-counting and multiple reporting are common. Moreover, it is often unreasonable to use the social and economic characteristics of the respondent as a proxy for those of the dead person. Thirdly, while the vast majority of the child deaths in the developing countries arise from a limited number of infectious diseases, the causes of death in adulthood may be more diverse and more difficult to diagnose. Finally, it is normally possible to obtain usable information on the ages of children and their ages at death. Accurate information on adult ages is more difficult to collect. Older people are less likely to have birth certificates than the young. Levels of education decline rapidly with age in most developing countries. Moreover, even if the person who died knew their own age, the informant who reports their death may not. Thus it is difficult to obtain reliable data that can be used to study age-specific adult mortality in developing countries. Ages tend to be exaggerated and ages at death are exaggerated even more.

The following sections of this chapter review the methods which can be used to measure adult mortality in developing countries. They are grouped under three major headings. The first group of methods yields direct measures of adult mortality. It comprises all approaches in which the information collected and analyzed refers to the deceased individual. These include the traditional sources of data on adult deaths such as vital registration systems, health services statistics and multi-round surveys. The alternative methods yield indirect measures of adult mortality. One group of these methods involves the analysis of census age distributions. In the other, the unit of analysis is not the deceased person but a clearly defined category of their relatives.

1.2 Direct measures

The term `direct' describes the unifying characteristic of this group of information sources, methods and measures, namely dead individuals. Levels, trends and differentials in mortality are derived directly from data on deaths rather than from data which can be related to deaths, as in the case of indirect estimation.

Direct methods and measures of mortality represent the conventional or classical approach to demographic estimation and form the basis of the statistical systems of developed countries. As a result, the establishment of systems for obtaining accurate direct measures of mortality has long been seen as a goal to be achieved by the poorer nations. The wisdom of this view is now being challenged (eg. Hill, 1984a), most obviously by the development of indirect methods for estimating mortality.

The basic counts of deaths from which all direct measures are derived may be gathered more-or-less continuously, as in the case of vital registration or health service statistics, or periodically, by including questions in a census or survey on deaths in the household during a recent fixed interval. Deaths may be identified by `passively' relying on relatives notifying the authorities, or by `actively' seeking events in the population. Additional information about the deceased may be supplied by the individuals themselves before their death, for example through a surveillance system, or afterwards by contact with their relatives. These features of the data vary according to the source of information and, in turn, influence the accuracy, degree of detail, availability and timeliness of the measures of adult mortality. Direct measures have the potential to describe recent mortality levels: constraints on their timeliness arise largely from administrative and data processing delays.

The completeness of registration of deaths in developing countries often falls far short of the 60 *per cent* limit generally regarded as the lowest level of coverage at which the data are usable (Preston, 1984). Registration completeness varies considerably within and between countries, depending, for example, on the cause of death and the age and socio-economic status of the deceased. Generally speaking, coverage is highest in Latin American countries and lowest in Africa. There are two major groups of reasons for incomplete registration of deaths: the first concern the reporting of deaths and the second the processing of the information. Typically, both factors operate in developing countries, though their relative importance varies.

The value of vital registration for studying adult mortality depends on the accuracy as well as the completeness of the reported information. Important dimensions of accuracy, besides the common problem of age misreporting, are the place, time and cause of death. Failure to allocate deaths to the place of usual residence of the deceased can lead to a significant `occurrence' bias towards urban areas or the higher levels of fixed health facilities. Similarly, failure to reassign deaths from their date of registration to their date of occurrence can lead to biases in the apparent death rates for particular periods, which may be exaggerated by a consequent mismatch with the denominator information used to calculate rates. Finally, a major limitation of vital registration data arises from imprecision in reported causes of death.

While sample registration, with more active registrars, appears to offer the possibility of overcoming some of the disadvantages of national systems, it is not used widely. A major operational problem with sample registration systems is that rather few events occur in registration areas of a size that a registrar can know well enough to find out about all deaths (Brass, 1971a). Thus, justifying sufficient remuneration to maintain the motivation of registrars is difficult. A baseline census of the population in the sample area may provide a denominator for the calculation of rates, but, in the absence of constant updating which allows for migration as well as births and deaths, this comparative advantage is soon lost. Dual-record systems, such as the Indian Sample Registration Scheme in which continuous recording of vital events is supplemented by biannual surveys, do appear to be successful at reducing the omission of deaths (Padmanabha, 1984). Unfortunately they are very expensive. Maintaining the independence of the two systems in the field is difficult and the matching of events during analysis is a formidable task.

In the absence of reliable systems for collecting data on deaths routinely, most information on mortality in developing countries has been collected in surveys and other *ad hoc* enquiries. An obvious and simple way of trying to measure mortality levels directly in a census or household survey is to include questions about deaths that have occurred in the household during a fixed period before the enumeration. Most enquiries ask about deaths in the preceding year; the World Fertility Survey (WFS), however, experimented with a two-year reference period asking about the exact date when each death occurred. Data is usually collected on age at death, the sex of the deceased and, sometimes, their relation to the household head. After making minor adjustments for population growth, the enumeration of the living population in the same enquiry provides appropriate denominators for the calculation of mortality rates.

Questions on recent deaths can only yield usable estimates of mortality in fairly large-scale surveys or censuses. Even in high mortality countries, death is a relatively rare event and the sampling errors of mortality rates are relatively large compared with their absolute size. Blacker and Scott (1974) suggest that a survey of about 20,000 households is about the minimum size that might yield reasonably precise estimates. Often even larger enquiries would be required.

There is considerable experience of this approach to measuring adult mortality. The results suggest that it is common for a substantial proportion of recent deaths to be omitted. Often only about a third to one half of the expected number of adult deaths are reported and sometimes far fewer. There seem to be several reasons for this. One major problem is that deaths only occur in a small minority of households and interviewers simply give up asking the question and leave that section of the schedule blank. In addition, reference-period errors may be important and also omissions, perhaps because of an unwillingness on the part of respondents to talk about the dead. In several WFS surveys the number of deaths reported each month declined rapidly as the interval between their occurrence and the survey increased (Timæus, 1987a). Coverage errors are also a problem. Not everyone is clearly attached to a single household and some people live alone. Such individuals may be among those most likely to die but are unlikely to have their deaths reported. Moreover, the death of an adult can precipitate household fission, so that households in which deaths occur may disintegrate before the survey is conducted.

The seriousness of these problems is illustrated by the experience of the WFS in the ten countries where their surveys asked about recent deaths (Timæus, 1987a). Despite the high standards of this programme of surveys, the data were not processed in Mexico, while in Sudan and Yemen AR very few deaths were reported. A substantial, but indeterminate, proportion of deaths were omitted in Cameroon, Mauritania and Morocco. In contrast, in Jordan and Syria the data on adult men, but not women, and in Korea and Lesotho the data on adults of both sexes, appear to be relatively complete.

It seems likely that, with better questionnaire design, field procedures and training of interviewers, more accurate data on recent deaths could be obtained than has been usual in the past. As it stands, however, this way of measuring adult mortality is unreliable. Moreover, its performance can seldom be assessed on the basis of internal evidence. In addition, sampling errors and errors in the reporting of ages at death mean that it is seldom possible to accept the age-specific mortality rates as they stand. Usually the data have to be smoothed by fitting a model life table. Thus, they are of little use for studying age patterns of mortality in detail. Finally, information on recent deaths are of limited value for the study of mortality differentials. Even if sample size constraints do not prohibit disaggregation, it is difficult to collect information retrospectively on the characteristics of the deceased.

Multi-round surveys attempt to circumvent the problems experienced when asking about deaths in the household in a single-round survey by enumerating and then re-enumerating the same population and enquiring about deaths in the intervening period. Since enquiries can be made about each individual present in the first enumeration but not the second, multi-round surveys eliminate most of the omissions and reference-period errors that vitiate the single-round approach. In addition, as age and other personal characteristics can be collected from individuals themselves in the first round, multi-round surveys investigating household change yield much more reliable data than single-round surveys for the study of age-patterns of mortality and mortality differentials.

In contrast to sample registration systems, multi-round surveys are based on an enumeration of the study population. As well as assisting in the detection of events, this means that they provide their own denominators for the calculation of rates. Such surveys are usually conducted over a fixed period, most often one year but sometimes longer, and make no attempt to remain in contact with the population in between rounds of the survey. The frequency of visits affects the level of omissions of events and rounds of the survey may be conducted at yearly intervals, six-monthly intervals or, ideally, every three months. One exception to this pattern is the Population Growth Estimation (PGE) surveys conducted around 1970. These combined continuous recording of vital events with a separate multi-round survey of the same population. These surveys were thus a form of dual-record enquiry with a similar design to that implemented on an ongoing basis in the Sample Registration System of India.

As with the single-round approach to measuring adult mortality, multi-round surveys need to cover some 100,000 person years of exposure to yield sufficiently precise data for the study of mortality levels. To take advantage of the potential that multi-round surveys offer for the study of differentials in adult mortality, the sample size would need to be considerably larger. This sample needs to be interviewed at least twice and the cost of each round often approaches that of a single-round survey.

Multi-round surveys have been conducted in most regions of the world and a great deal of attention has been devoted to their assessment (eg. Adlakha and Nizamuddin, 1984; Scott, 1973; Tabutin, 1984). The UN Latin American Demographic Centre (CELADE) devoted considerable resources to studies of this sort during the 1960's and 1970's and there is a strong tradition of such surveys in francophone Africa.

There is little dispute that, if well conducted, multi-round surveys are one of the more reliable ways of measuring mortality in countries that lack effective routine sources of data (eg. Seltzer, 1969). However, they reduce rather than eliminate errors such as age misreporting and suffer from distinctive technical problems. The field operations are inherently complicated, they require meticulous record keeping and supervision and loss to follow up is almost always a problem, especially in highly mobile groups such as the populations of many urban areas. Events only occur in a small minority of households between any two rounds of the enquiry and maintaining the motivation of interviewers is very difficult. At their worst, multi-round surveys combine the operational problems of both single-round and continuous data collection systems. Moreover, they are expensive and highly demanding of technical and managerial expertise. Finally, as multi-round surveys involve a protracted period of fieldwork, they inevitably suffer longer delays in the production of results than single-round surveys.

In recent years a range of analytic methods have been developed to assess the completeness of reporting of recent deaths and to adjust such data (eg. Brass, 1975; Preston *et al*., 1980; UN, 1983). The basic idea underlying such methods is that everyone who reaches any given age must die at an older age. Therefore, on the basis of a number of simplifying assumptions, it is possible to compare the age distribution of deaths with information on the age structure of the population and obtain measures of the completeness of death reporting. While such methods were conceived originally as a way of adjusting deficient registration data, they can also be applied to information on deaths in a fixed reference period collected in various forms of household survey.

The most basic assumption made by these analytic methods is that the level of under-reporting of deaths is invariant by age. It is generally agreed that this is unlikely to extend to child deaths and so the methods can only be applied to data on adult mortality. In order to estimate the number of deaths that should be reported above each age, both the growth balance method (Brass, 1975) and the Preston *et al*. (1980) method assume that the population has a stable age structure. The growth balance method compares deaths occurring above each age with the population of that age, while the Preston and Coale method uses detailed information on ages at death and a measure of the growth rate to assess the completeness of death reporting. The latter method is more vulnerable to exaggeration of ages at death than the growth balance method but is less sensitive to departures from stability caused by mortality decline. The assumption of constant under-reporting of deaths by age seems to hold widely and both procedures have the advantage that they indicate when the assumptions are invalid and their use inappropriate. Available evidence suggests that they give similar results with data of reasonable quality. However, both procedures are difficult to apply in populations that are heavily affected by migration or where the reporting of age is very poor. In particular, they can seldom be used to adjust data on sub-national populations.

Bennett and Horiuchi (1981) have suggested a procedure for assessing the completeness of data on recent deaths which does not assume stability and can be used when the age structure of the population is changing. To apply the method it is necessary to have information on the age distribution of the population from two separate enumerations so that the rate of growth of each age group can be calculated. The estimates are therefore vulnerable to changes in census coverage. Courbage and Fargues (1979) suggest a different way of avoiding the assumption of stability. Like the other analytic methods, their approach assumes that the under-reporting of deaths is constant by age. It proceeds by comparing the degree to which the reported deaths are concentrated in old age with equivalent measures in model life tables. While the method is rather sensitive to age reporting errors, it only requires one enumeration. Its major advantage is that it can be applied to populations that are affected by migration, such as those of urban areas where death registration may be relatively complete.

The existence of these methods implies that even deficient data on recent deaths may be useful for the estimation of the level of mortality in adulthood in specific circumstances. However, such data are unlikely to be usable in countries where only a minority of deaths are reported and can seldom be used to estimate mortality for population subgroups in order to study mortality differentials.

1.3 Indirect measures based on age distributions

An age distribution is the outcome of the history of births, deaths and migration experienced by a population over the preceding century. Unfortunately fertility and, in many countries, migration have a greater impact on age structure than mortality. Thus, in general, it is difficult to derive mortality estimates from single age distributions. In populations that have experienced little net migration, however, it is possible to attempt to estimate adult mortality if some independent information is available that can be used to control for the effects of fertility on age structure.

One approach to estimation, that requires minimal data, assumes that the population is approximately stable. It proceeds by matching the proportion of the enumerated population aged under 15 years with the equivalent proportion in a stable population model with the same rate of growth (UN, 1983). If the population's rate of growth and the proportion aged less than 15 have been measured accurately, life expectancy at age five in the enumerated population will be approximately equal to the same measure in the stable population even if the age pattern of mortality used to generate the model is rather different from that in the enumerated population.

Most attempts to estimate mortality from single age distributions date back to a time when little other information was available. Methods based on stable population models can only be applied in countries that have experienced near constant fertility for a long period. While it is possible to try and adjust for the impact of migration or a history of mortality decline on age structure, such attempts are seldom very satisfactory. Moreover, errors in the estimation of the growth rate and proportion aged less than 15 can seriously bias the adult mortality estimates obtained. Usually if data were available to estimate the growth rate (eg. from a second enumeration), one would be able to estimate mortality from intercensal survival or growth.

Methods based on the change in the size of a population between two censuses are among the oldest ways of attempting to measure levels of mortality in the absence of registration data. As children enumerated in the second census and aged less than the intercensal interval will have been born after the first census, the approach can only be used to measure adult mortality. Moreover, if international migration has a substantial impact on the population, estimating adult mortality in this way is usually impossible. Traditional procedures for estimating mortality using information from two censuses are based on cohort survival. In other words, the attrition of cohorts is examined by comparing the size of the population in each age group at the first census with the equivalent older age groups at the second enumeration. In practice, age reporting errors almost always raise major problems and a variety of methods have been proposed to adjust for and smooth out their impact (eg. Brass, 1975; UN, 1967, 1983).

Brass (1979) and Preston and Hill (1980) have proposed methods that combine information on intercensal survival with data on registered deaths by age in order to estimate both the completeness of the registration system and changes in census coverage. Despite the attractiveness of this idea, the methods appear very vulnerable to age exaggeration and are of more use as a means of estimating census coverage than mortality. If both registration and census information are available, the best way of analyzing the data is probably to apply the technique proposed by Bennett and Horiuchi (1981) and discussed in section 1.2.

The information required to estimate mortality from intercensal survival is available in most developing countries and, in straightforward applications, the method does not involve the use of demographic models. Intercensal survival was in use before the indirect methods based on special questions were developed and continues to be of importance for the estimation of mortality in large countries such as India, where migration has a limited impact on population trends. The resulting estimates are often too erratic to yield evidence on the pattern of mortality by age. A more fundamental drawback, however, is that the results are very sensitive to changes in enumeration completeness over time. These frequently introduce large errors. Rather than treating data on intercensal survival as an independent source of data on mortality, it is probably more realistic to use an intercensal population projection as a means of assessing the possible errors in mortality estimates derived from other sources.

A further limitation of methods of mortality estimation based on intercensal survival is that they can seldom be used to produce sub-national mortality estimates. Most socio-economic characteristics of individuals can change over time. Moreover, even if international migration is negligible, migratory movements within countries are almost always sufficiently large to preclude trying to estimate regional differentials in mortality in this way. Finally, it should be noted that, because it is only possible to measure mortality on the basis of intercensal survival once the second census has been processed, such methods will seldom be able to provide up-to-date estimates.

Preston and Bennett (1983) have proposed an alternative way of estimating mortality from two census age distributions that uses information on the growth in the size of the *same* age groups between two censuses to calculate a stationary life table age distribution from those reported in the censuses. All other life table measures of mortality can be derived from this distribution, though omissions and age reporting errors for young children mean that only those for adult mortality are likely to be reliable. A further development of the approach, which requires an independent index of the level of mortality in childhood, incorporates a procedure for smoothing the estimates by fitting a relational model life table (Preston, 1983). The method is as easy to apply if the censuses are not separated by exactly five or ten years as when they are. It also reduces the impact of equivalent, though not changing, age reporting errors on the mortality estimates. Like intercensal survival methods, Preston's procedures are subject to biases arising from migration and changes in enumeration completeness. They are more flexible, however, and somewhat more reliable than those based on cohort survival and are, in general, preferable.

1.4 Indirect measures based on questions about relatives

Methods of estimation based on questions about relatives represent an attempt to overcome the difficulties that arise in the collection of information about dead individuals and especially about individuals who died a considerable time ago. These methods are based on data that can be collected in single-round surveys or censuses about the current characteristics of respondents and, in particular, their age. Thus the defining characteristic of the methods considered in this section is that the unit of analysis is a relative of the deceased and not the person who has died. Moreover, it is the details of the informant that are used in the analysis rather than those of the dead individual. Analytical sophistication, involving models of demographic relationships, is substituted for sophisticated (and therefore more expensive) methods of collecting data.

The data used to estimate mortality indirectly, on the basis of special questions put to respondents, are the proportions surviving of some clearly defined category of relatives of the respondent. These proportions are analyzed according to the respondents' ages. For example, information on orphanhood can be collected using two simple questions that can be included on the interview schedules used in censuses and household surveys. They are `Is your mother alive?' and `Is your father alive?'. The primary concern of this thesis is with data of this sort. Tabulations can be produced of the proportion of respondents with living mothers and living fathers in each five-year age group. Data on widowhood and the mortality of other relatives can be obtained in a similar way. On the basis of such information and supplementary data, it is possible to work back to measures of the level of adult mortality that produced the proportions surviving observed among the respondents' relatives.

The proportion of relatives that remain alive at the time of data collection depends on the level of mortality, the ages of the relatives and the time over which the relatives were exposed to the risk of death. The average length of exposure can be calculated from information about the respondents. For example, a respondent's mother must have been alive when the respondent was born and, similarly, both spouses were alive at the time that their marriage occurred. The relationship between the ages of the respondents and those of any category of their relatives depends on demographic processes which can be represented adequately by simple models. Thus adjustment procedures can be derived whereby information on the survival of relatives can be translated into conventional life table measures on the basis of a simple index of the average difference between the ages of the respondents and their relatives.

The idea of obtaining indirect estimates of adult mortality from the proportions of people with parents who have died originated with Henry (1960). Henry's idea was taken up by Brass, who developed a variety of approaches for calculating life table indices from proportions of respondents with living parents, before publishing a straightforward method, expressing the life table measures as weighted averages of two adjoining proportions, that could be applied to both maternal and paternal orphanhood data (see section 2.1; Brass and Hill, 1973). A variety of regression-based procedures for estimating adult female mortality have since been developed, using the Princeton and the new UN model life table systems (see section 2.3; UN, 1983; Palloni and Heligman, 1986).

Although the rationale of the indirect methods of measuring adult mortality is a little difficult to understand, it is shown in chapter 4 that the main limitations of the approach do not arise from the approximations that are involved in the process of estimation. They stem from the basic nature of the data analyzed. The proportion of mothers, or any other category of relatives, that remain alive depends on the level of mortality over a range of ages and a number of years. There is no way in which the relative levels of mortality at older and younger ages and in the recent past and longer ago can be determined without making further assumptions. In practice what is done is to assume that the age pattern of mortality can be adequately represented by some model pattern. If it is further assumed that the level of mortality has declined linearly in terms of that model, it becomes possible to estimate the date at which the mortality of the cohort of relatives reported on by each age group of respondents equals the period measure of mortality prevailing in the population (see section 2.2; Brass, 1985; Brass and Bamgboye, 1981). As these dates are nearer to the present for younger respondents than for older ones, these `time location' methods make it possible to measure the general trend in mortality. Their development has greatly increased the value of the orphanhood and related techniques of estimation. However, indirect methods can yield only broad measures of the overall level and trend in adult mortality and not detailed schedules of age-specific mortality rates for particular years. Chapter 7 examines ways that more detailed information can be inferred from the average patterns encapsulated in systems of model life tables. However, the methods are inherently unable to detect abnormal age patterns of mortality within adulthood or short-term fluctuations in mortality.

While the fact that information on the survival of relatives refers to the experience of cohorts rather than deaths in a single year is a disadvantage for the study of mortality trends, it is an asset for examination of mortality differentials as many pertinent characteristics for the study of mortality remain fixed during adult life (Brass, 1980). On the other hand, the only data usually available on socio-economic characteristics concern the respondent at the time of the enquiry rather than their relatives. Insofar as these characteristics differ from those of deceased individuals at the time of death, any effect that they have on mortality will be masked by misclassification biases. Despite this problem, some studies of this sort have documented large differentials in adult mortality and more consideration needs to be given to the analysis and interpretation of such data.

One advantage of the orphanhood method over direct methods of estimating mortality is that the information used is based on respondents' lifetime experience. Thus fairly precise estimates of the proportions of respondents with living parents can be obtained from surveys of as few as 5000 individuals. Equally, because the questions involved are simple and can be included on census forms, it is operationally feasible to collect orphanhood data on a very large scale and to use them for detailed studies of mortality in comparatively small districts of a country.

Collection of orphanhood data in developing countries was initiated by the inclusion of the necessary questions in surveys conducted in Chad, Mauritania and West Cameroon in the mid-1960's. As an UN Adviser, John Blacker was instrumental in getting the questions asked in a number of African censuses and surveys conducted around 1970. Subsequent to the method being espoused by CELADE, orphanhood data have been collected in many Latin American countries. The questions were asked in 12 WFS studies that used the expanded household schedule and are also being asked in the World Bank's Living Standards Measurement Study (LSMS) surveys and in many of the Demographic and Health Surveys (DHS) sponsored by the US Agency for International Development. In an increasing number of countries, questions about orphanhood have been asked in more than one enquiry. This greatly assists assessment of the resulting estimates of mortality. Thus orphanhood data is quite widely available and there is considerable experience of its use. This is reviewed in chapter 3.

When two sets of orphanhood data have been collected in successive surveys of the same population, it is possible to derive a single set of proportions orphaned from them which reflects adult mortality during the period between the enquiries. Zlotnik and Hill (1981) propose an estimation procedure based on the calculation of a synthetic cohort, while Preston and Chen (1984) suggest a more flexible approach. These methods are discussed in more detail in Timæus (1986) and are further developed in chapter 5.

In another extension of the orphanhood method, Chackiel and Orellana (1985) examine the value of asking supplementary questions about the timing of parental deaths. This allows one to calculate the mean time at which parents died empirically, rather than by the time location methods proposed by Brass and Bamgboye (1981). The data can also be analyzed as if three separate enquiries had been conducted at five-yearly intervals, using the methods just mentioned. The main problem with the approach is that it requires respondents to remember accurately when their parents died. In an unpublished note, Brass has suggested that it may be possible to use information on whether a parent died before or after the respondent married to similar effect. This datum may be reported more accurately. The relevant questions have been included in a number of DHS surveys. Methods for analyzing them are discussed in chapter 6.

The widowhood method, which estimates adult mortality from information about the deaths of first spouses, was initially developed by Hill (1977). Techniques were derived for estimating male mortality from the reports of female respondents and *vice versa*. In order to circumvent the difficulty involved in modelling the effects of remarriage of the widowed and divorced, information is collected on the mortality of *first* spouses. To maximize the accuracy of the answers, the WFS and other surveys have found it useful to ask explicitly whether the respondent has been married more than once, prior to asking `Is your first husband (or wife) alive?'.

As with the orphanhood method, Hill's (1977) original technique, which uses a system of weights to adjust for variations in exposure, has been supplemented by regression-based approaches to the estimation of adult mortality from data on widowhood (eg. UN, 1983). In both cases, a series of conditional probabilities of surviving from one age in early adulthood to a range of later ages is obtained by adjusting the proportions widowed on the basis of indices of the average timing of first marriages.

While questions on the survival of first spouses have been asked in a number of enquiries, the experience gained with this method of measuring adult mortality is far more limited than that with the orphanhood method. In some surveys the method has given plausible results, judged on the basis of consistency with other adult mortality estimates. In other applications the results have been obviously biased. The method appears to be unreliable. While the orphanhood method sometimes indicates steep declines in mortality which seem implausible, estimates based on the questions about widowhood have often suggested constant or increasing levels of adult mortality (Blacker and Mukiza-Gapere, 1988). The biases seem to arise in the reported data rather than the estimation process. They can operate in either direction. One problem is that, in societies in which marriage is ill-defined and getting married more accurately conceived of as a process rather than an event, it may be difficult to establish whether a previous union was a marriage. Unions that ended because of the death of a partner may be ignored. On the other hand, where social stigma attaches to divorce, respondents whose unions have broken up may tend to state that they are widowed rather than divorced. More basically, divorced respondents may not know whether their first spouse is still alive. Problems often arise with the data on female widowhood in surveys where information on the whole household is provided by the household head. Men appear to be ignorant of, or unwilling to report, their wives' former marriages (Blacker *et al.*, 1983).

The principles on which the orphanhood and widowhood methods are based can be extended to the estimation of mortality from data on other categories of relatives. In particular, Hill and Trussell (1977) proposed a procedure using the proportion of surviving siblings. It has seldom been applied but appears to suffer from serious reporting errors as respondents are often unaware of the existence of siblings who died before they were born or when they were very young. More recently it has been suggested that better quality data would be obtained if questioning about deaths was restricted to those siblings who survive to marry. In the form of the sisterhood method (Graham *et al.*, 1989), this approach has been used to measure maternal mortality.

1.5 Strategies for meeting data needs

The health information systems of many developing countries are in desperate need of rationalization (Graham, 1986; Hill and Graham, 1988; Timæus *et al*., 1988). In too many countries different ministries and other organizations are duplicating, or engaged in conflicting, activities. Sophisticated and expensive data collection activities have little point if the information yielded is unreliable, incomplete or covers only a minority of the population. Much of the information that is collected is never processed thoroughly and even less is put to good use in the formulation and administration of health and other programmes. Often organizations and individuals are so overloaded with demands for statistical information that the quality of the data suffers. A few reliable indicators, that cover the whole population, are of much more use for health planning than a mass of information of doubtful accuracy and completeness. Developing countries need statistical systems that they have the financial and technical resources to sustain without permanent international assistance.

It is against this background that the measurement of adult mortality must be considered. The different sources of adult mortality data considered in this chapter vary in cost, in the accuracy, detail and timeliness of the information that they can supply, and in their capacity to throw light on causal hypotheses about the determinants of mortality and the impact of interventions. Only a few methods are basically unsound but none of them are fully satisfactory. Thus, there is a case for adopting an eclectic approach to trying to improve knowledge about adult mortality in developing countries. Ultimately, assessments of the most appropriate measures and methods should be carried out at a country level and in terms of the existing development of the health and statistical infrastructure and the resources potentially available to the health sector.

In general, rather broad indicators of the level and trend in adult mortality and of differentials between regions and social groups, will serve for the allocation of resources. Studies of socio-economic inequalities in mortality, such as the Registrar General's decennial supplements on occupational and areal differentials in mortality in Britain, complement and provide a background to more focused studies intended to establish aetiology or to quantify the impact on health of specific behaviours or services (Brass, 1980). Estimates that may be slightly out of date or of moderate accuracy are sufficient for most purposes as only large differences are likely to be of substantive significance (UN, 1984). Most developing countries could reasonably aspire to collect mortality data that are useful for such purposes.

Vital registration systems represent the traditional source of mortality data in the developed world but the basic rationale for the registration of births and deaths remains administrative rather than statistical. In many countries in Latin America and, to a lesser extent, Asia, registration of adult deaths is nearly complete but the compilation, processing and publication of statistical returns is in disarray. Administrative and financial reforms are needed rather than improvements in methods of measurement. On the other hand, while it would be unwise to consider dismantling an even partially adequate civil registration system, where notification of deaths is very incomplete one can question the high priority accorded to civil registration as a source of mortality data. There is no inherent reason why attempts to improve the coverage of the civil registration system for administrative purposes need to entail attempts to use it as a source of demographic estimates. In countries where only a minority of adult deaths are registered this seems a waste of resources.

There are, nevertheless, major advantages to collecting mortality data on a routine and continuous basis rather than in *ad hoc* surveys. However, it is debatable whether comprehensive national data is needed for most statistical purposes. Some form of sample enquiry may yield adequate district-level and national measures. Interdisciplinary and organizational divisions have often meant that sample registration and survey based approaches to measuring mortality have developed in isolation from the routine production of statistics by the health services. With existing experience, it is difficult to generalize about whether a more integrated approach would be feasible and beneficial. It would, however, have clear advantages in terms of the possibilities for making useful data available to those involved in planning and managing health services at every level. On the other hand, where coverage by such services is low or there is a flourishing private or traditional health care sector, it may be easier to obtain representative data if their collection is managed independently. Even if a separate statistical system for the collection of mortality data is envisaged, there are clearly major advantages if this can be based on comparable administrative units to those used by the health sector, a consideration that also applies to the estimates generated from censuses.

Whether or not the development of routine systems of providing mortality data is practical, most countries will also need to collect data on adult mortality in censuses and *ad hoc* surveys. While suffering from its own limitations, such information can contribute to assessments of the accuracy and completeness of that derived from routine sources. Surveys can also provide a richer and more detailed range of information on mortality differentials. Unfortunately, both the WFS and DHS programmes, which have tended to become the models for other sample surveys collecting detailed data relevant to the study of mortality, have had a major interest in contraception and family planning. This has led to the concentration of effort on gathering data about children, to the neglect of those concerning adults.

Perhaps the most unequivocal conclusion from this review of methods for measuring adult mortality is that questions put to household heads about recent deaths in the household should not be used in their present form. They very seldom yield useful data in censuses and other large-scale enquiries that are unable to deploy teams of well-trained and motivated interviewers. On the other hand, they are unable to measure mortality sufficiently precisely in surveys of a few thousand households. The success of fertility surveys, such as the WFS and DHS, at measuring both the levels and proximate determinants of fertility and child mortality has removed the justification for conducting large, single-round surveys to estimate vital rates, of the kind that were common in the 1970's (Timæus, 1987a).

The main advantage of the indirect methods for the measurement of adult mortality is that they are based on straightforward questions about respondents' lifetime experience that can be posed in single-round enquiries and are efficient in terms of sample size. Questions about orphanhood, in particular, offer the possibility of obtaining useful information fairly cheaply. They can be used even in countries which lack an effective means of collecting information on adult mortality rates in the population on a routine basis. The research presented here is based on a belief in both the value of including indirect questions about adult mortality in large-scale enquiries, such as the census, and of incorporating investigation of adult mortality into national programmes of household surveys. The main limitation of these methods is that they provide rather broad and non-specific measures of mortality.

In summary, in the large number of developing countries where complete registration of deaths remains a distant and questionable goal, the combination of sample routine health reporting systems and the inclusion of indirect questions in single-round enquiries may be the best way of collecting data on adult mortality that can inform national and health sector planning. The estimates that result from these approaches fall a long way short of the ideal of obtaining mortality rates by age and cause for sub-groups of the population. Arguably this does not matter. There are very few countries in which health planning is such a sophisticated process that it can use refined and precise measures to greater effect than broad indications of mortality patterns. Demographic research has shown that the similarity of age patterns of mortality in all populations means that it is possible to estimate mortality rates with reasonable accuracy from simple data on the proportions of surviving relatives of respondents. Further development of these methods may lead to techniques for obtaining more reliable and more specific measures of adult mortality. This thesis is an attempt to contribute to such developments.

Chapter 2

DEVELOPMENT OF THE ORPHANHOOD METHOD

2.1 Derivation of the method

The potential of information on orphanhood for the measurement of adult mortality was first explored by Henry (1960). He developed the ideas of Lotka (1939), who had considered the reverse problem, estimating orphanhood from data on mortality. If survivorship could be treated as a linear function of age, the proportion orphaned at any age *a*, O(a), would to a close approximation depend on the probability of surviving from the mean age ($\overline{M}$) of parents at the birth of their children to age $\overline{M}$+a.[1] For either sex:

$$O(a) = 1 - (1-\epsilon)l(\overline{M}+a)/l(\overline{M})$$

where $l(x)$ is the usual life table notation for the probability of surviving from birth to age *x* and ϵ is an adjustment, set to 1 *per cent*, for male deaths between conception and childbirth or female deaths in childbirth. For older children this relationship underestimates the survivorship ratio (${}_ap_{\overline{M}}$) because death rates increase rapidly with age. The lower risk of dying of those who bear children at young ages fails to fully compensate for the higher risk of dying of parents who bear children relatively late in life. To allow for this, Henry proposes an adjustment based on the variance in ages at childbearing. His estimating equation takes the form:

$${}_ap_{\overline{M}} = \{1-O(a)\}/0.99 + \lambda\sigma^2$$

Henry derives values for λ from the UN (1955) model life tables and presents an illustration of the procedure applied to historical data. He also points out the method's potential value in developing countries. Henry notes that the estimates only represent the mortality of those who bear children and that they will be biased by any association between the parents' fertility and their mortality, but argues that these effects will be slight. In addition, the paper points out that, if mortality has declined over the 20 or so years before the time that the data on orphanhood are obtained, the resulting estimates represent a rough average of mortality during those years.

[1] Notational conventions in the literature on indirect estimation are only partly standardized. A consistent notation is used in this thesis, based on what conventions exist (see appendix 1). Equations drawn from published papers are reformulated in this notation.

The derivation of simple, robust methods for estimating mortality from orphanhood is largely associated with Brass. Hill and Blacker also played an important role in deriving ways of estimating male mortality from paternal orphanhood. Development of the procedure occurred in a number of stages; several variants were circulated informally and eventually published in Brass (1975). The final version of Brass's method, however, is that published in Brass and Hill (1973). Its rationale is outlined in the following paragraphs. Details of the procedure that were omitted from the published paper are explained by drawing on unpublished sources. The relation between female mortality and the survival of mothers is considered first.

Let the number of children born *a* years before a demographic enquiry to women aged *y* be C(a,y). The probability of the children being alive at the time of the enquiry is $l(a)$ and the probability that their mothers are still alive is $l(y+a)/l(y)$. Assuming that the mortality of orphans and children with living parents is the same, the proportion of respondents aged *a* with living mothers, S(a), is:

$$S(a) = \frac{l(a)\int_s^w C(a,y)\,l(y+a)/l(y)\,dy}{l(a)\int_s^w C(a,y)\,dy} \qquad (2.1.1)$$

where integration is over all ages at child bearing *s* to *w*. The proportion of respondents in a five-year age group with living mothers is:

$${}_5S_x = \frac{\int_x^{x+5} l(a)\int_s^w C(a,y)\,l(y+a)/l(y)\,dy\,da}{\int_x^{x+5} l(a)\int_s^w C(a,y)\,dy\,da}$$

The number of children born to women aged *y* is a function of the number of women aged *y* and the fertility rate at age *y*. If we assume a stable age structure, then $C(a,y) = C(0)e^{-r(y+a)}l(y)f(y)$, where *r* is the rate of natural increase and f(y) fertility at age *y*. Therefore:

$${}_5S_x = \frac{\int_x^{x+5} l(a)\int_s^w e^{-r(y+a)}l(y)f(y)\,l(y+a)/l(y)\,dy\,da}{\int_x^{x+5} l(a)\int_s^w e^{-r(y+a)}l(y)f(y)\,dy\,da}$$

$$= \frac{\int_x^{x+5} l(a)e^{-ra}\int_s^w e^{-ry}f(y)\,l(y+a)\,dy\,da}{\int_x^{x+5} l(a)e^{-ra}\int_s^w e^{-ry}f(y)\,l(y)\,dy\,da} \qquad (2.1.2)$$

To represent the age pattern of fertility Brass and Hill use a cubic function, devised by Brass:

$$f(y) = c(y-s)(s+33-y)^2$$

where *s* represents the age at which childbearing starts and *c* determines the level of fertility and cancels in the expression for S(a). The function can readily be integrated to calculate fertility rates for five-year age groups.

There is no straightforward way of integrating either the numerator or denominator of this expression for ${}_5S_x$ and it has to be evaluated numerically. Brass and Hill proceed by summing the approximations for five-year age groups *s* to s+5, s+5 to s+10 *etc.*, taking the survivorship and growth functions at the mid-point of each age group and multiplying by the age-specific fertility rate for the age group. The result is an estimate of S(a) for an exact age *a* and a fixed pattern of fertility defined by the value of *s*:

$$S(a) \approx \frac{\sum_{z=s}^{w-5} e^{-r(z+2.5)} l(z+2.5+a) \int_z^{z+5} f(y)\,dy}{\sum_{z=s}^{w-5} e^{-r(z+2.5)} l(z+2.5) \int_z^{z+5} f(y)\,dy} \qquad (2.1.3)$$

where *r*, the rate of growth, was taken as 2 *per cent* and the *l*(x) values used were taken from Brass's General Standard (Brass, 1971b). To obtain the proportion of respondents in each five-year age group with living mothers, ${}_5S_x$, requires a further process of approximation. The calculations were repeated for each value of *a* from 2.5 years upwards by 2.5 year intervals. Values for ${}_5S_x$ could then be calculated as a weighted average of the estimates of S(a) for ages x, x+2.5 and x+5. The weights allow for growth and mortality across the age group and are defined as $e^{-ra}l(a)$.[2] Finally the calculations were repeated for values of *s* between 10 and 25 years by 2.5 year intervals.

To work backwards from proportions of respondents with living mothers, ${}_5S_x$, to life table measures of survivorship, requires an index of the value of *s* that determines their relationship. The appropriate measure is $\overline{M}$, the mean age at child bearing, *a* years earlier. With a stable age structure, $\overline{M}$ is defined as:

$$\overline{M} = \frac{\int_s^w e^{-ry} l(y) f(y) y\,dy}{\int_s^w e^{-ry} l(y) f(y)\,dy}$$

Given the age structure of the population, $\overline{M}$ is determined by *s*. Values of ${}_5S_x$ for round values of $\overline{M}$ at single years of age were obtained by interpolating between the values calculated from different values of *s*, using a Lagrange-type procedure.

Every value of ${}_5S_x$, for an age group with mid-point N, is equal to a survival ratio *l*(B+N)/*l*(B) in the standard life table used to generate the estimate of ${}_5S_x$. The value of B is close to $\overline{M}$ but is affected by N. Instead of working with B and N, however, it is preferable to make estimates for *b* and *n*, where *b* is a fixed round numbered age near $\overline{M}$

[2] Note that this is exactly equivalent to summation of the numerator and denominator before calculating the proportion.

and *n* is the dividing point between two adjacent age groups. For the estimation of mortality from data on maternal orphanhood, *b* is set to 25 years. The life table survival ratios are calculated as weighted averages of the proportions of respondents with living mothers in the two age groups surrounding age *n*:

$$_{n}p_{25} = W_{n} \cdot {}_{5}S_{n-5} + (1-W_{n})_{5}S_{n}$$

where $_{n}p_{25}$ is equivalent to $l(25+n)/l(25)$. The weighting factors, W_n, depend on *n* and the mean age of childbearing, $\overline{M}$, and can be calculated from the standard values of $l(25+n)/l(25)$ and the approximations for $_{5}S_{x}$. Brass and Hill present a table of these weights for central ages, *n*, from 10 to 60 years and mean ages at child bearing, $\overline{M}$, from 22 to 30 years that can be used to estimate life table measures of survivorship from data on maternal orphanhood collected in demographic enquiries.

Using information on two age groups has the theoretical advantage of allowing, in part, for differences in the age pattern of mortality between the survey population and the standard. If the slope of mortality around age *n* differs from that assumed, the relationships assumed between the proportions orphaned in the two age groups adjoining *n* and $_{n}p_{25}$ are biased in opposite directions. The effect of these biases on the results tends to cancel out. If the true value of the survival ratio, which would be obtained using a correct weight W^{*}_{n}, is $_{n}p^{*}_{25}$, then the error in the estimate is:

$$_{n}p_{25} - {}_{n}p^{*}_{25} = \{W_{n} \cdot {}_{5}S_{n-5} + (1-W_{n})_{5}S_{n}\} - \{W^{*}_{n} \cdot {}_{5}S_{n-5} + (1-W^{*}_{n})_{5}S_{n}\}$$

$$= (W_{n} - W^{*}_{n})({}_{5}S_{n-5} - {}_{5}S_{n})$$

The error in the estimate is not proportional to the error in the weighting factor, as it would be if a system of multipliers applied to information on a single age group was used, but to the error in the weighting factor multiplied by the difference between the proportions orphaned in the adjacent age groups. This is a much smaller quantity.

Estimation of adult male mortality from the proportion of respondents with living fathers is based on the same principles. A similar system of weighting factors is used, derived from equation 2.1.2. The major problem arises in modelling male fertility. Men bear children over a wider range of ages than women. Moreover, partly because of the prevalence of polygyny in some parts of the world, variability in both the timing and the dispersion of male fertility distributions is far greater than that in female fertility distributions. To calculate the weights, Brass and Hill represent male fertility by a polynomial of the form $f(y) = c(y-s)(s+60-y)^3$.

It is possible for fathers to die between the time that a child was conceived and when it is born. Thus the proportion of respondents with living fathers estimated by equation

2.1.3 is S(a-0.75). These proportions were calculated for values of *a* ranging upward from 2.5 years at 2.5 year intervals. In aggregating to the proportions surviving in five-year age groups, ${}_5S_x$, the nine month offset was allowed for by calculating a weighted average of the two values for ages x+1.75 and x+4.25. In this case, the values of ${}_5S_x$ for an age group with a mid-point N are equivalent to a survival ratio $l(B+N+0.75)/l(B)$. Again it is preferable to work with a fixed round age *b* and the age *n* that divides two five-year age groups and to estimate the survival ratios as a weighted average of the proportions of respondents with living fathers in two adjoining age groups. However, the amount of extrapolation from the observed proportions is reduced if survivorship is measured to the end of the upper age group. Therefore, the final relationship used to estimate life table survivorship from data on paternal orphanhood is:

$$l(b+n+2.5)/l(b) = W_n \cdot {}_5S_{n-5} + (1-W_n)\,{}_5S_n$$

Brass and Hill present two tables of weighting factors with age *b* set to 32.5 years and 37.5 years, depending on the mean age of fathers at the birth of their children.

2.2 Locating the time reference of orphanhood estimates

Implicitly, the discussion in section 2.1 is based on the assumption that the life table measures referred to cohorts. If mortality has changed over time, the survival ratios estimated reflect the mortality rates that have prevailed at a range of ages and dates. For each of these cohort measures, there is a point in time during the entire period of exposure when the estimated cohort measure of survival equals the equivalent measure calculated from period mortality rates. It has been shown by Feeney (1980), for child mortality, that this time can be identified on the basis of reasonable assumptions about the age pattern of mortality and the nature of the trend in mortality. Moreover, whatever the rate of change in mortality, the time at which the cohort and period measures are equivalent is the same. Different possible trends in mortality intersect at the same point. The interval since the time to which the estimated survival ratios refer increases with the length of time for which the relatives of respondents in a demographic enquiry have been exposed to the risk of death. It therefore increases with the age of respondents. Thereby, measures estimated from the reports of different age groups of respondent can be used to infer the broad trend in mortality over time. As the basic data refer to cohort survival, however, short-term fluctuations in mortality will tend to be smoothed out.

Equation 2.1.1 can be rewritten as a weighted average of the survivorship ratios, $S(a) = \int_s^w {}_a v_y \cdot {}_a p_y^c \, dy$, where ${}_a v_y = C(a,y)/\int_s^w C(a,y)\, dy$. Thus, ${}_a v_y$ represents the contribution made to S(a) by parents who bore children at age *y*. Time location methods aim to estimate the time T at which the cohort measures of survival, ${}_ap^c_y$, equal the equivalent period measures, ${}_ap_y(T)$, so that $S(a) = \int_s^w {}_av_y.{}_ap_y(T)\,dy$. Period survival is a function of the force of mortality by age at a point in time and cohort survival can be expressed as a function of the force in mortality by age prevailing when that age is reached:

$$\ln({}_ap_y(t)) = -\int_y^{y+a} \mu(z,t)\,dz$$

and

$$\ln({}_ap^c_y) = -\int_y^{y+a} \mu(z,z-(y+a))\,dz$$

where $\mu(z,t)$ is the force of mortality at age *z* and time *t* and the current time is zero. Brass and Bamgboye (1981) show that an approximate expression for cohort survivorship in time period terms can be arrived at by expanding a Taylor series about a point T. Ignoring the higher order terms:

$$\ln({}_ap^c_y) = \int_y^{y+a} \mu(z,T)\,dz - \int_y^{y+a} \{z-(y+a)-T\}\mu'(z,T)\,dz$$

where $\mu'(z,T)$ is the differential of $\mu(z,t)$ with respect to time, evaluated at T. If the trend in mortality is assumed to be linear in α, the level parameter of a relational logit system of life tables (Brass, 1971b), so that $Y(z,t) = K(t-t_o) + Y(z,t_o)$, where *o* is an arbitrary origin and $Y(z,t) = ½\ln[\{1-l(z,t)\}/l(z,t)]$, then it is a property of the system that $\mu'(z,T) = K.dl(z,T)/dz$. This is to say that there is a constant relationship between the rate of change in the force of mortality with time for a fixed age and the rate of change in survivorship with age for a fixed time. The cohort survival ratio can therefore be expressed as:

$$\ln({}_ap^c_y) = \ln({}_ap_y(T)) - K\int_y^{y+a} \{z-(y+a)-T\}dl(z,T)/dz\,dz$$

Cohort survivorship is equal to period survivorship minus a correction factor. The requirement is to identify the time T at which this correction factor vanishes. After some algebraic manipulation, Brass and Bamgboye (1981) show that this time is:

$$T = \frac{\int_x^{x+5} l(a) \int_s^w {}_av_y.{}_ap_y(T)\,\{a.l(y,T) - {}_aL_y(T)\}\,dy\,da}{\int_x^{x+5} l(a) \int_s^w {}_av_y.{}_ap_y(T)\,\{l(y,T) - l(y+a,T)\}\,dy\,da} \qquad (2.2.1)$$

where ${}_aL_y(T) = \int_y^{y+a} l(z,T)\,dz$. The bracketed term in the denominator is the difference between the number of survivors at ages *y* and y+a. This is the incidence of deaths over this age range. The bracketed term in the numerator is the number of person years that would be lived between ages *y* and y+a, if nobody died, minus the actual number of person years lived in this interval. This difference is the number of person years lost by those who died. Dividing one term by the other gives ${}_ag_y$, the mean time since death of those who died between fixed ages *y* and y+a. Integrating ${}_ag_y$ over *y* and *a* gives an alternative formulation for T:

$$T = \frac{\int_x^{x+5} l(a) \int_s^w {}_a v_y\, l(y+a,T)\{1-{}_a p_y(T)\}\, {}_a g_y \, dy\, da}{\int_x^{x+5} l(a) \int_s^w {}_a v_y\, l(y+a,T)\{1-{}_a p_y(T)\}\, dy\, da} \qquad (2.2.2)$$

Thus, if mortality schedules follow a 1-parameter logit system with a linear trend in α over time, the cohort estimates of survivorship obtained by indirect methods equal period survivorship at the average time of death of the respondents' relatives. This time is independent of K, the rate of change in α. While the derivation of equation 2.2.1 takes advantage of a relationship between changes in mortality with age and with time that is specific to a relational logit system of life tables, it possible to arrive at similar formulae for T on the basis of other reasonable assumptions about the trend in mortality with time by age (Palloni *et al*., 1984).

Equation 2.2.1 can be evaluated numerically, using values for ${}_a v_y$ and for the life table measures chosen on the basis of observed data. To develop a straightforward and systematic procedure for estimating T directly, a much simpler relationship must be assumed. Firstly, Brass and Bamgboye (1981) and Brass (1985) argue that the change in T with *a* over limited age ranges is sufficiently close to linear that all respondents in a five-year age group can be treated as of the central age N. Secondly, they argue, at the ages and levels of mortality at which indirect methods are used to estimate adult mortality, the force of mortality increases approximately exponentially with age. As a consequence, for such applications, variation in ${}_a g_y$ with *y* is slight. Therefore, the weighting factors for ${}_a g_y$ in equation 2.2.2 have little effect and all births can be treated as occurring at the mean age at child bearing, $\overline{M}$. To a satisfactory approximation[3]:

$$T = {}_N g_{\overline{M}}$$

To obtain an expression for ${}_N g_{\overline{M}}$ in terms of ${}_5 S_x$, the proportion of respondents with living parents, and other observed measures, Brass and Bamgboye again assume that the incidence of deaths increases exponentially with age. By definition:

$${}_N g_{\overline{M}} = \int_{\overline{M}}^{\overline{M}+N} (z-\overline{M}-N)\, l'(z)\, dz \Big/ \int_{\overline{M}}^{\overline{M}+N} l'(z)\, dz$$

If $l'(\overline{M}+N) = l'(\overline{M})e^{kN}$, by integrating and expanding the right-hand side they demonstrate that:

$${}_N g_{\overline{M}} \approx \tfrac{1}{2} N(1 - kN/6) \qquad (2.2.3)$$

However, using the subscript *s* to refer to a standard life table, in the logit life table system

3 Strictly this relationship applies only to estimates from maternal orphanhood. Because men can die between the conception and birth of their children, the equivalent approximation for paternal orphanhood is $T = {}_{N+0\ 75}g_{\overline{M}-0\ 75}$. For the sake of brevity and clarity, the following discussion is implicitly in terms of maternal orphanhood. For paternal orphanhood, the relevant adjustments can be applied readily to equations 2.2.3 to 2.2.7.

used to derive 2.2.1:

$$\frac{l'(\overline{M}+N)/l'(\overline{M})}{l'_s(\overline{M}+N)/l'_s(\overline{M})} = \frac{({}_Np_{\overline{M}})^2}{({}_Np_{\overline{M},s})^2}$$

By again substituting $e^{kN}l'(\overline{M})$ for $l'(\overline{M}+N)$ and taking the natural logarithms of both sides, one obtains $kN = 2\ln({}_Np_{\overline{M}}) - 2\ln({}_Np_{\overline{M},s}) + k_sN$. Substitution of this expression into equation 2.2.3 yields an estimate of ${}_Ng_{\overline{M}}$, and therefore of T, of:

$$T = ½N\{1 - \ln({}_Np_{\overline{M}})/3 + \ln({}_Np_{\overline{M},s})/3 - k_sN/6\} \quad (2.2.4)$$

Thus, in this formulation, the time references of measures of conditional survivorship obtained from orphanhood data are estimated as half the duration of exposure, N, reduced by a factor that depends on the incidence of mortality in the age range over which the parents are exposed, on average, relative to a standard mortality schedule. Having arrived at this expression for T on theoretical grounds, Brass and Bamgboye (1981) replace ${}_Np_{\overline{M}}$ by ${}_5S_x$ and the last two terms with a table of values indexed by $\overline{M}$+N, derived from a standard life table with $\alpha = -0.4$, and with a small adjustment for $\overline{M}$. The final estimates of T that they propose are obtained from:

$$T = ½N\{1 - \ln({}_5S_x)/3 - f_{\overline{M}+N} - 0.0037(27-\overline{M})\} \quad (2.2.5)$$

Brass (1985) adopts a different approach, using as his standard life table an extremely high mortality one in which $l_s(x)$ is linear over the adult ages and is taken as (1-x/80)/2. As $l_s(x)$ is linear, T = ½N and k_s becomes 0. Thus, T is estimated from observed data using:

$$T = ½N\{1 - \ln({}_5S_x)/3 + \ln[(80-\overline{M}-N)/(80-\overline{M})]/3\} \quad (2.2.6)$$

Both papers present tables that compare the estimates of T produced by the simplified procedures with those obtained by evaluating equation 2.2.1. The agreement is quite close in both cases until the age of the parents, $\overline{M}$+N, approaches 75 years. While the earlier approximation is somewhat more satisfactory, Brass prefers the simpler approach proposed in 1985 because the estimates of T can readily be adjusted to allow for differences in the age pattern of mortality.

Recently, Paget (1988) has recalculated the time references of orphanhood based estimates of male mortality using models of male fertility that extend up to age 80, rather than being truncated at 50 years. The values of T that Paget obtains by evaluating equation 2.2.1 agree less well with the estimates given by the simplified procedures than Brass and Bamgboye's (1981) direct calculations did. This suggests that it might be worth replacing the function $f_{\overline{M}+N}$, which was originally estimated from a range of fertility and nuptiality distributions intended to represent both male and female experience, by separate functions, estimated from appropriate models, for maternal orphanhood, paternal orphanhood and widowhood data.

Brass (1985) points out that, for older respondents, the age pattern of mortality has an appreciable impact on the mean time since the deaths of parents, ${}_{N}g_{\overline{M}}$. For example, if the force of mortality increases relatively slowly with age compared with the standard, a higher proportion of deceased parents will have died a relatively long time ago. If there is evidence of such an age pattern of mortality, it can be allowed for by defining a more appropriate standard by changing β, the second parameter in a relational logit system of life tables (Brass, 1971b). If this is done $l_s(x)$ is no longer linear with age in the extreme life table used as the standard and the term in k_s in equation 2.2.4 needs to be subtracted from the estimates of T obtained from 2.2.6. Making the same assumptions about the form of $l(x)$ as before, the adjustment is:

$$\delta T = -\frac{N}{12}(\beta-1)\ln\left[\frac{(80-\overline{M})(80+\overline{M}+N)}{(80+\overline{M})(80-\overline{M}-N)}\right] \qquad (2.2.7)$$

To infer mortality trends from a series of survival ratios, ${}_{n}p_{b}$, obtained from different age groups of respondents and referring to different dates, it is necessary to convert them all into the same index of mortality that can be compared over time. This is done by fitting a 1-parameter model life table to each measure and obtaining the comparable index from the model. A wide range of indices have been used for this purpose, including the level parameters of various systems of model life tables, ${}_{35}p_{25}$, life expectancy at ages 5, 15, 20 and 25 and temporary life expectancy up to age 70, ${}_{45}e_{25}$. Using the parameters of the models has the advantage of emphasizing that the full life table is being estimated by fitting a model, rather than measured directly. Each of the measures of life expectancy has its own rationale, while using survival ratios, such as ${}_{35}p_{25}$, or temporary life expectancies avoids extrapolation into old age from measures for younger adults. In this thesis, either the α parameter of a relational logit system of life tables or life expectancy at age 15 is used to make such comparisons, the latter on the basis that it is a synthetic measure of mortality over the whole of adulthood.

2.3 Alternative procedures for estimation

Subsequent to the derivation of the weighting methods for estimating adult mortality from orphanhood (Brass and Hill, 1973), several alternative regression based approaches to the estimation of female mortality from data on maternal orphanhood have been proposed. These are based on earlier work that used regression methods to fit models for the estimation of child mortality from data on proportions dead amongst children ever-born (Sullivan, 1972; Trussell, 1975). The basic idea is to use equation 2.1.2 to simulate proportions orphaned from a range of fertility and mortality schedules. Then equations are fitted by least-squares regression that relate the proportion of respondents with living mothers to life table measures of survivorship. Equivalent methods for the estimation of male mortality from paternal orphanhood have not been developed because of the lack of a satisfactory, flexible model of male fertility and scepticism about the robustness of the method.

The original regression method for estimating female mortality from maternal orphanhood was proposed by Hill and Trussell (1977). They simulated 900 sets of proportions orphaned, based on five levels and two age patterns of mortality, derived from Brass's General Standard, and 45 fertility schedules, derived using the Coale and Trussell (1974) model of fertility. The growth rate was calculated from the fertility and mortality schedules using two total fertility rates, 5.0 and 7.0. Hill and Trussell's research preceded the development of methods for allowing for changes in the level of mortality over time. If mortality trends can be ignored, estimates of child and adult mortality obtained in the same survey clearly pertain to the same life table. Thus they propose that, rather than estimate conditional survivorship from age 25, one should estimate unconditional survivorship measures by including an index of mortality in childhood in the regression equation. They experimented with various specifications of the regression model but argue that sufficiently accurate and robust estimates could be obtained from a model of the form:

$$l(25+x) = \lambda_0(x) + \lambda_1(x)\overline{M} + \lambda_2(x).{}_5S_{x-5}\,l(2)$$

Hill and Trussell's methods were subsequently revised and incorporated in the United Nation's manual on indirect estimation (UN, 1983). A slightly more complicated equation was suggested for the estimation of unconditional survivorship measures and coefficients were also published for estimating conditional survivorship ratios using a model of the form:

$$ {}_xp_{25} = \lambda_0(x) + \lambda_1(x)\overline{M} + \lambda_2(x).{}_5S_{x-5} \qquad (2.3.1)$$

The predictive power of a model with this simple form confirms Brass and Hill's (1973) finding that the proportion of respondents with living parents is a good index of adult mortality after controlling for just the mean age at childbearing.

An alternative series of equations for estimating female mortality from maternal orphanhood has been developed by Palloni and Heligman (1986). They simulated proportions orphaned using a slightly different set of Coale and Trussell fertility models, the new UN (1982) model life tables for developing countries and a growth rate consistent with a gross reproduction rate of 3.0. They show that, if, in general, the UN life tables represent the age pattern of mortality in developing countries more accurately than other systems of models, the coefficients that they estimate will produce appreciably better results than those in the UN's (1983) Manual X. The model that Palloni and Heligman fit is also of the form given in 2.3.1, but their methodology differs from that of Hill and Trussell in two ways. First, they estimate a separate set of regression coefficients for each of the five families of life tables developed by the United Nations `because there are non-trivial gains to be obtained whenever the model of mortality is identified properly'. Second, they propose a regression based method of estimating the time reference of the conditional probabilities of surviving. This is based on the assumption of a linear trend in life expectancy. In contrast to the estimates of conditional survivorship, Palloni and Heligman find that the time references are almost invariant across mortality patterns and present a single set of coefficients based on the UN General family of model life tables. Like Brass and Bamgboye (1981), they conclude that the main sources of systematic variation in the time references are the timing of fertility and level of mortality. However, they use several indices of the former and take into account the interaction of the two. Their model takes the form:

$$T = \lambda_0(x) + \lambda_1(x)\overline{M} + \lambda_2(x)P(1)/P(2) + \lambda_3(x)P(2)/P(3) + \lambda_4(x).{}_5S_x + \lambda_5(x).{}_5S_x.\overline{M}$$

where P(1) is the average parity of women aged 15 to 19 and P(2) and P(3) are the equivalent measures for those aged 20 to 24 and 25 to 29.

One practical disadvantage of the regression models in UN (1983) Manual X is that they have only been estimated for respondents aged between 15 and 50, whereas, with the weighting and Palloni and Heligman (1986) methods, an attempt can be made to estimate more recent mortality from data on children. While none of regression approaches has been extended to the estimation of male mortality from paternal orphanhood, the Brass and Hill (1973) weights can be used to obtain strictly comparable results for the two sexes. When both the weighting and Hill and Trussell regressions were used to estimate adult mortality from nearly 1000 set of simulated data (generated with parameters other than those used to derive the regression coefficients), the weighting method performed well for data on respondents aged less than 30 but the regression method performed better at higher ages (UN, 1983). While experience suggests that all the methods for estimating female mortality from maternal orphanhood yield similar results, it is argued in chapter 4 that there are theoretical reasons for concluding that, in general, regression based procedures will produce slightly more accurate results.

Chapter 3

ASSESSMENT OF THE ORPHANHOOD METHOD

3.1 Background

It was clear, even when the technique was first proposed, that mortality estimates from orphanhood would be subject to a number of limitations and errors. The initial papers describing the orphanhood method emphasize two such problems (Brass and Hill, 1973; Henry, 1960). The first of these is that the method measures mortality over an ill-defined period of exposure, rather than current mortality. This issue has been addressed to some extent since then by the development of ways of determining the time location of indirect estimates of adult mortality. The second limitation is the selectivity bias that arises from any association between mortality and numbers of surviving children. The mortality of individuals who do not marry and have children, or who have no surviving children, is not captured at all by data on orphanhood. Moreover, parents who have only one surviving child are reported on only once, while those with large surviving families are reported on many times.

Results from the first surveys to collect orphanhood data in developing countries began to emerge around 1970. There were two papers that discuss the method presented at the first African Population Conference held in 1971. In one of these, Clairin (1974) describes the findings from a survey in West Cameroon as `distinctly encouraging', while, in the other, Blacker (1974) concludes that the results obtained from the 1969 Censuses of Kenya and Uganda are `most promising'. Nevertheless, early experience with the method soon suggested a number of sources of error in orphanhood estimates. Brass (1975) discusses two of these. First, it may be difficult to determine the mean age of childbearing of women accurately, particularly if it has changed over time, so that it cannot be estimated from data on births during the year before the enquiry. Secondly, in many of the early trials of the method, mortality estimates obtained from the first few age groups were very low, relative to those made from data supplied by older respondents. This, Brass suggests, is because very young children who lose their mother are generally ‘adopted’ by other women. For example, they may be reared by a step-mother, their grandmother or some other foster-mother. Errors may result either because enumerators assume that a child's adoptive parent is the natural parent, without ever asking the questions about orphanhood, or because adoptive parents report that orphans are their natural children. The resulting bias in the data tends to diminish as the age of the respondents increases, both because older respondents are less likely to be enumerated in the same household as their foster-parents and because their foster-parents are increasingly likely to have died themselves. Such reporting errors are likely to be particularly severe in populations in which, as in much of Africa, child fosterage is common and the terms 'mother' and 'father' are used to address a wide range of older relatives (Blacker, 1977). This problem has become known as the 'adoption effect' (Hill and Trussell, 1977).

Estimation of the mean age of childbearing of men is especially problematic. Tables can be produced from some surveys that enable this index to be calculated directly. However, when large numbers of young married men are absentee migrant workers and their ages unknown, this is not a satisfactory procedure (Page and Wunsch, 1976). In most cases an index of the timing of male childbearing has to be estimated by adding either the difference between the male and female average age at first marriage or the difference between the average ages of currently married men and women to the mean age at childbearing of women. In polygynous societies, however, the average age gap between women and their spouses is usually wider for young wives, who bear most children, than for older women. This can be allowed for in a variety of ways, depending on the information available (eg. Blacker, 1977). Unfortunately, early on in process of experimentation with the orphanhood method, it became apparent that different ways of obtaining an index of the timing of male fertility can give different answers, without it being obvious which is to be preferred (Page and Wunsch, 1976).

Blacker (1977) published the first attempt to evaluate maternal and paternal orphanhood estimates against other sources of information on adult mortality. His paper considers data collected in a survey in Chad in 1964, which also collected fairly good information on recent deaths; data collected in the 1969 Census of Kenya, which could be compared with estimates from intercensal survival; and data collected in the course of a multi-round survey in Malawi which also measured adult mortality directly. Blacker finds the orphanhood estimates to be generally plausible and internally consistent. Nevertheless, he identifies two problems with the method in addition to those already mentioned. First, male respondents typically report more living parents than female respondents of the same ages. Secondly, the estimates obtained from older respondents are often very erratic and indicate increasingly light mortality as the respondents' ages increase. Blacker ascribes both problems to a systematic tendency for respondents to exaggerate their ages and points out that this is also liable to bias estimates of the mean age of childbearing. His analysis suggests that each of the independent sources of adult mortality estimates is itself subject to errors and biases. Moreover, they measure mortality at later dates than those referred to by the orphanhood estimates. Nevertheless, the agreement between the orphanhood and other estimates seemed reasonably good. Thus, the paper concludes:

> 'to obtain some rough index of adult mortality in a way that is quick, simple and cheap, the orphanhood approach would seem to be as good a bet as any.' (Blacker, 1977).

In Latin America, national life tables based on vital registration represent a more reliable source of information on adult mortality against which to evaluate orphanhood estimates. Somoza (1981) presents such comparisons for female mortality in four countries, Chile, Panama, Costa Rica and Guatemala, and for male mortality in Chile and Panama. In both cases the analysis is restricted to respondents aged 15 and over on the basis that:

> 'The application of the methods to information of children below 15 has systematically produced gross underestimates of mortality. This may be due to errors deriving from adoption as well as a process of selection (recent parents probably experiencing lower mortality than average).' (Somoza, 1981).

Somoza finds that the orphanhood estimates agree well with registration based estimates of survival from around age 25 up to age 70 for periods some 4 to 16 years before the enquiries. He concludes that methods based on orphanhood produce valuable results. However, on the basis of what is now known about the time reference of the estimates, it is apparent that they underestimate mortality by an appreciable amount in Panama and that the information from young respondents may also underestimate recent female mortality in Guatemala and male mortality in Chile.

Despite the optimistic tone of these early assessments of the orphanhood method, evidence was accumulating that, in some populations at least, it performs poorly. This led one of the originators of the method to emphasize its limitations, especially as a source of timely estimates, in a number of publications (Hill, 1981; Hill, 1984b; Zlotnik and Hill, 1981). The growing scepticism about the method stemmed both from the increasing availability of orphanhood data, and in particular of two sets of data on the same population, and from improvements in techniques of analysis, such as the development of time location methods for adult mortality (see section 2.2) and synthetic cohort methods for analyzing data on changes in orphanhood in between two enquiries (see chapter 5). Thus, applying the latter technique to orphanhood data from the 1972 Census and 1976 demographic survey of Peru, Zlotnik and Hill (1981) arrived at what they regard as implausibly light estimates of mortality. They ascribe this to systematic under-reporting of orphanhood. In addition, Hill (1984b) reanalyzed the data on maternal orphanhood from the 1973 Census of Guatemala, presented by Somoza (1981), using time location methods and correcting the registration based estimates for under-reporting of deaths. The results again suggest that the orphanhood method systematically underestimates mortality. The bias decreases with increasing age of respondents, but is very large for estimates based on respondents aged less than 30.

At about the same time, results were also appearing from enquiries in Africa that revealed disturbing inconsistencies between successive sets of orphanhood data collected from the same population. For example, Blacker (1984) presents orphanhood estimates from the 1969 and 1979 Censuses and the 1977 National Demographic Survey of Kenya. Each of the three sets of estimates suggests that adult mortality was declining rapidly, resulting in three parallel trend lines which are incompatible with each other. Estimates from data supplied by younger respondents in the later enquiries indicate lighter mortality than those made for the same date from data on older respondents in the earlier enquiries. Successive sets of orphanhood data for Malawi exhibit similar inconsistencies (Blacker and Mukiza-Gapere, 1988; Timæus, 1986), as do the estimates for Uganda in those districts for which the 1980 Census data survive (Mukiza-Gapere, personal communication).

During the 1980's a great deal of further experience has been accumulated with the orphanhood method. Two or more sets of orphanhood data are now available for quite a large number of African and Latin American populations. Moreover, orphanhood data were collected in a number of WFS surveys that achieved higher standards of fieldwork than most of the early enquiries that collected these data (Timæus, 1987a). It is now clear that the orphanhood method often works very well but that it sometimes performs poorly. Recent assessments of the method have tended to emphasize both its value and its limitations (Blacker and Mukiza-Gapere, 1988; Timæus, 1990; Timæus and Graham, 1989), though McDonald (1987) believes that selectivity biases and the method's inability to produce up-to-date estimates render it useless. The following sections of this chapter review recent experience with the method in order to assess what is now known about the influence of various sources of error on mortality estimates from orphanhood and about the conditions for applying the technique successfully. The discussion does not attempt to be exhaustive: questions about orphanhood have now been asked in far too many enquiries for that to be either practical or desirable.

3.2 Estimates from the World Fertility Survey

One of the optional modules developed for inclusion in WFS studies was a mortality module for use in the household surveys that preceded the individual interviews (Timæus, 1987a). This module included questions about orphanhood, widowhood and recent deaths in the household. In practice, with the exception of Peru, the module was only included in surveys that enumerated an enlarged household sample. However, not every country that enumerated an enlarged sample used all the mortality questions. Data on maternal and paternal orphanhood are available for Cameroon, Lesotho, Mauritania, Morocco and northern Sudan in Africa; for Jordan, Syria and Yemen AR in West Asia; for the Republic of Korea; and for Peru. In addition, Dominican Republic asked about maternal, but not paternal orphanhood.[1] The WFS data are of particular methodological interest because of the opportunity they offer to compare the results with those obtained from data on widowhood and recent deaths in the household. In addition, except in Jordan, Syria, Yemen AR and Korea, at least one other set of orphanhood data is available for all the countries that asked about the subject in the WFS.

The basic orphanhood data and estimates of adult mortality were presented in most of the First Country Reports on WFS surveys that asked these questions. In addition, several detailed analyses of the data on particular countries have been published (Blacker *et al*., 1983; Kim, 1986; Moser, 1985; Timæus, 1984, 1987b). A comparative study of the adult mortality data was also conducted, which has never been published in full, although the findings from it are discussed briefly in Timæus (1987a).[2]

As can be seen from figure 3.1, in all four sub-Saharan African countries, the WFS data on male and female orphanhood agree about the trend in adult mortality and suggest plausible sex differentials in mortality (Timæus, 1990). The estimates also imply similar trends in mortality to orphanhood data from previous enquiries. In Mauritania and Cameroon, the rapid drop in mortality, documented by the WFS results, ties up with the decline from very high levels indicated by the early surveys in 1964 and 1965 (Timæus, 1990). Similarly, in northern Sudan, although the results of the WFS study suggest slightly higher mortality than those from the 1973 Census, they indicate the same general level and trend in mortality. In Lesotho, the WFS estimates agree closely with those obtained in two earlier surveys. They show that, although adult morality in the 1950's was

---

[1] Orphanhood data were also collected in the multi-round survey enumeration in Benin that replaced the usual WFS household survey. These data are not considered here.

[2] This study was conducted at the Centre for Population Studies, London School of Hygiene and Tropical Medicine by John Blacker, Roquelio Fernández and Ian Timæus with partial funding from the International Statistical Institute.

not as high as in West Africa, there was little decline in adult mortality during the 1960's (Timæus, 1984).

In Lesotho and northern Sudan, the WFS also collected data on the survival of first spouses (Timæus, 1987a). In Sudan, the resulting estimates are rather erratic. They imply somewhat higher male mortality than those from orphanhood and suggest a slower pace of mortality decline. The estimates for females agree fairly well with the orphanhood estimates from older respondents, but suggest that mortality was rising during the late 1960's and early 1970's. In Lesotho, the widowhood based estimates for adult women (not shown in figure 3.1) indicate about the same level of mortality as those from orphanhood. The estimates of male mortality from female widowhood imply a very rapid fall in mortality and are clearly biased. All three surveys in Lesotho also enquired about recent deaths in the household and obtained useful data (Timæus, 1984). The results for males are a little erratic, probably because a substantial proportion of Basotho men are working in South Africa at any point in time, but suggest broadly similar mortality to the orphanhood data. The estimates for women imply that the orphanhood method may overestimate life expectancy at age 15 by about 2 years. They confirm that there has been little decline in mortality.

Peru is a country where extensive data on adult mortality are available that have been subject to intensive study (eg. Moser, 1985). The orphanhood estimates from the 1977 Peru Fertility Survey reveal a steady decline in both male and female adult mortality. This is also shown by orphanhood estimates from the 1976 demographic survey and, for females, from the 1972 and 1981 Censuses. Data on male widowhood collected in 1976 (not shown in figure 3.1), imply about the same level of female mortality as the WFS and other orphanhood data, but suggest that mortality rose, rather than fell, during the decade centred on 1968. The equivalent estimates of male mortality from female widowhood, in contrast, exhibit a similar trend to the orphanhood estimates but suggest that mortality was substantially higher than is indicated by the latter. Estimates of male mortality made from measures of lifetime widowhood calculated from the fertility survey marriage histories agree rather better with those from orphanhood. In Peru, adult mortality can also be estimated directly from data on registered deaths and data collected in the EDEN multi-round survey of 1974-6. The former source needs correcting for under-registration, but reporting of adult deaths in the survey appears to have been more-or-less complete (Moser, 1985). Estimates obtained from registered deaths in 1961 and 1972 agree fairly closely with those from orphanhood. So do estimates of female mortality made from the multi-round survey data. This source suggests, however, that life expectancy at age 15 for men may be about 3 years less than the orphanhood data imply.

The 1975 WFS study in the Dominican Republic only collected data on maternal orphanhood. The results can be compared with those from a second fertility survey, conducted in 1980, that also asked about paternal orphanhood. The estimates of female mortality obtained from data on young respondents in the WFS indicate slightly lighter mortality than those from older respondents in the 1980 survey. Extrapolating the trend in all the orphanhood estimates up to 1980 to compare them with estimates made from recent deaths in the household, also suggests that the orphanhood estimates may understate mortality. However, the discrepancies between the different sources of estimates are small, corresponding to differences in life expectancy at age 15 of one or two years.

In Korea, the orphanhood data reveal a large differential between adult male and female mortality. They indicate that mortality fell rapidly in Korea during the 1960's and that adult female mortality had dropped to a low level by 1970. Data on deaths in the year before the survey, which seem to be have been reported accurately, confirm that adult mortality was fairly low by the 1970's but suggest that the orphanhood data, especially for males, somewhat underestimate recent mortality (Timæus, 1987a).

The orphanhood data obtained by the WFS in Morocco and West Asian countries are more difficult to evaluate, but may be less satisfactory, than those considered so far (Timæus, 1987a). In Jordan, Morocco and Syria, the estimates for males indicate a steady decline in mortality, but those for females a more rapid decline in mortality to very low levels. The apparent sex differential in adult mortality widens over time. It iargued by Blacker (1984) and Timæus (1990), that similar findings in a number of East African countries probably reflect errors due to the adoption effect. Children whose mother dies are usually reared by another woman even if their father does not remarry. The foster mother will often be present when the questions about orphanhood are asked and may well answer them on the child's behalf. In contrast, if a widowed woman does not remarry, it is much easier to detect that her children's father is dead. Thus, reporting errors due to the adoption effect seem more likely to affect data on maternal, than paternal, orphanhood, producing an apparently larger drop in female mortality than male mortality.

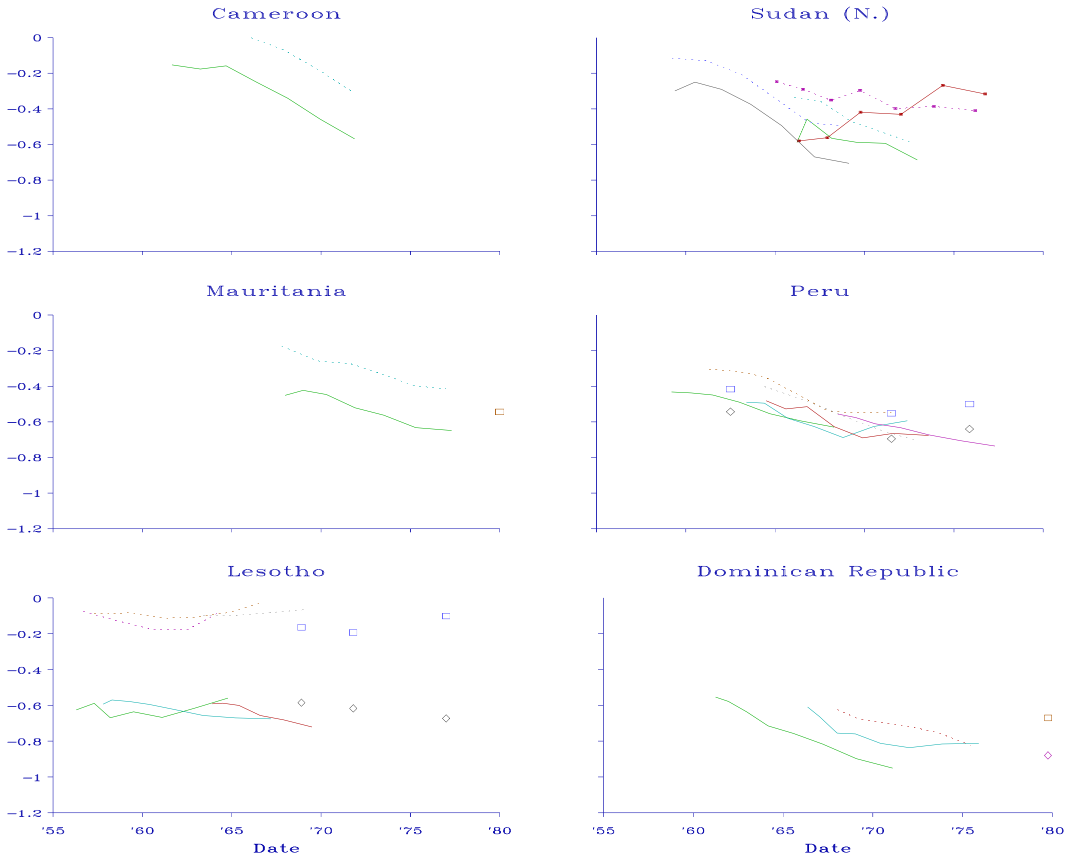

Figure 3.1: Estimates of the level and trend in adult mortality (α) in countries using the WFS mortality module.

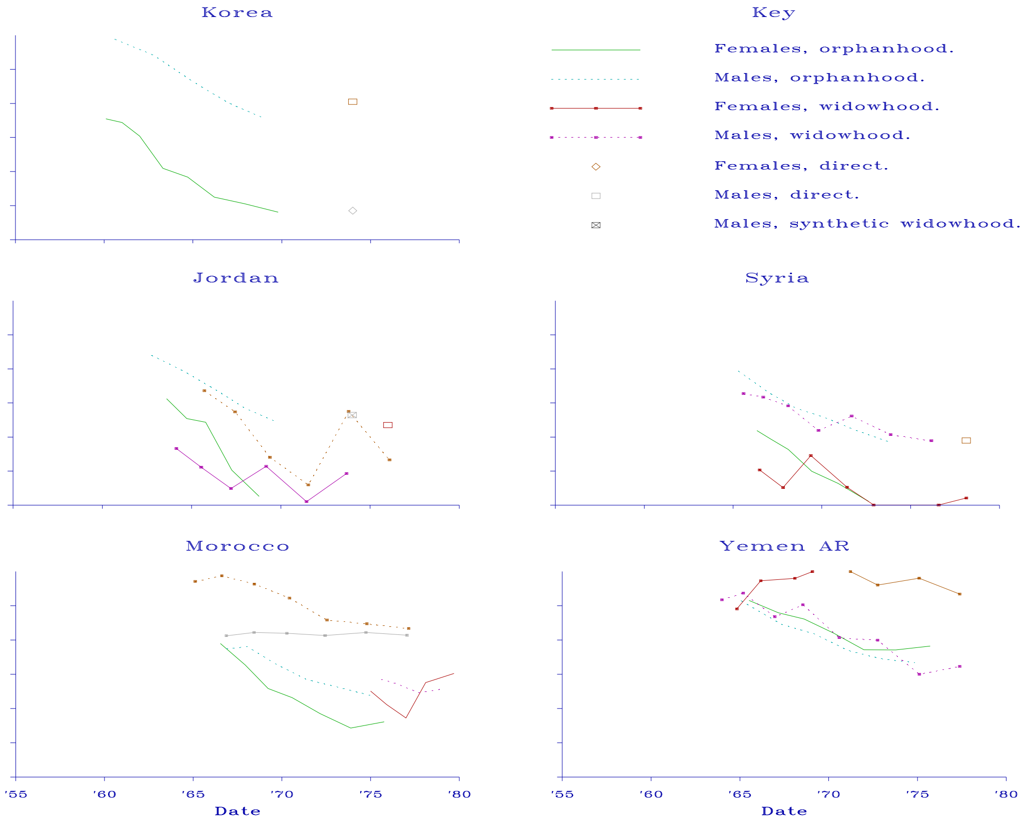

Figure 3.1 continued.

Other sources of estimates of adult mortality in these countries offer some support for this hypothesis. The surveys in Jordan and Syria collected data on widowhood. As elsewhere, the resulting estimates are somewhat erratic and they may underestimate mortality a little, yet they agree fairly well with those from orphanhood (Blacker *et al.*, 1983; Timæus, 1987a). They suggest slower mortality decline than the orphanhood estimates, especially for women. Information on recent male deaths suggest that even the data on paternal orphanhood indicate slightly too steep a decline in adult mortality. Unfortunately, the data on recent female deaths are too incomplete to be usable. In Morocco, data on recent deaths for both sexes are clearly incomplete, but too erratic to adjust. The widowhood estimates indicate much higher mortality than those from orphanhood. Those for women suggest that no decline in mortality has occurred. However, the orphanhood estimates from the WFS can also be compared with those from the DHS survey conducted in 1987. The DHS estimates obtained from younger respondents indicate an implausible trend in mortality and should be discounted. The survey was based on a sample of ever-married women. Early marriage and orphanhood are strongly associated in Morocco and the more recent estimates are not representative of the whole population. The estimates obtained from older respondents in the DHS are, however, are in broad agreement with those from the WFS, but again indicate that the more recent WFS estimates probably underestimate mortality, particularly among women (Al-Jem *et al.*, 1988).

Finally, in Yemen AR, the orphanhood data for the two sexes and the estimates of male mortality from widowhood agree closely. They suggest that, while mortality has fallen, it remains high. Notably, adult female mortality appears to be as high as, or higher than, male mortality. Once again the estimates made from male widowhood seem implausible. They indicate even higher female mortality and imply that mortality is increasing rather than decreasing. Any conclusion about the quality of these data has to be tentative in the absence of estimates from other surveys or recent deaths. Nevertheless, the close agreement between three series of estimates as to both the level and trend in mortality, suggests that the orphanhood questions may have worked quite well.

3.3 Evaluation studies

Rigorous evaluations of the external validity of the orphanhood method of estimating mortality are difficult to conduct. Countries in which reliable data on adult mortality exist tend to differ from those in which the orphanhood method is most useful in several pertinent ways. Often they have much lower fertility and mortality. This probably affects the nature and size of the selectivity biases that affect the orphanhood method, as well as the accuracy of the weighting factors. In addition, countries with sophisticated statistical information systems tend to have different family systems and more educated populations than those that do not. This is likely to affect the nature and extent of reporting errors, such as those that result from the adoption effect. Thus, it is difficult to determine the relevance of studies conducted where there is adequate direct information on adult mortality to the rest of the world. These problems are illustrated in the following discussion.

It is difficult to assess either the size or direction of the bias in orphanhood estimates introduced by multiple reporting about parents by all of their surviving children. There are direct linkages between women's fertility and mortality that operate in both directions. High fertility women are more likely to die in childbearing than women having fewer births. On the other hand, women who die during their childbearing years will tend to have fewer children than those who survive to middle age. Secondly, women whose health is generally poor are probably less fertile than average. Thirdly, in most countries, socio-economic status is associated with fertility, as well as adult mortality. Usually, high status women have lower fertility than average but, on occasion, the relationship is reversed. Fourthly, the mortality of parents and children is likely to be associated. In part, this is because young orphans suffer higher mortality than other children but, in addition, socio-economic, areal and other differentials in mortality are likely to have a similar impact on successive generations. Although it seems likely that the effects are less strong, except for the specific impact of maternal mortality, all of these forces will also tend to bias mortality estimates from paternal orphanhood.

To try and circumvent the problem of multiple reporting biases, the WFS mortality module and other surveys included a supplementary question about whether the respondent was their parent's eldest surviving child. Unfortunately, the data obtained in this way are of very poor quality. Far too many respondents claim to be eldest living children (Hill *et al.*, 1976; Timæus, 1987a). Thus, on average, the parents of this sub-group of respondents are older than is assumed by the procedure for estimation and the estimates of mortality obtained are biased upwards. In contrast to this experience, however, McDonald (1987) reports that the approach has given good results in Indonesia.

At least one survey has collected data that can be used to assess the size of the biases in orphanhood estimates that stem from multiple reporting of orphanhood. The 1980-81 Barbados Experimental Migration Survey asked about respondents' number of siblings ever-born and about how many of them remained alive, as well as about orphanhood. Thus it is possible to calculate proportions orphaned weighted by the reciprocals of the respondent's total number of siblings and number of living siblings (Blacker and Mukiza-Gapere, 1988). The results suggest that data in the usual form overestimate the life expectancy of women at age 15 by a year or two. The bias disappears in the data on older respondents. It arises largely because women who experience low mortality tend to have more children. The association between the mortality of children and of their mothers introduces a small additional bias. Although this finding is suggestive, it is difficult to generalize it to other populations. Both fertility and mortality in Barbados are quite low and underwent a rapid decline in the 1960's and 1970's. One result of this is that the average number of surviving siblings increases rather than declines with age (Blacker and Mukiza-Gapere, 1988).

Perhaps the most valuable evaluation of the orphanhood method in a contemporary high mortality population is that conducted in the Bandafassi area of Eastern Senegal (Pison and Langaney, 1988). This study combined a multi-round survey of vital events from 1970 to 1983 with the collection of detailed chronologies of events and genealogies intended to improve the quality of the data. Questions on orphanhood were included in several of the enumerations of the area. These intensive field procedures produced much better estimates of age than is usual in tropical Africa and the life table computed from the adjusted multi-round survey data seems plausible and has an age pattern of mortality in adulthood very similar to that in the Princeton West model life tables (Pison and Langaney, 1985, 1988). Life expectancy at age 15 is 47 years.

Comparison of the reports about orphanhood obtained in the 1975 enumeration with information compiled from the detailed genealogies, reveals that deaths of fathers were badly under-reported in the structured interviews. Of male respondents, 35 *per cent* of those aged less than 15, 26 *per cent* of those aged 15 to 19 and 6 *per cent* of those aged 20 to 49 with dead fathers were not reported as orphans. For female respondents, the equivalent figures are 28 *per cent* of those aged less than 15, 6 *per cent* of those aged 15 to 19 years and 3 *per cent* of older women. In striking contrast, motherless orphans were almost never misreported as having their mother alive.

Pison and Langaney (1988) go on to compare the corrected proportions orphaned by age group with model proportions calculated for a stable population with characteristics corresponding to the current ones in Bandafassi. In two of the three sub-areas covered by the study the agreement between the observed and model proportions is very close, while, in the third area, the observed data indicate a greater prevalence of orphanhood than the simulated data. This suggests that the corrected data are fairly accurate and that there has been probably been little change in mortality in at least two of the sub-areas. Finally, they compare the indirect estimates of mortality from orphanhood with the directly calculated life table for the area. As might be expected, the agreement between them is good, with the exception that the indirect estimates indicate much higher mortality at older ages than the current life table in the sub-area in which unexpectedly few respondents reported living parents. For example, the probability of surviving from age 35 to 55 for females is 70 *per cent* in the directly calculated life table for the whole study area and 74 *per cent*, 68 *per cent* and 58 *per cent* according to the orphanhood estimates for the three sub-areas.

Apart from being applied to data from developing countries, the orphanhood method of estimating mortality has been used widely in historical demography. In some cases, the results can be compared with mortality estimates calculated from death registers (eg. Nault *et al.*, 1986). Van Poppel and Bartlema (1985) go further and use nineteenth century data from The Hague to conduct a detailed and valuable evaluation of the orphanhood method. As in most historical applications, the information on orphanhood used by van Poppel and Bartlema was recorded at the time of marriage, implying that 79 *per cent* of the 8272 individuals on which information is available are aged between 20 and 35. During the 1870's, life expectancy at birth in The Hague was about 37 years, while at age 20 it was about 41 years. The total fertility rate was 4.85. Thus demographic conditions were broadly comparable with those in high mortality countries today. The mean age at childbearing of women, however, was 31.2 years, which is higher than in most contemporary populations and lies above the range of ages for which weighting factors have been tabulated for the maternal orphanhood method.

What is notable about this study is that, both because of the way the vital event and population registers were originally compiled and because of the checks made during the analysis, the directly calculated life tables, the data on orphanhood and ages at marriage and the information on ages at childbearing are all of very high quality. Therefore, reporting errors can be discounted as a significant source of discrepancies between the direct and indirect estimates. In addition, mortality in The Hague remained more or less constant from 1850 to 1880, rendering interpretation of the results more straightforward than if mortality had changed.

Van Poppel and Bartlema use the methods proposed by Henry (1960), Brass and Hill (1973) and the UN (1983) to estimate adult mortality from orphanhood. The relative error in most of the conditional survivorship ratios that they obtain, compared with those from the directly calculated life tables, is less than 6 *per cent*. Estimates obtained from respondents aged more than 40, using the weighting and regression methods, overestimate female survivorship more seriously, but Henry's method gives very good results. For male mortality, Henry's method performs poorly and the weighting method slightly overestimates survivorship from age 32.5. While van Poppel and Bartlema (1985) describe these as `satisfactory or even good results', 6 *per cent* errors in the conditional survivorship ratios would produce errors of 2 to 3 years in derived estimates of life expectancy at age 15. For paternal orphanhood, these errors consistently lead to the underestimation of mortality.

In part, as van Poppel and Bartlema point out themselves, the differences between the direct and indirect estimates probably reflect stochastic variation in orphanhood, stemming from the small number of respondents in the older age groups. Unfortunately, they do not present their results in a form that allows one to concentrate on ages between 20 and 34 years, the time at which most respondents married. In addition, the need to estimate appropriate weighting factors for a mean age at childbearing of 31.2 years probably has some adverse effect on the estimates of female mortality made using the Brass and Hill (1973) method.

To assess the impact of selectivity biases on the results, van Poppel and Bartlema calculate a life table from the death registers that refers only to the ever-married population, thereby excluding most of that part of the population without children. They find that married women of childbearing ages have higher mortality than single women, while older married women have lower mortality than the unmarried. Married men experience lower mortality at all ages than single men. Life expectancy at age 20 for ever-married women is about one year lower than for the total population. For ever-married men, the same measure is one to two years higher than for the population as a whole.

The orphanhood based estimates of male mortality agree better with the life table for ever-married men than that for the entire population. As might be expected, given the complex pattern of mortality differentials according to marital status by age, there is no consistent tendency for the orphanhood estimates for women to agree more closely with either one of the directly calculated life tables. It seems likely that a similar pattern of mortality differentials by marital status to that in The Hague exists in other high mortality populations. This suggests that, in populations where an appreciable proportion of men remain unmarried, the orphanhood method probably produces moderate underestimates of male mortality. However, in most developing countries very few men fail to marry and so this problem is unlikely to be of significance.

3.4 Simulation studies

In part because of the difficulty of evaluating the orphanhood method against direct measures of adult mortality, there have been a number of attempts to use simulated data and sensitivity analyses to assess the size of the errors in orphanhood estimates that are likely to stem from reporting errors and selection effects. Blacker has performed a number of simple calculations to good effect, while Palloni *et al*. (1984) have conducted a detailed analysis of the maternal orphanhood method.

The strength of the association between the mortality of parents and that of their children is, in principle, open to empirical investigation. The practical difficulty arises in identifying corresponding cohorts of parents and children in population strata with differing mortality. Although unable to do this, Blacker (1984) examines the bias using Kenyan data on maternal orphanhood and child survival according to the education of the respondent. If one assumes that respondents with a given level of education, answering the question about orphanhood, have experienced the same mortality as the children of women with the same level of education, it is possible to reconstruct the original size of the cohorts answering the question about orphanhood. The proportions with mothers alive can then be re-weighted by the number of births of which the respondents are the survivors. Although differentials in mortality according to parental education are large in Kenya, Blacker's calculations show that the revised proportion orphaned, in the population as a whole, differs from the observed one by only 0.2 *per cent*.

Palloni *et al*. (1984) arrive at the same conclusion on more formal grounds. If a population is divided into strata denoted by the subscript *i* and characterized by different mortality, they show that, assuming the age distribution of mothers at childbearing is the same in all strata, the aggregate proportion orphaned is:

$$S(a) = \Sigma_i \, S_i(a) l_i(a) C_i(a) \,/\, \Sigma_i \, l_i(a) C_i(a)$$

where $C_i(a)$ is the number of individuals born *a* years ago in stratum *i*. The bias depends on the product of the differentials in parental and child survival. Even when mortality is fairly high, absolute differentials in *survival* in the respondents' generation are not normally large enough to introduce a significant bias in the aggregate proportions with living parents. Even under the worst conditions, the proportions of respondents with living parents are biased downward by less than 4 *per cent* (Palloni *et al*., 1984).

Palloni *et al*. (1984) also develop expressions to examine selection effects arising from the high risk of dying of orphaned children. They show that, if all children aged under some age *z*, whose mothers die, die themselves, the bias in the estimates that results is more-or-less inversely proportional to $l(\overline{M}+z)/l(\overline{M})$. Under these assumptions, the size of the bias depends on the level of mortality and the value of *z*. Even assuming heavy mortality and that *z* is as high as 5 years, the upwards bias in the proportion of respondents with surviving parents is less than 5 *per cent*. Often it will be less than 1 *per cent* and, if a substantial proportion of young orphans survive, the bias is still smaller. If mortality has declined over time, the errors increase with the age group of respondents.

Thirdly, Palloni *et al*. (1984) examine the selectivity bias that stems from the fact that women who die during their childbearing years tend to have fewer children than those who do not. They point out that this bias will vary in its size between cohorts of respondents. The reports of young respondents underestimate mortality because they only concern women who survived long enough to bear the respondents. The result is to exaggerate the degree of decline in mortality. The maximum size of the bias can be estimated very approximately by examining the scope for differential attrition of groups of women between the mean age of childbearing and the oldest age at childbearing. It turns out to be small. The results that Palloni *et al*. (1984) present suggest that even data on young children could not overestimate life expectancy at age 15 by more than one year. Moreover, in practice, any bias is smaller because respondents' mothers are drawn from a range of cohorts of women at different stages of their reproductive lives.

The sensitivity of the orphanhood method to errors introduced by the adoption effect can also be examined analytically. Once again Blacker has produced a useful assessment based on Kenyan data. His approach is to calculate the proportion of orphaned respondents that would have had to have reported their mother as alive, to reconcile the results of the 1969 and 1979 Censuses. If the inconsistencies between the two enquiries stem solely from the adoption effect, then about one third of children aged 5 to 14 with dead mothers, 20 to 30 *per cent* of such respondents aged 15 to 19 and just under 10 *per cent* of those aged 20 to 24, must have been recorded as having living mothers (Blacker and Mukiza-Gapere, 1988). The idea that reporting might be so poor is discouraging, but not unbelievable.

Palloni *et al*. (1984) develop several analytic expressions for errors arising from the adoption effect, according to whether only orphaned neonates or all orphaned children under some age *z* are adopted and to whether foster-mothers are the same age as the natural mother or distributed by age in proportion to the fertility distribution. If foster-mothers are the same age as natural mothers and all children orphaned before age *z* are adopted and report on their foster-mothers, the relative error in the proportion with living parents, S(a), is approximately $z\mu(\overline{M}+\frac{1}{2}z)$. This implies that the errors are more serious when mortality is high, when childbearing is late and when adoption continues up to a fairly high age. Using a more complicated expression, based on the assumptions that the probability of being orphaned (and reporting about the foster mother) declines with age, that everyone orphaned after age 5 will report on their natural mother and that foster mothers may differ in age from the natural mother, Palloni *et al*. (1984) conclude that the relative errors in S(a) average around 5.5 *per cent* in populations with moderate mortality and 7 *per cent* when mortality is high. On this basis, the errors in S(a) are almost constant by age. This translates into errors in ${}_np_{25}$ that are somewhat greater at young ages and errors in estimates of the level of mortality that are considerably greater at young ages, exaggerating the apparent decline in mortality (Palloni *et al*., 1984; Timæus, 1986).

It should be noted that, because they are expressed in terms of the respondents with living parents, the likely errors due to the adoption effect indicated by Palloni *et al*. (1984) seem modest. They are not. The adjustments made by Blacker and Mukiza-Gapere (1988) to reconcile the two sets of census data on maternal orphanhood in Kenya imply relative errors in the proportions of respondents with living parents of only 1 or 2 *per cent*. Large discrepancies between successive sets of orphanhood data can readily be explained by errors due to the adoption effect that are well within the range that Palloni *et al*.'s (1984) analysis suggest are likely.

3.5 Synthesis

Detailed consideration of data collected by the WFS, supports the conclusion that the orphanhood method is capable of yielding useful estimates of the level and trend in adult mortality for both males and females. In light of the disappointing results from East Africa, the consistency of successive sets of orphanhood data in four sub-Saharan African countries, and the close agreement between direct and indirect estimates in Lesotho is particularly encouraging. Questions about orphanhood also seem to have yielded reasonable estimates of mortality in other WFS surveys, spanning most regions of the world. Although, in a few surveys, questions about widowhood and recent deaths yielded superior mortality estimates, in other surveys they were of little value. The WFS experience clearly indicates that the orphanhood method is the most reliable way of estimating adult mortality in a single-round survey.

Although the orphanhood method seldom yields wildly erroneous estimates of mortality, it clearly suffers from serious limitations and biases. First, there is a great deal of evidence that, even in countries where orphanhood data seem to be of high quality, the method usually overestimates life expectancy at age 15 by a year or so. This is suggested, for example, by data collected in the WFS about male deaths in Jordan, Korea, and Peru and female deaths in Lesotho. In general, the paternal orphanhood method appears to underestimate mortality more than the maternal orphanhood method. Secondly, the orphanhood method often produces exaggerated estimates of the decline in adult mortality, especially for women. This is not only apparent in the East African enquiries that are discussed in section 3.1, but also seems to affect the WFS data on Jordan and Syria and, to a lesser extent, on the Dominican Republic, Morocco and Sudan. Because such errors are liable to occur, it is difficult to interpret a single set of orphanhood data. Interpretation of the evidence is much easier once information on orphanhood has been collected in successive surveys of a population.

The reasons why the orphanhood method tends to underestimate mortality, particularly in the recent past, are difficult to disentangle. The selectivity bias that stems from the reporting of orphanhood by all of an individual's surviving children is probably part of the explanation. Errors that stem from this source are the net outcome of a complex set of relationships between fertility and mortality and the mortality of successive generations. Such errors undoubtedly vary in size, and possibly in direction, between populations. In general, greater biases can be expected in high mortality, high fertility populations and in societies where differentials in mortality and fertility are large.

Considered together, the findings from Barbados, from The Hague and from sensitivity analyses indicate that the predominant selectivity bias is that fertility and mortality are usually inversely related, leading the orphanhood method to underestimate mortality. In part, this relationship arises because death prevents further childbearing, but it also seems likely that prolific individuals are healthier than average. The inverse association between fertility and mortality is probably stronger for males than females because, for women, childbearing entails specific risks leading to maternal deaths. All the evidence suggests that the bias in the estimates that results from this source is unlikely to be large, even for men, but it could provide much of the explanation for the widespread tendency of the orphanhood method to overestimate life expectancy in adulthood by a year or two. Palloni *et al.*'s (1984) analysis suggests that the errors tend to produce an exaggerated impression of mortality decline.

Both the Barbados data and the sensitivity analyses suggest that, even if the children of individuals who die suffer above average mortality, this has little impact on orphanhood estimates. In general, differential mortality risks among respondents lead the orphanhood method to further underestimate mortality. However, to the extent that the children of more fertile women suffer high mortality, either because of crowded living conditions or due to the confounding effects of socio-economic status, the impact of the negative relationship between fertility and mortality in the parental generation will be offset.

A plausible degree of under-reporting of orphanhood due to the adoption effect could explain the inconsistencies between successive sets of orphanhood data observed in East Africa and elsewhere. If such reporting errors are most common among children who lose a parent when they are very young, this will bias data on young age groups more than those on older respondents, exaggerating the apparent degree of decline in mortality. To the extent that orphans are likely to die themselves, the adoption effect cannot operate (Palloni *et al.*, 1984). The two sources of errors are alternatives: they cannot compound one another.

On *a priori* grounds, data on maternal orphanhood seem more likely to be affected by the adoption effect than those on paternal orphanhood. This is consistent with the finding that estimates of female mortality from orphanhood are more likely than those for males to indicate an implausible downward trend in mortality. The only direct evidence that the adoption effect is a problem, however, is that from Bandafassi. It suggests that the data on fathers are affected but not those on mothers. There is no way to determine whether this is typical. The Bandafassi results also indicate that males report orphanhood less accurately than females. This factor, as well as different patterns of age mis-statement (Blacker, 1984), could explain the tendency for men to report a higher proportion of their parents as alive than women of the same age.

In summary, the orphanhood method usually provides reasonable estimates of adult mortality and is more reliable than other approaches that can be used in single-round surveys. It tends to produce moderate underestimates of mortality, especially for males, and sometimes the data obtained from young age groups severely underestimate recent mortality, exaggerating the degree of mortality decline. Both selection biases that are inherent in the method and reporting errors contribute to these problems. Despite the shortage of direct evidence to support the hypothesis, the main reason why the orphanhood method has performed poorly in some parts of the world and some surveys is almost certainly that, in such countries, reports about respondents who are orphaned at a young age tend to concern the individuals who raise them and not their natural parents.

Chapter 4

ROBUSTNESS OF THE ESTIMATION PROCEDURES

4.1 Background

Certain aspects of the rationale of the orphanhood method are not discussed in detail by Brass and Hill (1973). Considering equation 2.1.2, the proportion orphaned at any age is clearly a weighted average of the survival ratios $l(y+a)/l(y)$. The correct weighting factors depend on the age pattern of fertility, the initial age distribution of the parental generation and the rate at which its subsequent mortality increases with age. Brass and Hill argue that differences in age structure have a minor impact on the values of the weighting factors as their effect on the numerator and denominator tends to cancel, but suggest that different age patterns of fertility may adversely affect the accuracy of estimates of male mortality from paternal orphanhood. They do not demonstrate that the relationship between the $l(25+n)/l(25)$ ratios and ${}_5S_x$ proportions, that they estimate using Brass's General Standard, will hold at other levels of mortality. If it does not, one would have to know the level of mortality in order to estimate it!

To some extent, the subsequent derivation of regression based approaches to the estimation of mortality from maternal orphanhood (see section 2.3) has provided justification for procedures that determine indices of adult mortality from proportions orphaned and the mean age at childbearing. The goodness of fit (R^2) of models of this form is high, though the average relative error in the estimates increases with the age of the respondents and is rather large for those that are middle-aged (Palloni and Heligman, 1986). Such analyses, however, give little clue as to which of the assumptions made in the derivation of the estimation procedure are crucial. It is unclear which differences, between a survey population and the central patterns assumed, lead to errors in the estimates and whether there are particular conditions in which the accuracy of the estimation procedure is unacceptably low.

It is widely assumed that the paternal orphanhood method is less reliable than the maternal orphanhood method (eg. Brass, 1975; Hill, 1984b). The distribution of men's ages at childbearing is wider than that of women and differs more between populations. Moreover, the fathers of respondents of any age are usually older than their mothers and subject to higher mortality. However, little quantitative evidence about the sensitivity of this method exists. In this chapter, the robustness of the estimation procedures for both maternal and paternal orphanhood is assessed.

4.2 Maternal orphanhood

The size of the biases that could arise from inappropriate assumptions can be assessed by simulating proportions orphaned from known fertility and mortality schedules and estimating mortality from them using the original weighting factors. These estimates of mortality can then be compared with the schedule used to generate the proportions orphaned. The proportion of respondents by age with living parents in each simulated population is estimated first from equation 2.1.3 and by then interpolating to obtain values for five-year age groups, using the same methods that Brass and Hill (1973) employed to derive the weighting factors (see section 2.1). The effects of variations in the level (α) and age pattern (β) of mortality in the relational logit model life table system based on the General Standard (Brass, 1971b), age structure (as determined by r) and the width of the fertility distribution are examined in turn, holding the other parameters constant at the values used by Brass and Hill to derive the weights (see section 2.1). To assess the impact of variation in the width of the fertility distribution, the cubic devised by Brass is replaced with a relational Gompertz model of fertility, based on the standard developed by Booth (1984). A distribution generated in this way with the parameters $\alpha_f = 1$ and $\beta_f = 1.175$ is very similar in shape to the cubic function and can be regarded as the central pattern.[1]

Table 4.1 shows the absolute errors in α that result when adult mortality is estimated from the proportion of respondents with living mothers under different conditions from those assumed. Positive errors signify that the level of mortality is overestimated and *vice versa*. If the actual demographic characteristics of a population differ from those assumed in the derivation of the method, this only produces small errors in orphanhood based estimates of adult female mortality. The errors increase in size for central ages of respondent of 45 and over, but are trivial for the age groups usually used for estimation. For example, a bias in the estimated value of α of 0.1 represents about a 1.5 year error in the corresponding estimate of the expectation of life at age 15. Most of the biases in table 4.1 are an order of magnitude smaller than this.

[1] To simplify the calculations and their interpretation, the value of α_f used to generate the fertility distributions was fixed at zero and the timing of fertility altered by adjusting s, the age at which childbearing starts. For extreme mean ages at childbearing, a few births will occur at implausible ages. The effect of this is trivial and the results capture the size and direction of errors resulting from differences in the variance of ages at childbearing accurately.

Table 4.1 Errors in estimates of mortality (α) from maternal orphanhood in populations with different characteristics, selected mean ages at childbearing ($\overline{M}$) and central ages of respondents (n).

		$\overline{M}$=23			$\overline{M}$=26			$\overline{M}$=29		
		n=10	n=30	n=50	n=10	n=30	n=50	n=10	n=30	n=50
Mortality										
α=	0.4	-0.002	-0.012	-0.018	-0.019	-0.027	-0.037	-0.036	-0.038	-0.044
	-0.4	0.003	0.014	0.036	0.011	0.021	0.048	0.019	0.025	0.048
	-0.8	0.005	0.026	0.093	0.016	0.038	0.113	0.029	0.041	0.109
β=	0.6	0.008	0.013	0.100	-0.010	0.000	0.105	-0.028	-0.015	0.084
	1.4	-0.005	-0.016	-0.118	0.012	-0.004	-0.122	0.029	-0.010	-0.102
Growth rate										
r=	0.5%	0.004	0.004	0.005	0.006	0.005	0.005	0.006	0.006	0.005
	3.5%	-0.004	-0.005	-0.004	-0.006	-0.005	-0.004	-0.007	-0.006	-0.005
Fertility (Gompertz relational model, α_f= 0)										
β_f=	0.65	0.005	0.044	-0.049	0.024	0.051	-0.048	0.041	0.058	-0.049
	1	0.003	0.009	-0.011	0.006	0.010	-0.011	0.009	0.011	-0.012
	1.175	0.003	-0.001	0.000	0.000	-0.002	0.000	-0.001	-0.002	0.000
	1.525	0.002	-0.014	0.015	-0.009	-0.018	0.015	-0.015	-0.020	0.016

The direction of the errors sometimes changes as the age of the respondents increases because of the way that the parameters affect the distribution of ages at childbearing. If the variance of ages at childbearing is high, young respondents are more likely to be orphaned than is assumed because the proportion of their parents who are old and suffer high mortality is larger. Mortality will be overestimated. By the same token, however, middle-aged respondents are less likely to be orphaned, because most older parents will already have died and the proportion of their parents who are relatively young is also larger. Thus mortality will be underestimated.

Differences in the rate of natural increase of 1.5 *per cent* have a large impact on population age structure but an insignificant impact on the estimated level of mortality, confirming Brass and Hill's argument that the orphanhood method is very robust to errors in the assumptions made about the age distribution of mothers. Similarly, variation in the width of the fertility distribution from that assumed only has a small impact on the estimates of mortality as, in all populations, once the mean age at childbearing is fixed, most births occur to women in a narrow range of ages. The estimates are more sensitive to differences in the level and pattern of mortality. If mortality is higher overall than was assumed in the estimation of the weights or increases rapidly with age, the survivorship ratios estimated from proportions orphaned are biased upwards. Mortality will be underestimated and the decline in mortality over time understated. Conversely, if mortality is lower than was assumed when deriving the weights or increases slowly with age, it will be overestimated and the decline in mortality exaggerated. Compared with the likely impact of reporting errors, these biases are small.

While the results presented in table 4.1 only examine the effect of deviations from each assumption in turn, experimentation suggests that the errors resulting from plausible combinations of parameters are more likely to offset, than to compound, one another. For example, if light mortality (α = -0.8) is combined with rapid increase (*r* = 3.5 *per cent*), the root mean squared error (RMSE) in the estimated values of α over mean ages of childbearing from 24 to 29 years and central ages, *n*, of 10 to 40 years declines from 0.039 to 0.034. If, on the other hand, equally light mortality is combined with low growth (*r* = 0.5 *per cent*) and a more concentrated distribution of ages at childbearing (β_f = 1.525), the RMSE in the estimates of α falls to 0.020.

4.3 Paternal orphanhood

Comparable measures of the impact of the characteristics of a population on the accuracy of estimates of adult male mortality from paternal orphanhood to those in table 4.1, are shown in table 4.2. The effect of different age patterns of fertility is examined using a relational Gompertz model and a standard developed recently by `stretching' the female standard to age 80 (Paget, 1988). The model fits male fertility distributions remarkably well. The standard fertility schedule is broadly similar in shape to the polynomial used by Brass and Hill to represent male fertility and is taken as the central pattern, although it exhibits somewhat higher fertility at late ages. A β_f of 0.675 produces a broad fertility distribution, such as those characteristic of polygynous societies, and a β_f of 1.75 a narrow fertility distribution, such as those characteristic of countries with fairly low fertility.

Table 4.2 Errors in estimates of mortality (α) from paternal orphanhood in populations with different characteristics, selected mean ages at childbearing ($\overline{M}$) and central ages of respondents (n).

		$\overline{M}$=29			$\overline{M}$=34			$\overline{M}$=39		
		n=10	n=25	n=40	n=10	n=25	n=40	n=10	n=25	n=40
Mortality										
α=	0.4	-0.016	-0.032	0.022	-0.057	-0.067	-0.039	-0.075	-0.075	0.007
	-0.4	0.019	0.043	0.029	0.043	0.064	0.061	0.059	0.074	0.032
	-0.8	0.036	0.096	0.101	0.078	0.132	0.152	0.111	0.160	0.105
β=	0.6	0.015	0.050	0.154	-0.011	0.037	0.180	-0.003	0.074	0.250
	1.4	-0.014	-0.056	-0.139	0.009	-0.043	-0.164	-0.002	-0.077	-0.202
Growth rate										
r=	0.5%	0.008	0.011	-0.004	0.013	0.012	-0.006	0.016	0.011	-0.028
	3.5%	-0.010	-0.012	0.006	-0.014	-0.013	0.007	-0.016	-0.011	0.029
Fertility (Gompertz relational model, α_f= 0)										
β_f=	0.675	0.079	0.052	-0.076	0.087	0.047	-0.068	0.083	0.019	-0.124
	1	0.004	-0.003	-0.008	0.000	-0.006	-0.006	-0.002	-0.008	0.011
	1.375	-0.040	-0.042	0.035	-0.051	-0.043	0.040	-0.055	-0.032	0.136
	1.75	-0.065	-0.066	0.062	-0.078	-0.065	0.069	-0.085	-0.047	0.232

As has been suggested before, the paternal orphanhood method is more vulnerable than the maternal orphanhood method to deviations in the characteristics of the population surveyed from those assumed in the method's derivation. Errors in the estimated level of mortality follow broadly the same pattern as those in orphanhood based estimates of female mortality, but are larger. For example, if mortality is lower than is assumed in the derivation of the weighting factors or rises slowly with age, it will be underestimated and *vice versa*. Differences in the growth rate have little impact on the results, suggesting that, as for female mortality, orphanhood based estimates of male mortality are very robust to variations in age structure. However, in contrast to those for females, estimates of male mortality will be appreciably biased if the width of the fertility distribution is very different from that assumed. If men's ages at childbearing are very dispersed, the level of mortality estimated from data on young respondents will be too high and the level estimated from data on older respondents too low. Thus, the trend in mortality over time inferred from the estimates will be biased upward. In contrast, if childbearing is very concentrated, the extent of mortality decline will be exaggerated.

The finding that empirically common combinations of demographic characteristics are more likely to produce offsetting errors in mortality estimates from orphanhood than errors that compound one another applies to males as well as females. As a very approximate guide, for central ages of respondents, *n*, of 35 or less, deviations from the assumptions made in the derivation of the paternal orphanhood weights are unlikely to introduce errors in estimates of male life expectancy at age 15 of much over one year. The estimates obtained from data supplied by older respondents are less robust. This is satisfactory for most purposes, but indicates that one should be cautious about drawing conclusions about sex differentials in adult mortality from orphanhood data.

4.4 Sensitivity analysis

An analytic examination of the sensitivity of the orphanhood method supports these conclusions. The probability that a parent survives from the birth of a child to when that child is interviewed, ${}_ap_y$, can be expressed as a Taylor series (Mammo, 1988):

$$ {}_ap_y = {}_ap_{\overline{M}} + {}_ap_{\overline{M}}'(y-\overline{M}) + {}_ap_{\overline{M}}''(y-\overline{M})^2/2 + ... \qquad (4.3.1)$$

The higher order terms of the series are moderately large in size, increase with age and have signs that can vary, explaining the complex pattern of errors shown in tables 4.1 and 4.2. However, if one is concerned only to identify the main limitations to the robustness of the orphanhood method, they can be illustrated by the initial terms. Mammo (1988) also shows that:

$$ {}_ap_{\overline{M}}' = \frac{d\{e^{-\int_{\overline{M}}^{\overline{M}+a}\mu(z)\,dz}\}}{dx} = {}_ap_{\overline{M}}\{\mu(\overline{M}) - \mu(\overline{M}+a)\}$$

and similarly that:

$$ {}_ap_{\overline{M}}'' = {}_ap_{\overline{M}}[\{\mu(\overline{M}) - \mu(\overline{M}+a)\}^2 + \{\mu'(\overline{M}) - \mu'(\overline{M}+a)\}] \qquad (4.3.2)$$

Substituting the Taylor series defined in 4.3.1 into equation 2.1.1 and integrating, yields:

$$ S(a) \approx {}_ap_{\overline{M}} + {}_ap_{\overline{M}}''\sigma^2/2 \qquad (4.3.3)$$

where σ^2 is the variance of ages at childbearing:

$$ \sigma^2 = \int_s^w C(a,y)(y-\overline{M})^2\,dy \,/ \int_s^w C(a,y)\,dy $$

Substituting expression 4.3.2 for the second derivative of ${}_ap_{\overline{M}}$ in equation 4.3.3 and rearranging, one obtains:

$$ {}_ap_{\overline{M}} \approx S(a) / (1 + ½[\{\mu(\overline{M})-\mu(\overline{M}+a)\}^2 + \{\mu'(\overline{M})-\mu'(\overline{M}+a)\}]\sigma^2) \qquad (4.3.4)$$

Thus Mammo's (1988) analysis shows that, after controlling for the mean age at childbearing, variation in the relationship between proportions orphaned and life table survivorship is determined largely by the variance of ages at childbearing multiplied by a factor, ${}_ap_{\overline{M}}''/{}_ap_{\overline{M}}$, which reflects the way in which mortality increases with age.

Evaluation of equation 4.3.4, with alternative values for the mortality rates and their differentials and for the variance of women's ages at childbearing to those used in the derivation of the weighting factors, indicates a similar pattern of biases to that shown in table 4.1. When *a* is low, $\mu(\overline{M}+a)$ is similar to $\mu(\overline{M})$ and ${}_ap_{\overline{M}}''/{}_ap_{\overline{M}}$ is small. The proportion of respondents with living parents is closely related to the parents' probability of surviving to $\overline{M}+a$, limiting the impact that mis-specification of either the pattern of mortality or the variance of ages at childbearing (which is multiplied by this factor) can have on the mortality estimates. At older ages ${}_ap_{\overline{M}}''/{}_ap_{\overline{M}}$ is much larger, indicating that orphanhood estimates become increasingly sensitive to errors in the assumptions used to derive the weighting factors as the age of the respondents increases. Moreover, examination of a range of fertility and mortality schedules suggests that the higher order derivatives of ${}_ap_{\overline{M}}$ differ more between populations than the higher moments of the distribution of women's ages at childbearing, supporting the conclusion that the maternal orphanhood method is more sensitive to deviations from the assumptions made about mortality in deriving the weights, than to other biases.

In very broad terms, the variance of men's ages at childbearing tends to be about 50 *per cent* larger than that of women. Moreover, the distribution of ages at childbearing of men differs rather more between populations than that of women, as it is a convolution of the distributions of female ages at childbearing and the relative ages of spouses. Thus, the *absolute* differences that are likely for men between the actual variance of ages at childbearing in a survey population and that assumed are considerably larger than those for women. This is reflected in the much larger errors in the mortality estimates in table 4.2, than in table 4.1, that result when the value of β_f used to model fertility is inappropriate. Secondly, because respondents' fathers tend to be older than their mothers, both the size of ${}_ap_{\overline{M}}''/{}_ap_{\overline{M}}$ and likely errors in it are somewhat larger for paternal orphanhood data than those on mothers, especially for older respondents. These considerations explain why the paternal orphanhood method is less robust than the maternal orphanhood method, particularly to differences in fertility patterns.

4.5 Revised estimation procedures

The sensitivity of mortality estimates obtained using the weighting factors to variation in the level and pattern of mortality is somewhat perturbing. The expectation of life at birth in the General Standard is about 43 years, which represents rather heavier mortality than is commonly found in the developing world today. In general, therefore, if reporting is accurate, the weighting factors will produce slight overestimates of mortality. For female mortality the bias is trivial but, for males, it is large enough to be of concern when mortality is either very heavy or very light. The fact that the size of the bias is systematically related to the level of mortality, and therefore orphanhood, suggests that it should be possible to improve the accuracy of the estimates.

The source of the bias becomes clear when simulated proportions of respondents with living parents are plotted against corresponding life table measures of survivorship at a range of levels of mortality. This relationship is illustrated in figure 4.1 for paternal orphanhood and several age groups, using mortality schedules generated from the General Standard by varying α, that have expectations of life at birth that range from 30 to 74

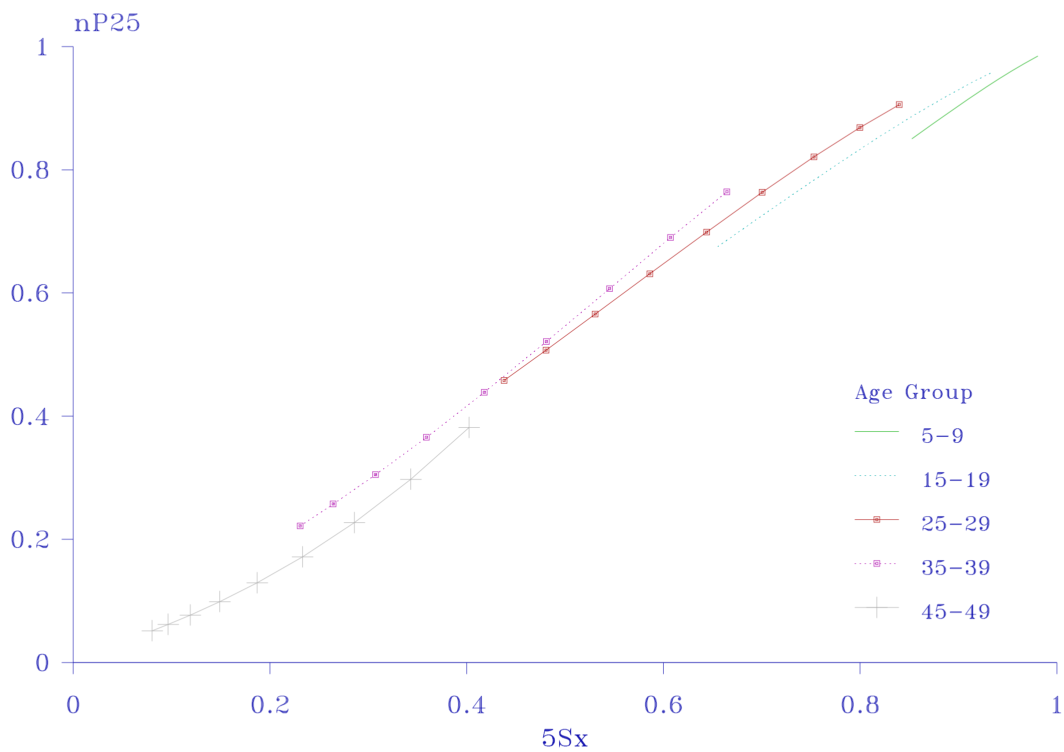

Figure 4.1: Relationship between paternal orphanhood and life table survivorship at different levels of mortality in selected age groups ($\overline{M}$ = 35 years).

years. The errors do not arise from non-linearities in the relationship between orphanhood and mortality. Except at extremely high and low mortality and in old age, the relationship is close to linear. In fact, the bias in orphanhood estimates obtained using the weighting factors stems from the lack of an intercept term. The relationship *is* non-linear at the extremes, which largely fall outside the range of human experience. Thus, as mortality declines, the proportion of respondents with living parents increases more slowly than a line fitted to a single mortality schedule and passing through the origin would suggest. The problem becomes more serious as the age of respondents increases because the proportion of them with living parents varies more with the level of mortality. Thus, mortality is systematically underestimated, when it is high, and overestimated, when it is low. This problem does not arise with regression based procedures, as they are sufficiently flexible to model accurately relationships such as those shown in figure 4.1.

As the analysis in sections 4.2 and 4.3 suggests, the size of the errors that arise from lack of an intercept term are small for the maternal orphanhood method. For paternal orphanhood they are much larger. This implies that worthwhile improvements in the accuracy of measures of male mortality from paternal orphanhood would result from using a regression based procedure for estimation. Further minor increases in the overall precision of the method would result from replacing the functions used to represent fertility in the original derivation of the weighting factors with more sophisticated model fertility distributions.

New regression coefficients for the estimation of adult mortality from orphanhood are presented in appendix 2. They were estimated from simulated data created from relational logit model life tables, based on the General Standard, and fertility distributions generated, using the relational Gompertz model, from Booth's (1984) standard for females and Paget's (1988) standard for males. These standards were developed for use in populations with high fertility. As in the original calculations, the age distribution of the parents is represented by a stable population with a rate of increase of 2 *per cent*. The coefficients have been calculated for maternal, as well as paternal, orphanhood to provide a consistent basis for estimation for the two sexes. While there is no reason to suppose that the coefficients for maternal orphanhood will perform any better than those published previously, it is reassuring to note that they are usually intermediate in size to those published in the UN's Manual X (1983) and those estimated by Palloni and Heligman (1986), using the UN's General family of model life tables.

The parameters used to generate the model populations are shown in table 4.3. They result in populations with life expectancies at birth that range from 36 to 73 years and life expectancies at age 15 that vary between 39 and 62 years. The mean age at childbearing of mothers falls between 23.4 and 31.5 years, while that of fathers ranges from 29.3 to 40.3 years.

The proportion of respondents with living parents by five-year age group in each simulated population was calculated by evaluating equation 2.1.3, using essentially the same procedure as Brass and Hill (1973). The steps involved are described in section 2.1. One refinement was adopted that made an appreciable difference to the results. The proportion of living fathers in each five-age group, ${}_5S_x$, was calculated by interpolating between estimates for three exact ages, x-0.75, x+1.75, and x+4.25, rather than just the latter two ages, as in the original method.

Table 4.3 Parameters that define the simulated populations for estimation of the relationship between parental survival and life table measures.

A) **Mortality**

α = 0.2, -0.2, -0.6, -1.0

β = 0.8, 1.1

B) **Fertility - females**

α_f = -0.5, -0.2, 0.1, 0.4

β_f = 0.7, 1.0, 1.3

C) **Fertility - males**

α_f = -0.4, -0.1, 0.2, 0.6

β_f = 0.7 (α<0.6), 1.0, 1.3 (α>-0.4)

The coefficients have been estimated from a more limited set of populations than those published previously. Inspection of the residuals and of the bivariate relationships between the mean age at childbearing and the proportions of surviving parents, on the one hand, and the predicted life table measures, on the other, confirms that the relationships are linear and that the variance of the independent variables is unrelated to the dependent measures for both males and females. Coefficients can be estimated reliably, therefore, from a few evenly spaced points. The β parameter of the mortality models, which affects the age pattern of mortality, and the β_f parameter of the fertility models, which largely determines the variance of the distributions, were set to two and three different values respectively, to check that this did not affect the relationships of central concern. It did not seem worthwhile, however, to introduce a large number of intermediary populations, when it was not proposed to model this variation.

Although in the course of fertility decline the ages of women bearing children usually change from being widely distributed with a late mean to an earlier and narrower schedule, exceptions to this pattern exist. Therefore the coefficients for maternal orphanhood were estimated using the full set of fertility distributions, resulting in a sample of 96 populations. In contrast, the male fertility schedules assembled by Paget (1988) suggest a strong relationship between the α_f and β_f parameters of the relational Gompertz model. Late ages at childbearing are mainly found in polygynous societies that also have wide male fertility distributions. Therefore, no populations with early, but very wide, or late, but very narrow, fertility distributions were included in the set used to model the relationship between paternal orphanhood and male mortality. A sample of 80 populations results.

The model used to predict female mortality takes the same form as those published previously:

$$_np_{25} = \lambda_0(n) + \lambda_1(n)\overline{M} + \lambda_2(n)_5S_{n-5}$$

The equivalent equation, however, sometimes produces poor estimates of male mortality. Better results can be obtained, especially for populations with light mortality and late ages at childbearing, by including information on a second age group in the model. It is demonstrated in section 2.1 that use of weighting factors, rather than multipliers, increases the robustness of the orphanhood method to variation in the age pattern of mortality. The model proposed here for males combines this advantage of the weights with those of a regression based approach.[2] While the proportion of respondents in two adjoining age groups is highly correlated, the additional parameter captures the effect of variation in the `slope' of mortality. Thus, the specification of the model used to estimate male mortality is:

$$_np_{35} = \lambda_0(n) + \lambda_1(n)\overline{M} + \lambda_2(n)_5S_{n-5} + \lambda_3(n)_5S_n$$

The relative errors in the estimated survival ratios increases rapidly at older ages. For this reason, no coefficients are presented for ages, *n*, of over 50 years for female mortality or of over 40 years for male mortality.

---

[2] If the weighting factors are thought of as embodying two stages, first the application of a multiplying factor, appropriate for age *n*, to convert the measures of orphanhood into measures of life table survival and second the averaging of these two measures, then:

$$\begin{aligned} _np_b &= W_n \cdot {}_5S_{n-5} + (1-W_n)_5S_n \\ &= w_n.f_n \cdot {}_5S_{n-5} + (1-w_n)f_n \cdot {}_5S_n \end{aligned}$$

If an intercept term is added, one obtains:

$$\begin{aligned} _np_b &= w_n.(a_n+f_n \cdot {}_5S_{n-5}) + (1-w_n)(a_n+f_n \cdot {}_5S_n) \\ &= a_n + w_n\, f_n \cdot {}_5S_{n-5} + (1-w_n)f_n \cdot {}_5S_n \\ &= \lambda_0(n) + \lambda_1(n)_5S_{n-5} + \lambda_2(n)_5S_n \end{aligned}$$

Thus, a regression model using information on two adjoining age groups can be thought of as equivalent to the use of weighting factors with an intercept term.

The size of the errors in estimates of α for males that result, when mortality is estimated from further simulated populations by means of the procedures presented here, is illustrated for a range of populations in table 4.4. While these results can be compared with those shown in table 4.2, they are not exactly equivalent. First, other things being

Table 4.4 Errors in regression based estimates of mortality (α) from paternal orphanhood in populations with different characteristics, selected mean ages at childbearing ($\overline{M}$) and central ages of respondents (n).

		M=31			M=34			M=37		
		n=10	n=25	n=40	n=10	n=25	n=40	n=10	n=25	n=40
Mortality										
α=	0	-0.065	0.038	0.060	-0.033	-0.018	0.009	0.015	-0.020	-0.060
	-0.8	0.066	0.001	-0.053	0.006	0.009	-0.001	0.004	0.006	0.004
	-1.2	0.221	-0.009	-0.033	0.060	-0.081	0.043	-0.066	-0.125	0.055
β=	0.6	0.062	-0.023	-0.017	0.000	-0.069	0.003	-0.032	-0.113	-0.016
	1.4	-0.083	-0.037	-0.087	-0.029	0.018	-0.107	0.054	0.046	-0.130
Growth rate										
r=	0.5%	-0.011	0.000	-0.039	-0.006	0.031	-0.022	0.052	0.042	-0.048
	3.5%	-0.030	-0.028	-0.045	-0.034	-0.009	-0.023	-0.025	-0.013	-0.026
Fertility										
β=	0.675				0.081	0.088	-0.049	0.095	0.039	-0.070
	1	0.012	0.024	-0.043	-0.021	0.010	-0.024	-0.030	-0.009	-0.016
	1.75	-0.044	-0.052	-0.046	-0.099	-0.081	-0.006			

equal, the regression method should yield better estimates in low mortality populations and worse ones when mortality is high, as it was derived from populations with an average life expectancy at birth 12 years greater than that in the General Standard. Thus, except when their impact is being assessed, the mortality parameters used to generate the data are set to α = -0.4 and β = 0.95, rather than to 0 and 1 respectively. Secondly, while the fertility model used to derive the weighting factors has a fixed shape, the shape of the distributions generated by the Gompertz model varies systematically as the timing of fertility changes. Except when the β_f parameter is explicitly varied, the parameters used in table 4.4 are α_f = 0.6 and β_f = 1.3, for populations with a mean age of childbearing of 31 years, α_f = 0 and β_f = 1, for mean ages at childbearing of 34 years and α_f = -0.3 and β_f = 0.85, for mean ages at childbearing of 37 years. Thirdly, the coefficients for males yield poor results from data on the youngest age groups when male childbearing is unusually early or late. They perform well though, for mean ages at childbearing of 31 to 37 years, a range that encompasses most developing country populations.

Bearing these *caveats* in mind, the estimates of male mortality produced using the new regression method and shown in table 4.4 are, as expected, significantly more accurate than those yielded by the weighting factors at extreme levels of mortality. In particular, the errors in the estimates obtained from older respondents are much smaller. On occasion, the new approach produces less accurate results than the weighting factors in populations with an unusual age pattern of childbearing, but the errors remain reasonably small. Estimates of female mortality produced using the new regression coefficients are also more accurate than those obtained from the weighting factors in populations subject to extreme levels and patterns of mortality. As earlier research has suggested (UN, 1983), the weighting method performs quite well for female mortality at young ages, but a regression based method produces better results from data on older respondents.

When the new coefficients are applied to real data and the results are compared with those from other variants of the orphanhood method, the expected pattern of differences is found. The estimates of female mortality are similar to those obtained using the weighting factors or other regression methods. When mortality is light, however, the majority of the estimates indicate lower mortality than those arrived at using the weighting factors. At the moderate levels of adult mortality now prevailing in many developing countries, the regression method of estimating mortality from paternal orphanhood usually yields estimates of life expectancy at age 15 that are about one, or sometimes two, years higher than is indicated by estimates made using the weighting factors.

Chapter 5

ORPHANHOOD IN ADULTHOOD

5.1 Background

In a paper published in 1981, Zlotnik and Hill present procedures whereby indirect methods of estimation can be applied to hypothetical cohorts constructed with data from two censuses or surveys conducted five or ten years apart. Knowledge of these techniques has been disseminated widely through their incorporation in the UN (1983) manual on indirect estimation. The methods avoid the complexities introduced into analysis and interpretation by the impact of trends in the variable of interest (fertility or mortality) on measures based on respondents' lifetime experience. Instead equivalent measures are calculated from the experience of five-year age cohorts during the interval between two enquiries and estimates of fertility or mortality are made from these.

Procedures for analyzing recent experience in this way can provide more up-to-date demographic estimates than those based on respondents' lifetime experience. They also have the advantage that, if events occurring during the inter-survey period are fully reported, omission of more distant events will have little impact on the results. In addition, because they eliminate the impact of trends and at least some under-reporting of events, such procedures can reveal inconsistencies in census and survey data that would not be apparent if information was only available from a single enquiry. On the other hand, as Zlotnik and Hill acknowledge, a major limitation of the approach is that measures of cohort experience in between two enquiries are sensitive to differential reporting and sampling errors in the characteristic of interest. Furthermore, they can be biased severely by age reporting errors, which will usually mean that the groups of respondents being compared are not fully equivalent.

One of the indirect techniques that Zlotnik and Hill (1981) apply to data on hypothetical cohorts is the maternal orphanhood method of estimating adult mortality. More recently, Chackiel and Orellana (1985) have pointed out that, by adding questions about the date of, or interval since, parental deaths to the usual ones about parental survival, one can collect data in a single enquiry that can be used to make up-to-date estimates in the way proposed by Zlotnik and Hill. If the dates when parents died are reported reasonably accurately, these data can be used to reconstruct the proportion of respondents who had living parents five or ten years earlier. From these successive crosssections, one can construct period measures of the incidence of orphanhood that are formally identical to those generated from data collected in a series of separate enquiries.

Zlotnik and Hill's procedure is taken up in this chapter, together with a related procedure, described by Preston and Chen (1984), based on recent generalizations of stable population theory (Preston and Coale, 1982). The discussion extends that in an earlier paper (Timæus, 1986) to consider paternal orphanhood and analysis of the dated reports experimented with by Chackiel and Orellana. It presents a new and simpler method for arriving at reliable estimates when data on children are biased by under-reporting of orphanhood.

5.2 Methods of analysis

The principles involved in the procedures suggested by Zlotnik and Hill (1981) are straightforward. They are described here only as they apply to the estimation of adult mortality from orphanhood. If two surveys have been conducted five years apart, the proportion of respondents with either living mothers or living fathers in each five-year age group, according to the first survey, can be compared with the equivalent proportion for respondents five years older, five years later. Assuming that the mortality and propensity to migrate of respondents is unrelated to the mortality of their parents, the ratio of each pair of proportions reflects adult mortality during the inter-survey period. These cohort changes can be chained together to obtain a set of proportions that also reflect mortality over the five-year interval between the enquiries. This represents, for a synthetic birth cohort, the proportion of the population that would have living mothers or fathers, depending which parent the data being analyzed refer to, at the average level of mortality prevailing during the period between the surveys. Thus, if ${}_5S_x(t)$ is the proportion of respondents with a living mother or father at time *t* and ${}_5S_x(\tau)$ is the equivalent for a hypothetical cohort, calculated from, in this case, data referring to times *t* and t+5, then:

$$ {}_5S_0(\tau) = {}_5S_0(t+5) \qquad (5.2.1a) $$

and

$$ {}_5S_x(\tau) = {}_5S_{x-5}(\tau) \,.\, {}_5S_x(t+5)/{}_5S_{x-5}(t) \qquad (5.2.1b) $$

The proportion of respondents with living mothers or fathers in a synthetic cohort can be analyzed using methods developed for data referring to real cohorts (see sections 2.1, 2.3 and 4.5). Extension of the procedure to two enquiries conducted ten years apart is simple.

The generalization of the relationships between age structure, growth and mortality stated by stable population theory to all populations (Preston and Coale, 1982) provides the basis for an equivalent, but more convenient, procedure to that described by Zlotnik and Hill (Preston and Chen, 1984). Stationary measures of orphanhood can be obtained from observed ones by adjusting the latter to remove the impact of past trends in mortality as reflected in the growth rates by age in the proportions orphaned. In any closed population defined by age (Preston and Coale, 1982):

$$N(0) = N(a) \,.\, e^{\int_0^a \mu(z)\,dz} \,.\, e^{\int_0^a r(z)\,dz} \tag{5.2.2}$$

where N(a) is the number of individuals aged *a* and $\mu(z)$ is the force of mortality and r(z) the rate of growth at age *z*. Attrition of the population of individuals with living mothers or of the population with living fathers, *NO*, can be decomposed into the mortality of the parents and the mortality of the population itself:

$$NO(0) = NO(a) \,.\, e^{\int_0^a \mu(z)\,dz} \,.\, e^{\int_0^a \gamma(z)\,dz} \,.\, e^{\int_0^a r^{NO}(z)\,dz}$$

where $\gamma(z)$ represents the instantaneous rate of orphanhood and $r^{NO}(z)$ the rate of growth of the population with living parents at age *z*. Assuming that orphans and the rest of the population have identical mortality and using the fact that $N(0) \equiv NO(0)$, division of the expression for non-orphans by that for the total population produces:

$$e^{-\int_0^a \gamma(z)\,dz} = \frac{NO(a) \,.\, e^{\int_0^a \{r^{NO}(z) - r(z)\}\,dz}}{N(a)} \tag{5.2.3}$$

Taken as a whole, the left-hand term represents the stationary probability of an individual aged *a* having a living mother or father, $S(a,\tau)$. With survey data it is more convenient to work with the rate of increase in the proportion of the population with living parents, $r^S(z)$, than with its equivalent, the difference between the rates of increase of the non-orphaned and total populations. If the interval between two surveys is *h* years, then, in discrete form:

$${}_5S_x(\tau) = {}_5S_x(\bar{t}) \,.\, e^{5\sum_{z=0}^{x-5} {}_5r_z^S(\bar{t}) + 2\,5\,{}_5r_x^S(\bar{t})} \tag{5.2.4a}$$

where:

$${}_5S_x(\bar{t}) = \sqrt{{}_5S_x(t).{}_5S_x(t+h)} \tag{5.2.4b}$$

and:

$${}_5r_x^S(\bar{t}) = \ln\{{}_5S_x(t+h)/{}_5S_x(t)\}/h \tag{5.2.4c}$$

Thus, this approach uses changes in orphanhood experienced by each age group to estimate period measures, while Zlotnik and Hill use the changes experienced by each age cohort for the same purpose. Both approaches can be applied readily to the analysis of dated information on orphanhood collected in a single enquiry, of the type proposed by Chackiel and Orellana (1985). Adjustment using age-specific rates of increase in parental survival has the advantage, when the data come from two enquiries, that it can be applied easily to any inter-survey interval *h* and not just to five- or ten-year periods. The calculations involved are fairly straightforward and are illustrated, using data from Kenya, in table 5.1.

Table 5.1 Estimation of the stationary proportion in each age group with living mothers from two enquiries, using the rate of increase in parental survival, 1969 and 1979 Censuses of Kenya.

Age Group x-x+5	Proportion not orphaned 1969 ${}_5S_x(t)$	Proportion not orphaned 1979 ${}_5S_x(t+h)$	Annual growth rate ${}_5r^S_x(\bar{t})$	Cumulated growth rate $\Sigma_5 r^S_x(\bar{t})$	$e^{\Sigma_5 r_x(\bar{t})}$	Average proportion not orphaned ${}_5S_x(\bar{t})$	Adjusted proportion not orphaned ${}_5S_x(\tau)$
0- 4	0.9922	0.9932	0.0001	0.0002	1.0002	0.9927	0.9929
5- 9	0.9806	0.9853	0.0005	0.0017	1.0017	0.9829	0.9846
10-14	0.9644	0.9733	0.0009	0.0052	1.0052	0.9688	0.9739
15-19	0.9304	0.9515	0.0022	0.0131	1.0132	0.9409	0.9533
20-24	0.8777	0.9107	0.0037	0.0279	1.0283	0.8940	0.9194
25-29	0.8068	0.8588	0.0062	0.0528	1.0542	0.8324	0.8775
30-34	0.7135	0.7745	0.0082	0.0889	1.0929	0.7433	0.8124
35-39	0.6276	0.6927	0.0099	0.1341	1.1435	0.6593	0.7540
40-44	0.5088	0.5798	0.0131	0.1915	1.2110	0.5431	0.6577
45-49	0.3961	0.4723	0.0176	0.2681	1.3075	0.4325	0.5655
50-54	0.2646	0.3286	0.0217	0.3662	1.4423	0.2949	0.4253
55-59	0.2022	0.2323	0.0139	0.4550	1.5762	0.2167	0.3416
60-64	0.1260	0.1388	0.0097	0.5139	1.6718	0.1322	0.2210

Source: Timæus (1986).

It is shown in Timæus (1986) that changes in the proportion of each cohort with a living mother between two surveys conducted five (or ten) years apart (ie. ${}_5S_{x+5}(t+5)/{}_5S_x(t)$) are measures of the mothers' mortality that can be examined directly, without being chained to form a synthetic cohort. Although these measures are rather erratic, estimates obtained from recent increases in orphanhood among young adults consistently suggest higher mortality than either the Zlotnik and Hill or Preston and Chen approaches (Timæus, 1986). This probably stems from under-reporting of the extent of orphanhood among young children, presumably because of the adoption effect. Chaining cohort changes or growth rates transmits this bias to older age groups. Working directly with the ratios of the proportions of young adults with living parents reduces such errors. Unfortunately, the estimation procedure used by Timæus (1986) to analyze these ratios is inconvenient to apply. It involves working backwards from a life table and an estimate of the mean age of childbearing to estimates of orphanhood in order to reproduce the observed cohort changes in parental survival. Moreover, like Zlotnik and Hill's procedure, it requires extrapolation from, or interpolation between, the two sets of data if they were not collected exactly five or ten years apart. The following paragraphs outline a more flexible way of estimating period adult mortality from the incidence of orphanhood experienced by young adults.

The estimation procedure proposed here involves constructing a synthetic cohort, based at

20 years of age, from the data on orphanhood collected in two enquiries. This cohort indicates the proportion of the adult population whose mothers or fathers would remain alive, amongst those who had a living mother or father at exact age 20, at current levels of mortality. Such a synthetic cohort can be constructed using only the more reliable data on young adults. The cohort is based at 20 years, rather than 15 years, for two reasons. First, this minimizes the possibility of underestimating orphanhood at the base age and consequently overestimating subsequent orphanhood and adult mortality. Secondly, it makes it possible to apply the method to data collected in surveys such as those conducted by the DHS programme, in which only women aged 15 to 49 are asked about orphanhood.

It is possible to construct such a synthetic cohort either from cohort changes in orphanhood, as is proposed for all ages by Zlotnik and Hill (1981), or by using the rates of increase in parental survival to adjust the proportions of respondents with living parents. With identical data, the two procedures are equivalent but the latter approach is, in general, more convenient when the information comes from two surveys.

It is fairly clear that the population above any given age can be treated as a self-contained one and that the relationship between age structure, increase and mortality stated in equation 5.2.2 will continue to hold for such populations. Thus, using the notation already established, the equivalent relationship for the population aged 20 and over, to that given in equation 5.2.3, is:

$$\frac{NO(20).e^{-\int_{20}^{a} \gamma(z)\,dz}}{N(20)} = \frac{NO(a).e^{\int_{20}^{a} \{r^{NO}(z) - r(z)\}\,dz}}{N(a)}$$

for a > 20. In discrete form, expressing the rates of increase in proportionate terms, this becomes:

$$\frac{{}_5S_x(\tau)}{S(20,\tau)} = \frac{{}_5S_x(\overline{t}) \,.\, e^{5\sum_{z=20}^{x-5} {}_5r^S_z(\overline{t}) + 2\,5\;{}_5r^S_x(\overline{t})}}{S(20,\overline{t})} \qquad (5.2.5a)$$

where the proportion of respondents orphaned at age 20 can be estimated as:

$$S(20,\overline{t}) = \sqrt{{}_5S_{15}(\overline{t}).{}_5S_{20}(\overline{t})} \qquad (5.2.5b)$$

In order to simplify the estimation of life table measures of mortality from these proportions, a new set of coefficients is required. These are presented in appendix 2. The estimation equation for maternal orphanhood after age 20 is analogous to that used for orphanhood since birth. The only difference arises from the fact that the mothers are 20 years older at the beginning of the period of exposure. Therefore the base age, *b*, is set to 45 years. The equation used to make the estimates takes the form:

$$l(25+n)/l(45) = \lambda_0(n) + \lambda_1(n)\overline{M} + \lambda_2(n)\,{}_5S_{n-5}(\tau)/S(20,\overline{t}) \qquad (5.2.6)$$

While, for female mortality, very little predictive power is gained by including information on a second age group, for male mortality this is even more important than in the basic method. Survival is calculated from a base age of 55 years, using a model of the form:

$$l(35+n)/l(55) = \lambda_0(n)+\lambda_1(n)\overline{M}+\lambda_2(n)\,{}_5S_{n-5}(\tau)/S(20,\overline{t})+\lambda_3(n)\,{}_5S_n(\tau)/S(20,\overline{t}) \qquad (5.2.7)$$

The coefficients were estimated using an identical set of simulated populations to those employed to obtain the revised coefficients for estimating mortality from orphanhood data of the normal type. Mortality is represented by relational logit life tables, based on the General Standard; the

Table 5.2 Estimation of mortality from maternal orphanhood above age 20 between two enquiries, 1969 and 1979 Censuses of Kenya.

Age Group x-x+5	Average proportion not orphaned ${}_5S_x(\overline{t})$	Annual growth rate ${}_5r^S_x(\overline{t})$	$e^{\Sigma_5 r_x(\overline{t})}$	Adjusted proportion not orphaned since 20 ${}_5S_x(\tau)/S(20,\overline{t})$	n	$\frac{l(25+n)}{l(45)}$ ($\overline{M}$=26.7)	α
$S(20,\overline{t})$				0.9172			
20-24	0.8940	0.0037	1.0093	0.9838	25	0.9732	-0.775
25-29	0.8324	0.0062	1.0347	0.9390	30	0.9324	-0.731
30-34	0.7433	0.0082	1.0727	0.8694	35	0.8683	-0.682
35-39	0.6593	0.0099	1.1223	0.8068	40	0.8162	-0.800
40-44	0.5431	0.0131	1.1885	0.7038	45	0.7178	-0.832
45-49	0.4325	0.0176	1.2832	0.6051	50	0.6158	-0.961

age structure of parents is a stable one, generated with r = 2 *per cent*; and fertility is represented by relational Gompertz models (see table 4.3). The average size of the errors in the predicted measures depends largely on the age of the parents and so, once again, coefficients are presented for ages, *n*, of up to 50 years for female mortality and 40 years for male mortality. Estimation of female adult mortality from stationary proportions of those with living mothers at age 20, whose mothers remain alive, by means of age-specific growth rates in parental survival is illustrated in table 5.2, using the Kenyan data presented in table 5.1.

The accuracy of the estimation procedures for this form of orphanhood data is similar to that of those proposed for the basic orphanhood method in section 4.5. The results are slightly less sensitive to the age pattern of mortality, as survival is being measured over a more restricted range of ages. On the other hand, the errors that result from unusual age patterns of childbearing are a little larger. Once again, the coefficients for paternal orphanhood perform best in populations where the mean age at childbearing of men lies in the range 31 to 37 years.

5.3 Applications

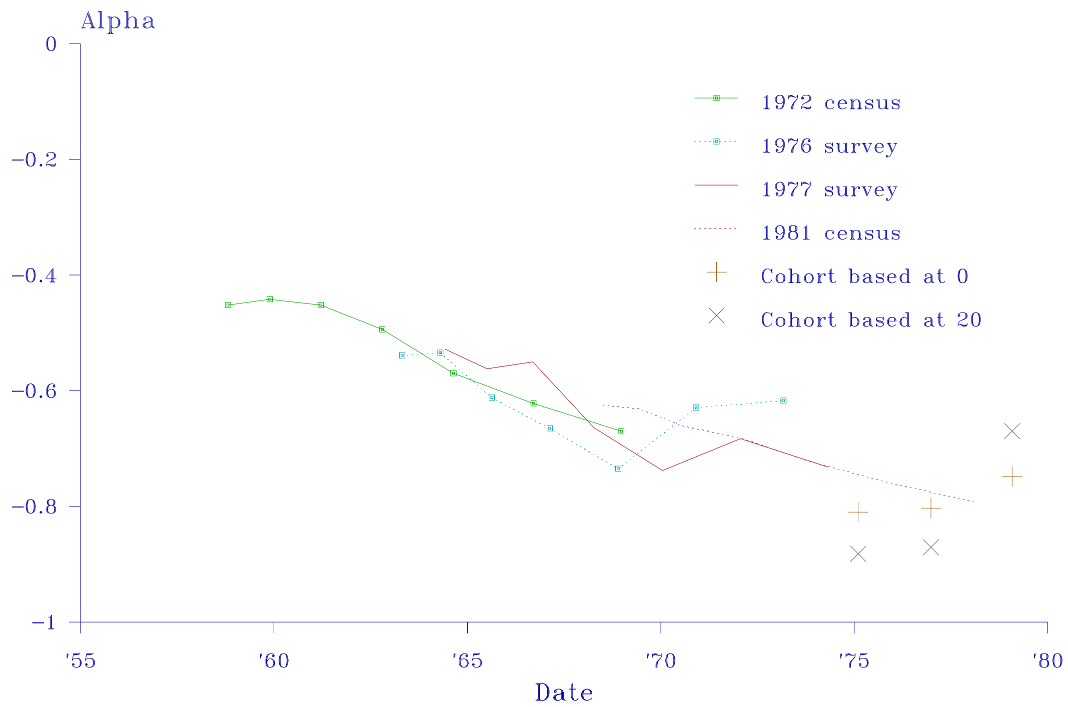

Figure 5.1: Female adult mortality (α) estimated from cohort and period orphanhood data, Peru.

Estimates of adult female mortality in Peru are shown in figure 5.1. The estimates are from cohort data on lifetime orphanhood and from synthetic cohort data based at ages 0 and 20, calculated using equations 5.2.4 and 5.2.5. The results are presented in terms of α, the level parameter of the family of relational logit model life tables based on the General Standard (Brass, 1971b). Orphanhood data from four different enquiries are considered, making it possible to calculate three sets of period measures. The first set is based on changes in orphanhood between the 1972 Census and 1977 Peru Fertility Survey, the second on the 1972 and 1981 censuses and the third on the 1976 Demographic Survey and 1981 Census. The basic data are taken from Moser (1985) and Timæus (1986), but the final estimates made from data on lifetime orphanhood differ slightly from those published previously, as they have been recalculated using the new coefficients described in section 4.5. The agreement between the four series of estimates from cohort data is close and suggests that there has been a moderate but steady decline in the level of female adult mortality in Peru, with life expectancy at age 15 rising from about 51 to 56.5 years between 1960 and the end of the 1970's.

The period estimates in figure 5.1 are averages of individual measures obtained from respondents of different ages. The estimates from orphanhood since birth are based on respondents aged 15 to 39 and those from orphanhood after age 20 on respondents aged 15 to 34. Within these age ranges, all the estimates are very similar, though those from older age groups indicate increasingly light mortality, presumably because respondents tend to exaggerate their ages (Timæus, 1986).[1] The estimates from data on synthetic cohorts constructed from birth onward were calculated using the coefficients presented in appendix 2, but are otherwise equivalent to those published by Timæus (1986). The earliest of these measures indicates slightly lighter mortality than the trend in the cohort measures, but the estimate for the period around 1979 slightly higher mortality. The discrepancies correspond to about half a year's difference in life expectancy at age 15. The middle estimate, which was obtained from the two censuses, agrees well with the results from the basic orphanhood method. The estimates obtained from synthetic cohort data based at age 20, using equation 5.2.6 and the coefficients presented in appendix 2, follow the same pattern, but diverge more from the trend in the cohort data. For the first period, data on orphanhood after age 20 yield an α of -0.88, compared with one of -0.74 indicated by cohort data from the 1981 Census, while, in the last period, the corresponding estimates are -0.67 and -0.82 respectively. In addition, application of this method to data from the two censuses also produces a rather low estimate of mortality. As there is little reason to doubt the accuracy of the basic data on orphanhood in Peru, it seems likely that the period measures are affected by differences between the various censuses and surveys in the quality of the results obtained. In both surveys, but not the censuses, the estimates obtained from cohort data on respondents aged 15 to 25 drop appreciably below the general trend of the series. One can hypothesize that biases in the composition of the samples may be responsible. Thus, changes in orphanhood between 1972 and 1976 underestimate the growth in the proportion of the population with living parents, producing overestimates of mortality. By the same token, the period estimates for 1977 to 1981 are biased in the opposite direction. The small discrepancies between the sets of cohort data have a substantial impact on the period estimates based on differences in the proportion of respondents with living parents between successive enquiries. Data on orphanhood after age 20 are more affected than those on orphanhood since birth, because they are calculated solely from the biased data concerning young adults.

Estimates of female adult mortality obtained from orphanhood data collected in the 1969 and 1979 Censuses of Kenya are shown in figure 5.2. This is the first of the cases considered by Timæus (1986), in which the basic orphanhood method has yielded inconsistent results from two successive enquiries. Once again, measures obtained from cohort and period data about lifetime

---

[1] The pattern of the estimates for Kenya, shown in table 5.2, is fairly typical of all the sets of data considered here. In the smaller surveys, the estimates for the first age group are rather erratic.

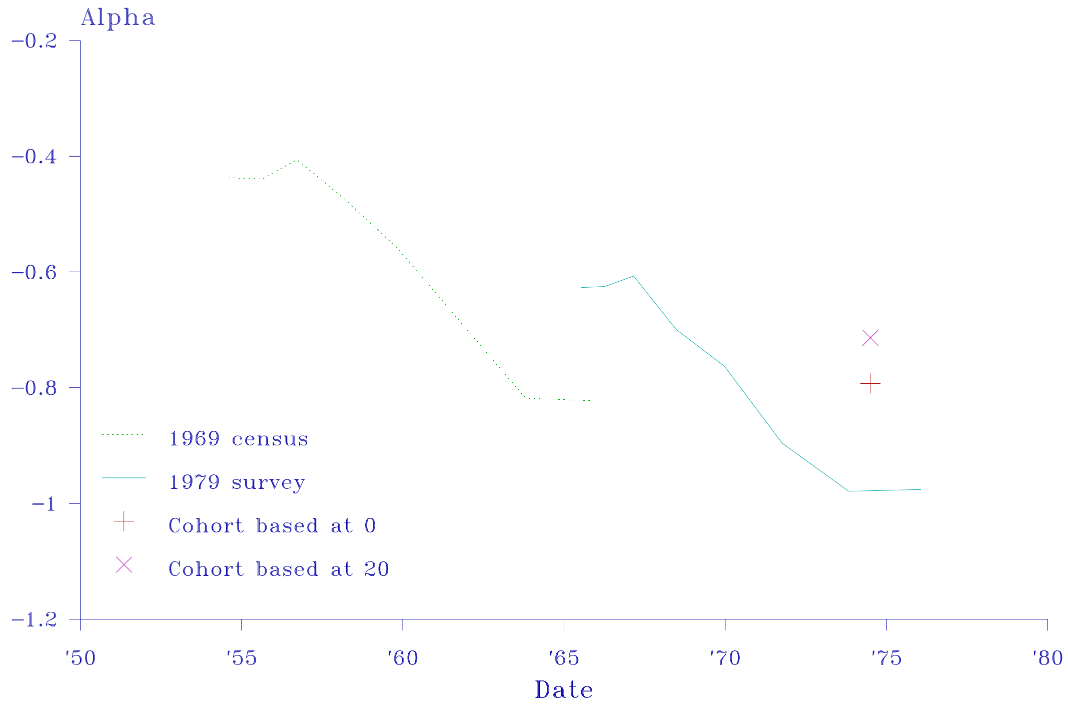

Figure 5.2: Female adult mortality (α) estimated from cohort and period orphanhood data, Kenya.

orphanhood have been estimated using the coefficients described in chapter 4. The data on which the estimates are based are presented in table 5.1 and calculation of the estimates from orphanhood after age 20 is shown in table 5.2. In Kenya, the period estimates indicate considerably higher mortality than those from cohort data on younger respondents in 1979. Moreover, an appreciably lighter estimate of mortality is obtained from orphanhood after age 20, than from the period data on lifetime orphanhood. The new method suggests that life expectancy at age 15 in Kenya in the mid-1970's was 54.9 years, while, if the 1979 Census data were taken at face value, one would obtain an estimate of 58.8 years for the same index of mortality. It seems likely that orphanhood of young respondents was under-reported in the censuses. The period data on lifetime orphanhood reduce the impact of this bias on the results and the data on orphanhood after 20 are even less affected, if at all. The estimates from older respondents in each census and the estimate from orphanhood after age 20 fall approximately in line, suggesting that there was a moderate decline in the mortality of adult women in Kenya between the mid-1950's and mid-1970's.

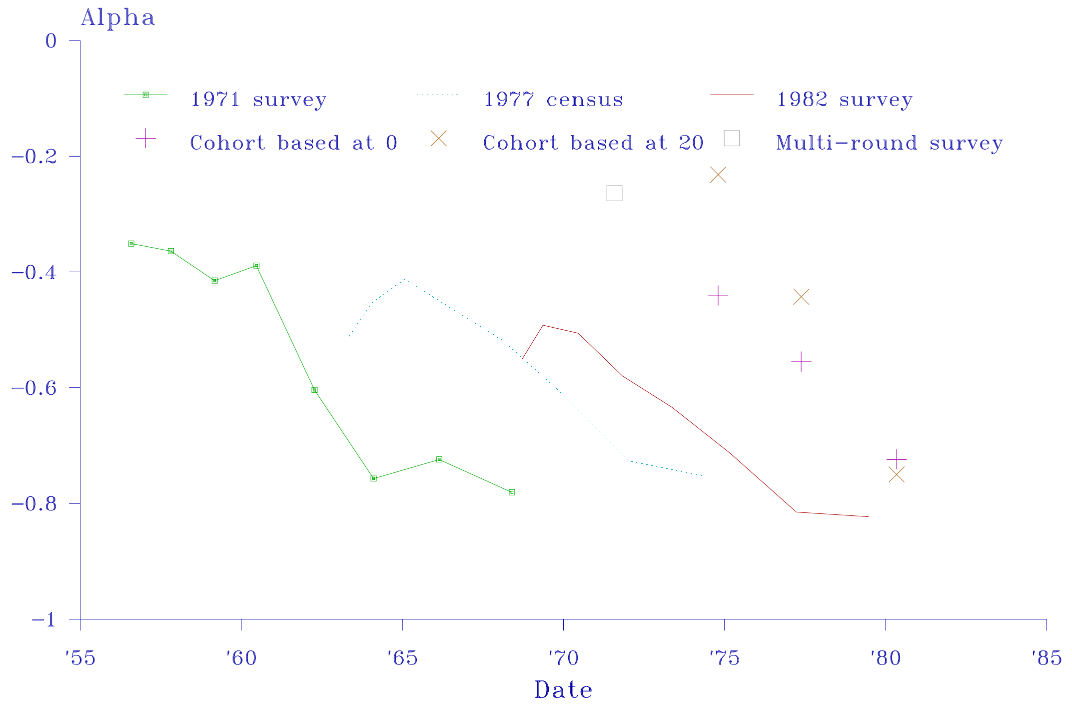

Figure 5.3: Female adult mortality (α) estimated from cohort and period orphanhood data, Malawi.

Results for Malawi are presented in figure 5.3. Malawi is another East African country in which estimates of adult mortality from orphanhood data collected in different enquiries have yielded inconsistent findings. In figure 5.3, the two sets of data considered by Timæus (1986) are supplemented by data on maternal orphanhood collected in the Demographic Survey of 1982 (National Statistical Office, 1987). This makes it possible to calculate three sets of period measures of adult female mortality. While the period estimate obtained from data on lifetime orphanhood for the years around 1975 is very similar to that published by Timæus (1986), the estimate from orphanhood after age 20 indicates substantially higher mortality. The estimates for the period around 1980 agree better, but indicate much lighter mortality. The middle pair of estimates, based on changes in orphanhood between 1971 and 1982, fall in between these extremes. While the results for Malawi are rather erratic, they can readily be explained by a combination of errors due to the adoption effect and differences between the enquiries in the quality of the data. The estimates obtained from cohort data on older respondents in each of the three enquiries fall in line and suggest that a slow decline in adult female mortality occurred between the mid-1950's and the end of the 1960's. If this trend is extrapolated forward to the 1970's, it suggests that the first and last sets of period estimates are biased in opposite directions.

The middle set of estimates, obtained from data collected in the two surveys, seems more plausible. It is likely that age exaggeration and other reporting errors were more serious in the 1977 Census than the 1971 and 1982 surveys, leading the first set of period measures to overestimate mortality and biasing the last set downward. As in Peru, the estimates from orphanhood after age 20 are affected more than those made from period data on orphanhood since birth. Superimposed on these errors are those due to the adoption effect, which affect period estimates from orphanhood since birth, but not those from orphanhood after age 20. Thus, the two estimates for the period 1977 to 1982 agree because they are both biased downward for different reasons but the estimate of mortality from orphanhood after age 20, calculated from data collected in 1971 and 1982, is probably reasonably accurate.

Adult female mortality in Malawi can also be estimated directly from data collected in the Population Change Survey during 1971 (National Statistical Office, 1973). These data are very erratic and the estimate shown in figure 5.3 has been obtained by smoothing the mortality rates for the age range 20 to 59 years by fitting a 2-parameter logit model life table. The direct estimate suggests that most of the orphanhood based results are biased downward. This estimate, however, and that from orphanhood after age 20 according to the two surveys, agree quite well, suggesting that a moderate decline in adult mortality occurred in Malawi in the early 1970's. According to these figures, the life expectancy of women at age 15 rose from around 47.5 to 50.5 years between 1971 and 1977. Taken as a whole, the estimates for Malawi provide rather convincing evidence that the inconsistencies in the cohort estimates arise from under-reporting of orphanhood, concentrated among children, that this bias also affects period estimates based on lifetime data, but not those based on orphanhood after age 20, and that estimates from orphanhood after age 20 are particularly sensitive to differences between the enquiries on which they are based, in the accuracy of age reporting.

Results from a recent survey of parts of the Buganda and Basoga regions of Southern Uganda are shown in figure 5.4. While only one set of orphanhood data is available (see appendix 3), the survey included questions on the dates when parents died. Thus, one can calculate the proportion of respondents with living parents five and ten years earlier and estimate mortality from orphanhood during the two intervening periods. Estimates are presented of adult male mortality, produced using equation 5.2.7 and the coefficients in appendix 2, as well as of female mortality. As in Kenya and Malawi, orphanhood data from the 1969 and 1980 censuses of Uganda indicate implausibly steep declines in mortality (Timæus, 1990). This problem also seems to have affected the smaller scale survey considered here. The estimates for males, in particular, suggest a very steep decline in mortality to an impossibly low level. As in the last two

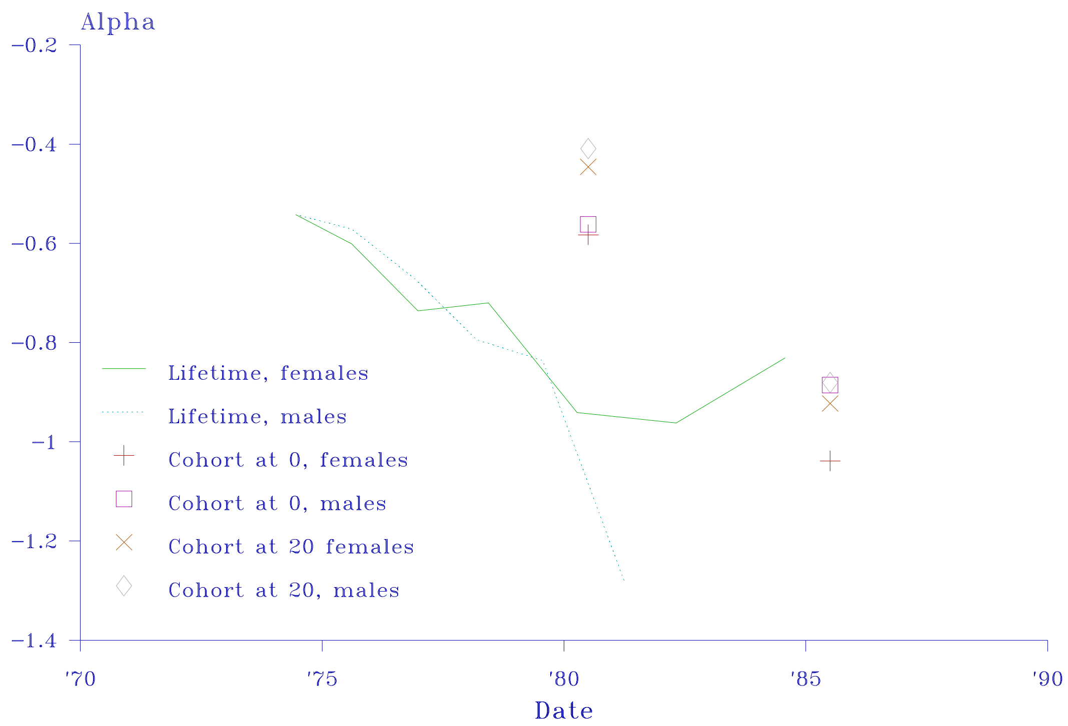

Figure 5.4: Adult mortality (α) estimated from cohort and period orphanhood data, Buganda.

applications, the period estimates indicate higher mortality than those based on cohort data, especially for men. Moreover, except for the estimate of male mortality for the period 0 to 4 years before the survey, estimates from orphanhood after age 20 indicate higher mortality than those made from data on synthetic cohorts based at birth. This once again suggests that the steep decline in the cohort estimates is the product of the adoption effect and that this bias in the data has little impact on estimates made from orphanhood after age 20. However, although the period estimates indicate much higher mortality than the cohort data, they also suggest that a very rapid decline in adult mortality has occurred in this area. Period estimates made using data collected in a single enquiry about dates of orphanhood are not subject to the errors resulting from differences in the quality of age reporting and sample selection biases that affect estimates made from data collected in two enquiries. They are subject to an equivalent problem. If retrospectively reported dates of orphanhood tend to be shifted backward in time, the incidence of orphanhood in the recent past will be understated and orphanhood in the period 5 to 9 years before the survey may be overstated. Such errors will exaggerate the apparent fall in the level of adult mortality. Equally, if dates of orphanhood are shifted towards the date of interview, the estimates obtained will indicate too small a decline in mortality. While there is no direct evidence of this, the very steep decline in mortality yielded by these Ugandan data suggest that they may be affected by the former bias. Combining the data on orphanhood after age 20 for the two periods, produces estimates of α of -0.757 for males and -0.714 for females. Although these indices may be biased by shifting of dates of orphanhood across the ten-year boundary, the most robust estimate of mortality that can be obtained from these data is that life expectancy at age 15 for both sexes was about 55 years in the early 1980's.

These analyses of data from three East African countries provide strong evidence that the inconsistencies and errors in mortality estimates from orphanhood data collected in this part of the world stem from under-reporting of orphanhood, concentrated among young respondents. When either two sets of data on a population are available or questions have been asked about the dates when parents died, synthetic cohort data on orphanhood after age 20 are a valuable way of measuring adult mortality, that greatly reduces the impact of such reporting errors on the results. Unfortunately, the new method is sensitive to differences in the quality of data between enquiries and to the accuracy of retrospective reporting of dates of orphanhood. However, comparisons of estimates from the new method with those from period and cohort data on lifetime orphanhood provide the analyst with a wealth of material, often allowing the final estimates of adult mortality to be made with some confidence.

Chapter 6

ORPHANHOOD BEFORE AND SINCE MARRIAGE

6.1 Background

The previous chapter discusses the supplementary question about orphanhood, proposed by Chackiel and Orellana (1985), on the date of death of a parent that has died. The importance of this question arises from the potential that it offers for obtaining recent and accurately dated estimates of adult mortality from data on maternal and paternal orphanhood. In addition, if data could be collected of sufficient quality, it might be possible to use the information supplied by different age groups of respondents to make inferences about the age pattern of mortality experienced by their parents. The major problem with this approach is that respondents may not be able to report the dates of death of their parents with sufficient accuracy for the information collected to be of much value. Initial trials of the method in Bolivia, Costa Rica and Honduras have yielded promising results (Chackiel and Orellana, 1985). So far, however, experience with this question is limited. In particular, though the Ugandan data discussed in section 5.3 offer some grounds for optimism, it is unclear whether the approach will prove useful for estimation of adult mortality in poorly educated populations, such as those of many sub-Saharan African countries.

Information that is of potential value as a way of obtaining recent and accurately dated measures of adult mortality from information on orphanhood has also been collected by the DHS programme of surveys. The core questionnaire contains questions that establish whether parents and parents-in-law, that have died, were still alive at the time of the respondent's first marriage. These questions were asked in 15 of the surveys conducted during the first round of the DHS programme. The main reason why the questions were included in the questionnaire was as a filter for further questions, that establish whether respondents lived with their parents(-in-law) after marriage, so that the impact of residential patterns on contraceptive use and fertility can be investigated. In an unpublished note, Brass (1986) points out that this supplementary information on when orphanhood occurred is of potential value for the study of adult mortality. Marriage is an event that can be used to distinguish, for each age group of respondents, more recent parental deaths from those that occurred longer ago. While the information on the timing of deaths is less precise than that yielded by direct questions about the date of death of parents, it may be more accurately reported. Even if respondents cannot remember exactly when their parents died, it seems likely that nearly all of them will be able to report the timing of parental deaths relative to marriage, another event of major significance in their lives.

Brass (1986) points out that the proportion of women with living mothers at marriage is approximately equal to the life table probability of surviving from the mean age at childbearing of women to that age plus the mean age at first marriage of the cohort of women concerned. Trends in this index across cohorts of respondents aged between 25-29 and 45-49 can be compared directly, to estimate the trend in adult mortality over a period some 10 to 30 years before the survey. On the other hand, the proportion of mothers that have remained alive since female respondents married is closely related to the probability of surviving from the sum of the period mean age at childbearing and the cohort mean age at marriage until $n-\overline{m}$ years later, where n is the mean age of respondents and $\overline{m}$ the mean age at marriage of the cohort. In form and characteristics, estimation of life table indices from data about women of this type is equivalent to obtaining mortality estimates from survivorship of first husbands, when men marry at much later ages than their wives. Estimates made from data on orphanhood since marriage will measure more recent mortality than those based on respondents' lifetime experience of orphanhood. In addition, Brass points out, because parental deaths since marriage occurred recently, when respondents were sufficiently old to remember them clearly, such data could be less subject to reporting errors than those concerning the overall level of orphanhood.

In response to Brass's suggestion, the questions proposed by Chackiel and Orellana (1985) were added to the questionnaire used in Burundi in the survey conducted as part of the DHS programme. This makes it possible to compare the results of a direct question on dates of parental deaths with those obtained from information on the relative timing of orphanhood and first marriage. These data have been examined by Makinson (1988), in an unpublished paper. She presents adult mortality estimates from widowhood and lifetime orphanhood data and from synthetic cohort data on orphanhood for the periods 0 to 4 years and 5 to 9 years before the survey. A major problem was encountered with the direct question on the timing of parental deaths in Burundi, that does not appear to have arisen in the Latin American trials of the same question (Chackiel and Orellana, 1985). About 25 *per cent* of orphaned respondents failed to report the date of death of their father or mother. In the majority of these cases, Makinson (1988) is able to calculate whether the parent was, in fact, alive 5 and 10 years earlier from the information on orphanhood since marriage and the question on date of first marriage. She does not attempt to use the data on whether orphanhood occurred before or after marriage as the primary source of information on the timing of parental deaths, along the lines outlined by Brass (1986). The paper concludes that all the data analyzed and methods used yield similar results, corresponding to estimates of life expectancy at age 15 of 47 to 50 years.

This chapter takes up the strategy for analyzing data on the timing of parental deaths relative to respondents' marriages suggested by Brass (1986). Experimentation reveals that a simple procedure that equates orphanhood before marriage with survival from the mean age of childbearing to the sum of that age and the mean age of marriage of respondents can be improved on. Moreover, the age gap between respondents and their parents is too large for it to be possible calculate life table indices from orphanhood since marriage using the coefficients developed for data on the survival of first spouses. Thus, new coefficients are estimated and tested here for estimating adult mortality from orphanhood before and after marriage.

Although the DHS surveys collected individual-level information on women's date of first marriage, future surveys that use the new questions about orphanhood might not ask this question. In addition, retrospectively reported dates of first marriage are often very inaccurate (Goldman *et al.*, 1985). Because of this, the methods developed here use aggregate information on ages at marriage to control for variation between populations in the ages of parents and the length of time that they are exposed to the risk of death. Future surveys seem unlikely to include questions on the timing of deaths of parents-in-law. Moreover, the DHS questionnaire did not ask women about the age of their first husbands. Therefore, no attempt is made to develop a procedure for analyzing these data. Finally, the coefficients developed are intended for the analysis of data supplied by female respondents. In the DHS, only women interviewed individually were asked the relevant questions and, in any case, women usually supply more reliable information than male respondents about orphanhood (see chapter 3).

6.2 Orphanhood since marriage

The derivation of the basic orphanhood method involves the assumption that the mortality of respondents is uncorrelated with the mortality and fertility of their parents. To derive methods for the estimation of adult mortality from orphanhood before and after marriage, further assumptions are made that the probability that a woman marries at any age is unaffected by whether her parents are alive or dead and that her age at marriage is uncorrelated with her parents' ages at the time that she was born. The validity of these assumptions and the impact of breaches of them on the results is discussed in section 6.5. In addition, the mortality estimates obtained apply only to the parents of ever-married respondents. In nearly all developing countries, almost all women aged 25 or more fall into this category and the results can be treated as representative of the whole population.

Using the notation established in chapter 2 (see also appendix 1), the proportion of women aged exactly *z*, with living mothers, depends on the probability that the mothers gave birth at any age *y* and the probability that they survive a further *z* years:

$$S(z) = \int_s^w {}_z v_y \cdot {}_z p_y \, dy$$

If the first marriage rate at age *z* is defined as m(z), then the proportion of ever-married respondents aged *a*, whose mothers were alive at the time of the respondents' first marriage, is:

$$S^m(a) = \int_0^a m(z) \int_s^w {}_zv_{y}\cdot{}_zp_y \, dy \, dz \Big/ \int_0^a m(z) \, dz \qquad (6.2.1)$$

By division of the proportion of ever-married respondents aged *a* with living mothers by the proportion of them who had living mothers at marriage, one obtains the probability that a mother has survived from the time of marriage to the time of interview:

$$\frac{S(a)}{S^m(a)} = \frac{\int_0^a m(z) \, dz \, . \int_s^w {}_av_{y}\cdot{}_ap_y \, dy}{\int_0^a m(z) \int_s^w {}_zv_{y}\cdot{}_zp_y \, dy \, dz}$$

The equivalent relationship, for five-year age groups of respondents, is:

$$\frac{{}_5S_x}{{}_5S_x^m} = \frac{\int_x^{x+5} l(a) \int_0^a m(z) \, dz \, . \int_s^w {}_av_{y}\cdot{}_ap_y \, dy \, da}{\int_x^{x+5} l(a) \int_0^a m(z) \int_s^w {}_zv_{y}\cdot{}_zp_y \, dy \, dz \, da} \qquad (6.2.2)$$

Once the possibility of death between conception and birth is allowed for, equations 6.2.1 and 6.2.2 also express the probability that a woman's father is alive when she marries and the probability, for a five year age group, that he has remained alive since then. Although the weights, ${}_av_y$, applicable to the probabilities of survival from age *y* to y+a differ for fathers and mothers, the marriage distribution is in both cases that of female respondents.

Coefficients for the estimation of male and female adult mortality from orphanhood since marriage are presented in appendix 2. They have been estimated from orphanhood in a set of simulated populations, calculated by evaluating equation 6.2.2. These populations have the same set of age distributions at childbearing and mortality schedules as those on which the coefficients for orphanhood estimation presented already are based (see section 4.5 and table 4.3). The numerical procedures used are also those described in section 4.5. In particular, the integral of ${}_av_{y}\cdot{}_ap_y$ is estimated using equation 2.1.3.

To evaluate equation 6.2.2, however, it is also necessary to model the distribution of women's ages at marriage. It turns out that, for respondents in the age range at which most marriages occur, the proportion orphaned at marriage is extremely sensitive to the exact shape of the first marriage distribution. At these ages it is impossible to control for the age pattern of marriage sufficiently accurately to estimate mortality. Acceptably robust estimates can only be obtained from age groups in which nearly all women are married. For respondents aged 25 and over, in contrast, a simple control for the mean age at marriage allows one to estimate mortality reliably. Almost identical results are obtained, whatever assumptions are made about the shape of the distribution of first marriages.[1] Thus, even if the ages at which women marry have been

---

[1] This can be verified by comparing estimates obtained using the coefficients for orphanhood since marriage, for a mean age at marriage of 20 years, with those obtained from the same data using the coefficients for orphanhood since age 20. The latter coefficients are equivalent to those for orphan-
(continued...)

rising, mortality can be estimated if the mean age at marriage of each cohort is known. In addition, in models of the form that are proposed here, the relationship between the mean age at marriage and the predicted life table measures is close to linear. Thus, coefficients for orphanhood since marriage can be estimated using three marriage distributions in which the mean age at marriage spans the range found in most developing countries. Very little would be gained by including populations in which the variance of the marriage distribution is different.

The distributions used to derive the simulated populations are Coale and McNeil (1972) marriage models in which $\overline{m} = a_0 + 11.36k$ and the incidence of first marriage is a mathematical function of the parameters a_0 and k. The scale factor, *k*, is fixed at 0.55 and the age at which marriages first occur, a_0, is set to 8.75, 13.75 and 18.75 years, producing marriage distributions with means of 15, 20 and 25 years. Levels of orphanhood were calculated for each marriage pattern, for all of the populations defined in table 4.3. Thus, the coefficients for maternal orphanhood since marriage are based on 288 populations and those for paternal orphanhood on 240 sets of data.

Because of the sensitivity of estimates for younger respondents to the age pattern of marriage, the earliest central age of respondents, *n*, for which one can estimate a survival ratio, ${}_np_b$, from data on orphanhood since marriage is 30 years. In contrast to the estimation of female mortality from orphanhood at all ages and from orphanhood after age 20, a model incorporating measures of orphanhood in two adjoining age groups produces more accurate estimates of life table survival from data on orphanhood since marriage, than a model using information on just one age group. The improvement in fit is particularly large for estimates from data on younger respondents. Part of the explanation of why information on an extra age group proves valuable is that it captures the impact of an interaction between the level of orphanhood and the timing of marriage. A model including a term for ${}_5S_x$, however, explains more of the variation in ${}_np_b$ than one that includes an interaction term of the form $\overline{m}.{}_5S_{x-5}$. Moreover, the latter term has a distinctly non-linear relationship with the dependent variable. Thus, the model used to estimate life table measures from the survival of mothers since the respondents married is:

$$l(25+n)/l(45) = \lambda_0(n) + \lambda_1(n)\overline{M} + \lambda_2(n)\overline{m} + \lambda_3(n){}_5S_{n-5}/{}_5S^m_{n-5} + \lambda_4(n){}_5S_n/{}_5S^m_n$$

The only difference in the equation for the estimation of male mortality from paternal orphanhood since marriage arises from the fact that men are, on average, about 10 years older than women at the birth of their children. The estimates are made using:

$$l(35+n)/l(55) = \lambda_0(n) + \lambda_1(n)\overline{M} + \lambda_2(n)\overline{m} + \lambda_3(n){}_5S_{n-5}/{}_5S^m_{n-5} + \lambda_4(n){}_5S_n/{}_5S^m_n$$

(...continued)
hood since marriage in a population in which all women marry at the same age.

Coefficients are presented for estimating survival over the age ranges ${}_{10}p_{45}$ to ${}_{30}p_{45}$ for females and ${}_{10}p_{55}$ to ${}_{20}p_{55}$ for males.

As one might expect, the performance of the coefficients for orphanhood since marriage is broadly similar to that of those for orphanhood since age 20, but the errors in the estimated life table measures tend to be slightly larger. When the coefficients for males and females were used to estimate mortality from further simulated data, including those on populations with extreme characteristics, the errors in the results only represented biases in measures of life expectancy at age 15 of more than two years when mortality was exceptionally heavy. In most of the simulated populations the error in the estimates is much smaller. As for the other orphanhood methods, the coefficients for paternal orphanhood since marriage produce relatively unreliable estimates when the mean age at childbearing of men is unusually early or late.

6.3 Orphanhood before marriage

Estimation of adult mortality from orphanhood before marriage is based on equation 6.2.1. The equivalent relationship for five-year age groups of respondents is:

$$ {}_5S^m_x = \int_x^{x+5} l(a) \int_0^a m(z) \int_s^w {}_zv_{y\cdot z}p_y \, dy \, dz \, da \Big/ \int_x^{x+5} l(a) \int_0^a m(z) \, dz \, da \qquad (6.3.1) $$

For age groups of women who are still marrying, this relationship changes slightly with age. However, in a stable population, exactly the same proportion of women will have been orphaned before marriage in all age groups above the oldest age at which first marriages occur.

Coefficients for the estimation of male and female adult mortality from orphanhood before marriage are also presented in appendix 2. They were derived using the same data and methods as the coefficients for estimating mortality from orphanhood since marriage. The sensitivity of the level of orphanhood since marriage to the age pattern of marriage among women aged less than 25 years is mirrored in data on orphanhood before marriage. Therefore, coefficients are only presented for use with data supplied by respondents older than 25 years. Because the responses of each age cohort reflect the survival of parents over more-or-less the same range of ages at different points in time, there is nothing to be gained by using information on more than one age group to make the estimates. Experimentation with different models again indicates that the relationship between the level of orphanhood and mortality depends on the timing of marriage. For orphanhood before marriage, however, an interaction term between the mean age at marriage and proportion orphaned improves the fit of the model more than for orphanhood since marriage. Moreover, in this case, the interaction term is linearly related to the dependent variable. Thus, the probability of surviving from age 25 to age 45 for women can be estimated from orphanhood before marriage using:

$$l(45)/l(25) = \lambda_0(n) + \lambda_1(n)\overline{M} + \lambda_2(n)\overline{m} + \lambda_3(n)_5S_n^m + \lambda_4(n)\overline{m}_5S_n^m$$

The same considerations apply to the estimation of adult male mortality from paternal orphanhood before marriage. Estimates of the probability of surviving from age 35 to age 55 are made from:

$$l(55)/l(35) = \lambda_0(n) + \lambda_1(n)\overline{M} + \lambda_2(n)\overline{m} + \lambda_3(n)_5S_n^m + \lambda_4(n)\overline{m}_5S_n^m$$

In the other variants of the orphanhood method discussed in this thesis, the accuracy with which mortality can be estimated declines as the age of respondents, and therefore of their parents, increases. For orphanhood before marriage this problem does not arise. The estimates from respondents of all ages are based largely on the experience of young and middle-aged parents. In theory, the final sets of coefficients presented for maternal and paternal orphanhood before marriage in appendix 2 can be used to estimate mortality from the reports of any age group of respondents aged 40 and above. Only the accuracy of reporting about parental deaths and respondents' own ages impose an upper age limit on the data on orphanhood before marriage that can be used to estimate mortality.

6.4 Locating the time reference of the estimates

Like those from lifetime orphanhood, estimates made from the data supplied by different age cohorts of respondents, about orphanhood before and after marriage, reflect mortality over varying and ill-defined periods of time. As for the basic method (see section 2.2), for each cohort there is a date at which the cohort estimate of mortality obtained from orphanhood data equals the equivalent period measure of mortality. Thus, if the data are sufficiently accurate, the estimates obtained from respondents of different ages can be used to measure the trend in mortality over time.

The procedure developed by Brass and Bamgboye (1981), for estimating the time location of adult mortality estimates calculated from lifetime orphanhood and widowhood, can also be used to calculate the time reference of mortality measures obtained from data on orphanhood prior to marriage and orphanhood since marriage. The justification for this is clear if the discussion in section 6.1 is recalled. The proportion of women orphaned prior to marriage is similar to the overall proportion of female respondents orphaned in the age group encompassing the mean age of marriage of women in the population concerned. In general terms, the distribution of intervals between orphanhood and marriage is also similar to the distribution of intervals between orphanhood and the time of interview, for women whose age is the same as the mean age at marriage. Thus, reasonable estimates of the time location of estimates made from data on orphanhood before marriage can be obtained using the method developed by Brass and Bamgboye (1981) for orphanhood at all ages. Similar considerations apply to estimates from orphanhood since marriage. These data are similar in form and characteristics to those concerning the proportion of women whose first husbands have died, in populations in which men marry at much later ages than women (Brass, 1986). Once again, use of the procedure developed by Brass and Bamgboye seems justified. It is suggested in section 2.4 that appreciable increases in the overall accuracy of the time references calculated for indirect estimates of adult mortality could probably be achieved by replacing the adjustment factors, f_u, proposed by Brass and Bamgboye, by separate adjustments for use with data on maternal orphanhood, paternal orphanhood and widowhood respectively. No doubt tailor-made procedures would also produce better results for orphanhood before marriage and since marriage. Nevertheless, the generic method is adequate for the initial assessment of the characteristics and value of mortality estimates made from these new forms of orphanhood data.

For orphanhood before marriage, the time reference of the mortality measures is a convolution of the distribution of intervals between parental deaths and marriage and the distribution of intervals between marriage and interview. To the degree of precision required, time references for maternal orphanhood equal the sum of the mean interval from marriage to interview and the mean interval since orphanhood, at the time of first marriage. For cohorts of women who have nearly all married, the ages at which the parents are exposed to the risk of dying are concentrated between the mean age at childbearing and the sum of that age and their daughters' mean age at first marriage. Using the notation already established, the time reference of the mortality measures can be calculated as:

$$T = (N - \overline{m}) + {}_{\overline{m}}g_{\overline{M}} \tag{6.4.1}$$

The second term on the right-hand side of equation 6.4.1 can be estimated from the right-hand side of equation 2.2.5 and the adjustment factors, f_u, published by Brass and Bamgboye (1981). As for lifetime orphanhood, the time reference of estimates of male mortality should allow for deaths between conception and birth and are calculated using ${}_{\overline{m}+0\ 75}g_{\overline{M}-0\ 75}$.

For orphanhood since marriage, the age at which parents enter exposure to the risk of death is a convolution of their distribution of ages at childbearing and their daughters' ages at marriage. Following Brass and Bamgboye's argument, for age groups that have largely completed marriage, one can collapse these distributions to their means. The parents' exposure continues till the time of interview, N-$\overline{m}$ years later. Thus the time references of the mortality measures are:

$$T = {}_{N-\overline{m}}g_{\overline{M}+\overline{m}} \tag{6.4.2}$$

They can also be estimated from equation 2.2.5. Because the age range over which fathers are exposed to the risk of death commences well after the birth of their daughters, equation 6.4.2 is appropriate for estimates of both male and female mortality.

6.5 Applications

Estimates of adult mortality from data on orphanhood since marriage in Morocco, collected in the survey conducted in 1987 as part of the DHS programme, are shown in table 6.1. The life table indices have been calculated using the coefficients presented in appendix 2 and the time reference of the estimates from equation 6.4.2, using the adjustment factors proposed by Brass and Bamgboye (1981). As in the previous chapter, the results are translated into α, the level parameter of the relational logit system of model life tables based on the General Standard (Brass, 1971b).

Table 6.1 Estimation of adult mortality from orphanhood since marriage, Morocco.

Age (n)	Proportion not orphaned since marriage ${}_5S_{n-5}/{}_5S^m_{n-5}$	$\frac{l(b+n)}{l(b+20)}$ ($\bar{m}$=18.3)	α	Date
Female mortality:		(b=25, $\bar{M}$=28.3)		
30	0.9650	0.9564	-0.997	1982.9
35	0.9083	0.8997	-0.862	1981.4
40	0.8022	0.8638	-1.002	1980.3
45	0.7234	0.7927	-1.061	1979.8
50	0.6114			
Male mortality:		(b=35, $\bar{M}$=36.67)		
30	0.8801	0.8686	-0.799	1983.4
35	0.7574	0.7669	-0.850	1982.1
40	0.6133	0.6316	-0.925	1981.2
45	0.4919			

In Morocco, cross-sectional data on marital status yield a mean age at marriage for women of over 21 years. However, ages at marriage are known to have risen in Morocco in recent years. The dates of marriage reported retrospectively by women aged 25 years and over in the DHS indicate that they married at younger ages than suggested by the cross-sectional data. As the mean ages at marriage reported by different cohorts are very similar, an average value of 18.3 years, calculated from data on respondents aged 25 to 44, was used to produce all the estimates of mortality.

It can be seen from table 6.1 that the time references of all the mortality estimates fall within a period that is little more than three years long. Given the nature of the data and method, it seems unlikely that the results are sufficiently accurate to detect trends in adult mortality over such a brief period. Thus, the best way to treat these estimates is probably to average those that seem most reliable, to produce a point estimate of mortality. In most applications it would be sufficiently accurate to take this average as applying to a date six years before the data were collected. In Morocco another consideration supports this argument. The estimates of mortality for both males and females from older respondents indicate significantly lighter mortality than those obtained from the first couple of age groups. This is unlikely to be a real trend. Moreover, it cannot be the product of an increase in ages at marriage across the cohorts, as this would produce a bias in the opposite direction. The most likely explanation of this apparent trend is age exaggeration by older respondents. Like estimates from orphanhood after age 20, estimates from orphanhood since marriage are more sensitive to age reporting errors than those from lifetime orphanhood because they are based on parental exposure at relatively advanced ages. In addition, the first estimate for females is a little high, while in some of the applications presented later it is rather low. This is probably because the number of deaths on which this proportion is based is fairly small.

Table 6.2 Estimation of adult mortality from orphanhood before marriage, Morocco.

Age (n)	Proportion not orphaned before marriage ${}_5S_n^m$	$\frac{l(b+20)}{l(b)}$ ($\bar{m}$=18.3)	α	Date
Female mortality		(b=25, $\bar{M}$=28.3)		
25	0.9199	0.9248	-0.659	1970.2
30	0.8985	0.9051	-0.514	1965.2
35	0.8947	0.9019	-0.493	1960.1
40	0.8496	0.8562	-0.226	1954.9
45	0.7981	0.8038	0.033	1949.8
Male mortality		(b=35, $\bar{M}$=36.67)		
25	0.8522	0.8862	-0.673	1970.4
30	0.8229	0.8570	-0.518	1965.3
35	0.8002	0.8326	-0.402	1960.2
40	0.7980	0.8301	-0.391	1955.0
45	0.6952	0.7170	0.087	1949.8

Estimates of adult mortality in Morocco from the data on orphanhood before marriage collected in the DHS survey are shown in table 6.2. They have been obtained using the final sets of coefficients presented in appendix 2. The time reference of the measures was estimated on the basis of equation 6.4.1. It is clear from the differences by age in the proportions of respondents with living mothers and fathers at marriage that a substantial decline in adult mortality has occurred in Morocco. This is confirmed by the estimated life table measures and the values of α derived from them. The time reference of the estimates based on respondents aged 25 to 29 years, who married around 9 years before the survey was conducted, is about 7 years before the time that the data were collected. As one would expect, the time reference of the each of the following estimates is about 5 years earlier than the one before. The estimates from respondents aged 45 to

49 measure the level of adult mortality more than 37 years before the data on orphanhood prior to marriage were collected.

These estimates of mortality from orphanhood before and after marriage in Morocco are shown in figure 6.1, together with the estimates from lifetime orphanhood according to the WFS

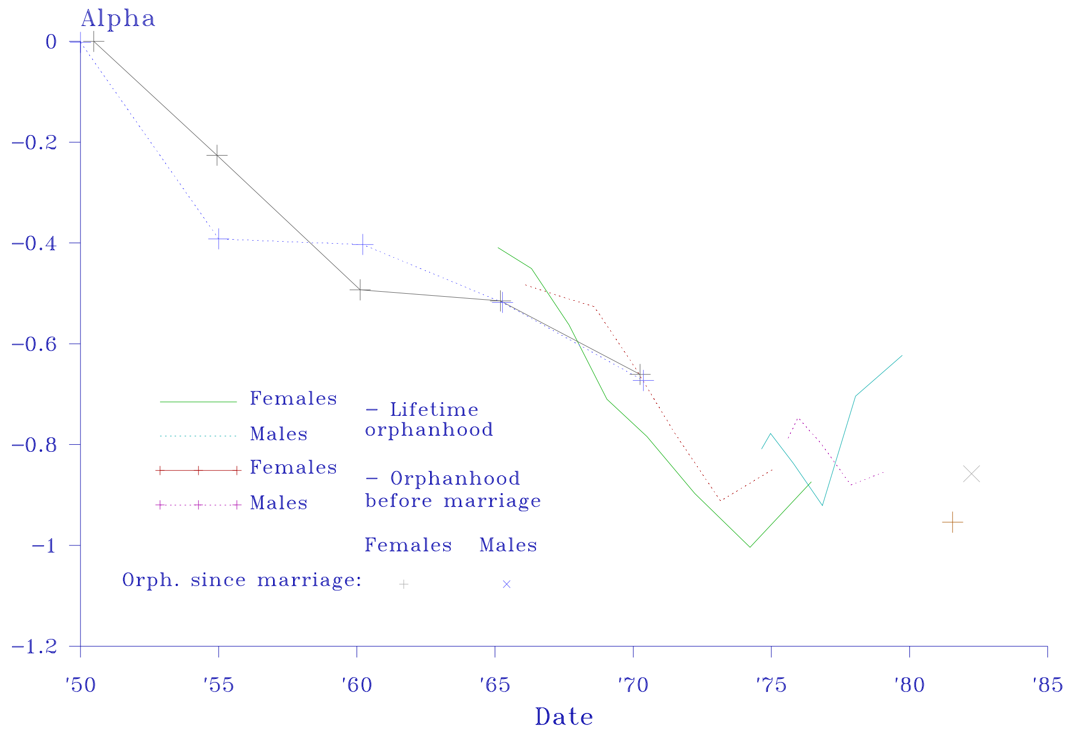

Figure 6.1: Adult mortality (α) estimated from orphanhood before and since marriage, Morocco.

and DHS surveys, that are discussed in section 3.2. Comparison of the two sets of cohort data suggests that the more recent estimates from the WFS may be biased downward somewhat by the adoption effect. In addition, as the earlier discussion mentions, the DHS survey interviewed a sample of ever-married women. Early marriage is associated with both maternal and paternal orphanhood in Morocco. This biases the DHS data on the survival of parents downward for respondents aged less than 25, resulting in the sharp recent increase in female mortality indicated by these data and the smaller upturn in the most recent DHS estimate for males. Because of this problem, no attempt has been made to produce estimates for Morocco based on the incidence of orphanhood after age 20 in between the WFS and DHS surveys.

The existence of an association between orphanhood and entry into marriage in Morocco breaches one of the assumptions made in the derivation of the coefficients for estimating life table measures from orphanhood before and after marriage (see section 6.2). The strength of the association is difficult to determine, but can be gauged in broad terms from table 6.3. This table compares the prevalence of orphanhood at marriage, reported retrospectively by women interviewed in the DHS who married between mid-1977 and mid-1982, with cross-sectional data on women of all marital statuses according to the WFS survey conducted at the end of 1979. Although the WFS data may be biased upwards by the adoption effect and the DHS figures are based on fairly small numbers, it is clear that women who marry at young ages in Morocco are more likely to have lost a parent than other girls of their age. The extent to which this is because high mortality and early marriage are concentrated among the rural poor and the extent to which orphanhood itself causes early marriage would be difficult to establish from the relatively small, cross-sectional DHS survey, but does not matter for the current analysis.

Table 6.3 Percentages by age of all women, and of women who marry, with living fathers and mothers, Morocco 1979.

Age group	All women (WFS)	Women who marry (DHS)
Percentage with mother alive:		
10-14	97	91
15-19	95	93
20-24	91	90
Percentage with father alive:		
10-14	93	87
15-19	87	88
20-24	80	82

If mortality was estimated from orphanhood before and since marriage of cohorts of women who had not largely completed entry into marriage, the association between ages at marriage and orphanhood would introduce a major bias into the results. For respondents aged 25 and over this is less of a problem. What the association does imply, however, is that aggregate data on ages at marriage misrepresent the length of time for which parents have been exposed to the risk of dying. For orphanhood before marriage the bias is not of any significance. The estimates are extremely robust to the exact shape of the marriage distribution and the shorter than average potential exposure of the parents who died before their daughters married is offset by the longer exposure of parents who were alive when their daughters married. The impact of the bias is also limited for orphanhood since marriage because the majority of parents are alive when their daughters marry. Although, on average, parents who live till their daughters marry are exposed for a shorter period subsequently than is indicated by the overall mean age at marriage, in Morocco the mean ages at marriage of women with living fathers and mothers only differ by 0.05 years from the overall mean age at first marriage. The bias that results from using an estimate of $\overline{m}$ applying to the whole population is minimal. Given the enormous demographic differentials that exist between the urban and rural areas and educated and uneducated in Morocco (Al-Jem *et al*., 1988), it seems probable that the association between marriage and orphanhood is particularly strong in this country. For example, the DHS data on Burundi, discussed next, reveal a weaker tendency for orphaned women to marry early. Thus, it is unlikely that, in general, this problem is an significant source of biases in mortality estimates from orphanhood before and since marriage.

The other assumption, in addition to those involved in the basic orphanhood method, made to derive coefficients for estimating mortality from orphanhood relative to marriage, is that women's ages at marriage are unrelated to their parents' ages. In many populations such a relationship probably exists. For example, the youngest daughter may tend to remain at home to look after her aging parents. Despite this, in light of the discussion of the impact that any association between orphanhood and age at marriage has on the results, it seems unlikely that breaches of this second assumption are a significant source of bias in estimates of mortality from orphanhood relative to marriage.

The results shown in figure 6.1 are most encouraging. The information on orphanhood before and since marriage eases interpretation of the results based on lifetime orphanhood and provide estimates for a much longer span of time than the basic method. In particular, the information on orphanhood since marriage yields mortality estimates that are at least five years more up-to-date than the most recent acceptable estimates from the data on lifetime orphanhood collected in the DHS survey. The estimates made from cohort data on older respondents in the WFS and the DHS and those from orphanhood since marriage form fairly consistent series for both male and female mortality. If anything, the estimates from orphanhood since marriage suggest rather higher mortality than the trend in the cohort data. They indicate that life expectancy at age 15 in the early 1980's was a little under 58 years, compared with an estimate from the WFS, for 15 years earlier, of about 51 years. In broad terms, the estimates from orphanhood before marriage, according to the DHS, provide confirmation of the earlier estimates from the WFS. These data indicate slightly lower mortality than the trend in the more recent figures but, given that the estimates from older respondents are probably biased downward to some extent by age exaggeration, they offer good evidence that a prolonged decline in adult mortality has occurred in Morocco, from a very high level at the beginning of the 1950's.

Estimates of mortality in Burundi from orphanhood before and after marriage and other sources are shown in figure 6.2. The estimates from the 1971 multi-round survey are based on the life table published by Condé *et al*. (1980) and the orphanhood estimates on data collected by the DHS survey (see appendix 3). The synthetic cohort measures estimated from orphanhood after age 20 have been obtained from the proportions orphaned calculated by Makinson (1988). For about 25 *per cent* of respondents she had to estimate the date of orphanhood (see section

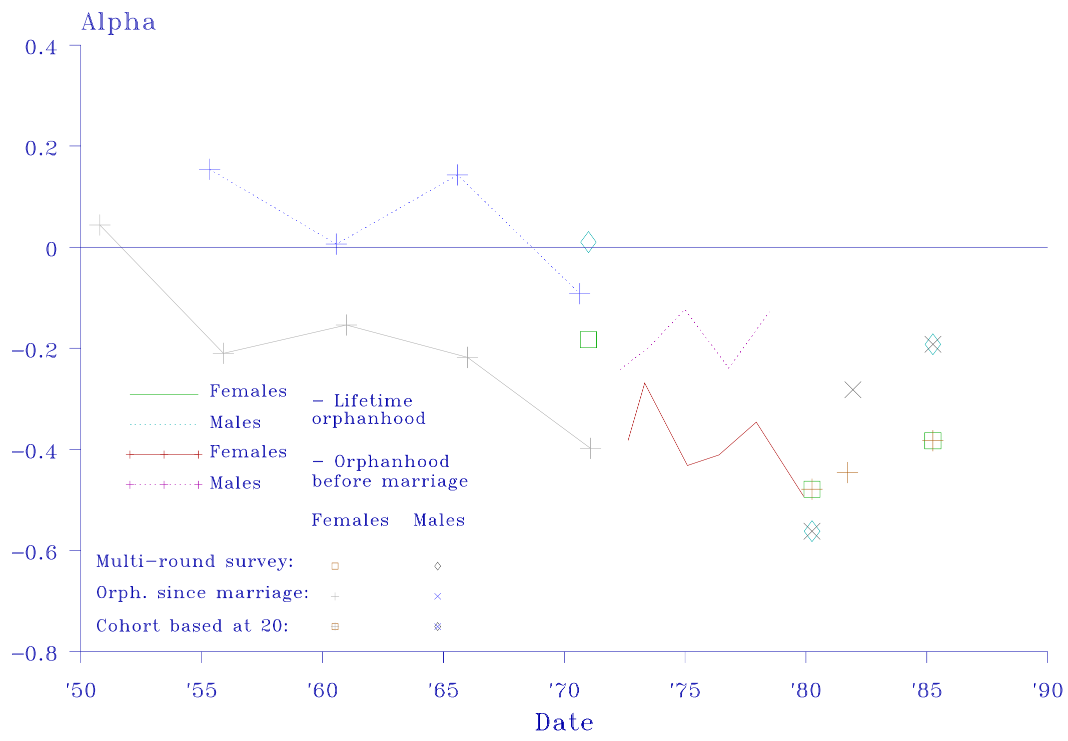

Figure 6.2: Adult mortality (α) estimated from orphanhood before and since marriage, Burundi.

3.1). She reports that this probably biases the most recent estimates upwards. This seems to have occurred but, in addition, the estimates for the period 5 to 9 years before the survey appear biased in the opposite direction. This suggests that, as in the Ugandan survey discussed in section 5.3, these period estimates are affected by errors in the reporting of dates of orphanhood. In this survey, dates of orphanhood seem to have been shifted forward. The data on deaths of fathers are more affected than those on mothers.

As in Morocco, the results obtained from orphanhood before and since marriage seem very useful and provide a much sounder basis for determining the level and trend in adult mortality in Burundi than would exist if only data on lifetime orphanhood and the multi-round survey results were available. Although the estimates from lifetime orphanhood and orphanhood before marriage are rather erratic, they agree well for both males and females and offer clear evidence that, while adult mortality in Burundi remains fairly high, it has undergone a slow decline extending from the early 1950's to the late 1970's. If the data from the multi-round survey can be relied upon, they indicate that these orphanhood estimates understate the life expectancy of young adults by a modest amount.

Extrapolating forward from the estimates from the multi-round survey, on the basis of the trend in the data on lifetime orphanhood and orphanhood before marriage, it seems likely that the estimates from orphanhood since marriage are fairly accurate. They yield a life expectancy at age 15 for the early 1980's of 49 years, compared with 45 years for 1971. Although the estimates from orphanhood after age 20 are clearly only partially independent of the estimates from orphanhood since marriage, averaging the points for 0 to 4 years and 5 to 9 years before the survey also suggests that the latter results are reasonably accurate.

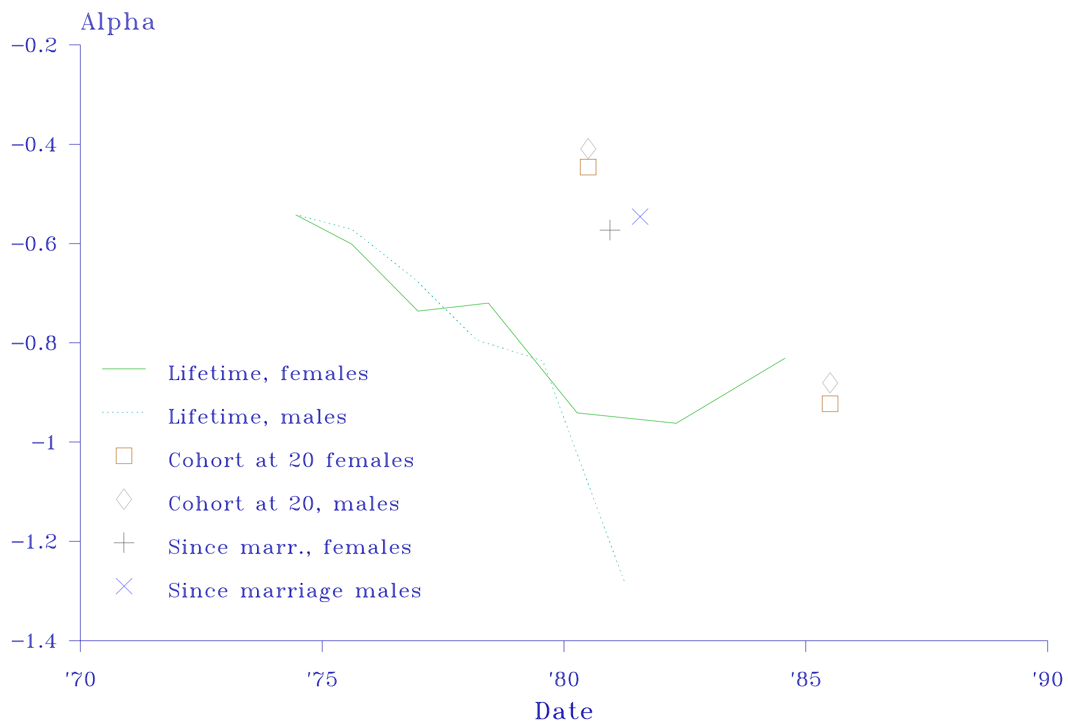

Figure 6.3: Adult mortality (α) estimated from orphanhood before and since marriage, Buganda.

In figure 6.3, results are presented from a survey conducted in the Buganda region of Uganda, that is also discussed in section 5.3. Of the synthetic cohort estimates presented in section 5.3, only those from orphanhood after age 20 are included in this figure. It is argued in the earlier discussion that the orphanhood data from this survey are severely biased by the adoption effect. Further evidence of this is provided by the finding that only a small proportion of respondents in this survey reported that they had been orphaned before marriage. These data did not yield useful estimates of adult mortality.[2] The estimates from orphanhood since marriage are more plausible than any of the other estimates. Although there is no independent source of data against which they can be validated, they indicate very similar mortality for males and females and much higher mortality than the estimates from lifetime orphanhood. Thus, although some of the `parents' that were alive when these women married may not be their biological genitors, it does not appear that adults are liable to redefine someone else as their parent when the person who reared them as a child dies. The values of α obtained from the data on orphanhood since marriage correspond to an expectation of life at age 15 of about 52 years, an estimate which seems reasonable for the early 1980's. Partial confirmation of the estimates from orphanhood since marriage is provided by the fact that they fall in between the two sets of points obtained from orphanhood after age 20. (As is suggested in section 5.3 though, it is likely that errors in the retrospective reporting of dates lead these results to exaggerate the decline in mortality that has occurred in this region).

The initial applications of the procedures for estimating adult mortality from orphanhood before and since marriage that are presented in this section provide only a tentative basis for assessing the usefulness of the methods. Nevertheless, the results are very encouraging. In the three analyses presented here, estimates from orphanhood since marriage are either consistent with the other estimates available or seem more believable than them. The sex differentials in adult mortality indicated by these data are plausible and agree closely with those from other sources. These applications also reveal four important potential advantages of data on orphanhood since marriage for the estimation of adult mortality. First, these data make it possible to estimate mortality about 5 to 6 years before the survey even if, as in the studies conducted by the DHS, only respondents aged 15 and over are asked about orphanhood. Secondly, a question on the timing of orphanhood relative to marriage seems a more robust way to refine the time reference of the mortality estimates than a question on the dates when respondents were orphaned. The latter question appears to have been badly affected by errors in the retrospective reporting of dates in the surveys in both Burundi and Uganda. In the former study, it was also subject to a high degree of item non-response. Collection of data on orphanhood relative to marriage makes fewer demands on the respondents. Thirdly, synthetic cohort data derived from surveys that only interview ever-married women are vulnerable to selectivity bias, stemming from association of women's ages at marriage with orphanhood. Estimates based on information supplied by older respondents about orphanhood relative to marriage are less liable to bias arising from this source. Fourthly, the application to Ugandan data suggests that, in countries where the basic data on orphanhood are biased severely by under-reporting of the incidence of orphanhood in childhood,

---

[2] This confirms that adults who were orphaned at young ages tend to answer questions about orphanhood in terms of their foster- or step-parents. In Buganda, at least, errors stemming from the adoption effect are not limited solely to proxy reports on behalf of young children.

it may be possible to collect much more accurate data on orphanhood since marriage. One problem with this method of estimating mortality that these initial applications have revealed, on the other hand, is that the initial estimate for a central age, *n*, of 30 years tends to be rather erratic, probably because of sampling errors, while the estimates based on older age groups are sensitive to age exaggeration by respondents. Thus mortality estimates have to be based on the information supplied by respondents in a narrow range of ages.

The usefulness of the information on orphanhood before marriage, that also results from questions on whether respondents were orphaned before or after they first married, appears to depend on whether or not it is affected by the adoption effect. In Uganda these data were useless but, in Morocco and Burundi, they appear to be reasonably accurate, although they probably underestimate mortality a little. The estimates provide useful confirmation of the trend in mortality indicated by data on lifetime orphanhood and seem capable of extending the time series of mortality estimates provided by the orphanhood method backward to some 30 or 35 years before the data were collected. It is likely that selectivity biases arising from any association between the mortality of parents and children becomes increasingly severe as the age of respondents increases and a rising proportion of them die themselves. Nevertheless, it seems possible that, in populations where older people report their age reasonably accurately, data on orphanhood before marriage could be used to reconstruct the historical trend in mortality over a period of 50 or more years.

Chapter 7

FITTING A LIFE TABLE TO INDIRECT ESTIMATES

7.1 Background

Two major limitations are inherent in the indirect methods for estimating mortality. First, determination of the trend in mortality is only possible on the assumption that the age pattern of mortality is known. Unless this is done there is no way to distinguish relatively high mortality at younger ages from relatively high mortality recently and *vice versa*. Second, the estimates measure mortality over limited and varying age spans. These limitations of the indirect methods must be met head on in any attempt to construct a full life table from such estimates of mortality. As the data are incomplete, this can only be achieved by fitting an appropriate model schedule. The principle involved in this is to fit a sufficiently flexible model to extract the maximum amount of genuine information from the data, while smoothing out the impact of various forms of error.

Questions about the proportion of women's children that have died and the survival of parents provide independent information about child and adult mortality levels and trends over time. Brass's 2-parameter relational logit life table system represents a convenient model for deriving a life table from such data.[1] The model life tables are defined by two parameters, α and β, that respectively determine the overall level of mortality and the comparative severity of early age and later mortality in relation to a standard pattern (Brass, 1971b). The differences between the logits of the $l(x)$ values in any two life tables are approximately linear in x, so that simple graphical or regression methods can usually be used to fit a model to a survivorship function. In particular, $Y(x) = \alpha + \beta Y_s(x)$, where $Y(x)$ represents the logit of an observed $l(x)$ value and $Y_s(x)$ represents the equivalent measure in the standard life table. Any life table can be used as a standard but, if no single life table appears particularly appropriate, the General Standard developed by Brass (1971b), and subsequently smoothed and modified (Carrier and Hobcraft, 1971; Ewbank *et al*., 1983), represents a typical mortality pattern.

Straightforward methods of fitting a line to the observed data cannot be used in the case of estimates obtained by indirect methods because the orphanhood technique measures ${}_np_b$, survivorship conditional on having reached a base age in adulthood, b, which is fixed for each variant of the method. Brass (1975) suggests a simple iterative method of fitting a model life table

[1] Other approaches based on Ledermann and Princeton model life tables are discussed by Page and Wunsch (1976) and the UN (1983).

to estimates of this sort. First $l(2)$ is estimated from information on the proportion of children that have died. On the assumption that $\beta = 1$, the probability of surviving to the base age, l(b), can be calculated using the value of α corresponding to $l(2)$. By multiplying the conditional survivorship ratios by this estimate of l(b), absolute measures of l(b+n) and, therefore, Y(b+n) can be obtained. The differences between the Y(b+n) values and Y(2), compared with those in the standard, yield a series of new estimates of β. Averaging the values of β that seem most reliable and then recalculating α, a revised estimate of l(b) can be made from $l(2)$ and the procedure repeated. Further iterations only improve the results slightly. It is important to note that, although the estimates of β are obtained from data on orphanhood, they do not measure the rate at which mortality increases with age in adulthood. They reflect the proportion of mortality by age b+n that occurs before age two and between ages *b* and b+n, relative to the standard.

The procedure outlined in the preceding paragraph was proposed before methods were developed for estimating the time location of indirect measures. Because the estimate of $l(2)$ and each of the ${}_np_b$ ratios estimated from orphanhood refer to different dates, erroneous and implausible relationships between child and adult mortality were often obtained. To estimate mortality trends it is assumed that they have been linear and that the age pattern of mortality corresponds to that in a 1-parameter life table system (see section 2.2). On this basis, the mortality estimates made from the data supplied by different age groups of respondent can be translated into equivalent measures and compared to establish trends. To compare those for adults and children, the same measures have to be used. A straightforward way of doing this is in terms of the level parameter of the family of model life tables used to derive the estimates.

Typically estimates of the level of mortality obtained from child and parental survival data for the same date will differ. This implies that the age pattern of mortality in the population in question does not correspond to that in the 1-parameter family of models being used. One response to this is to assume that it will not greatly bias either set of estimates of the level of mortality considered individually. Separate 1-parameter model life tables can then be fitted to the child and adult estimates and spliced together at some age (usually 15 years) where mortality is low and the discontinuity in the two series slight (UN, 1983). Unfortunately, evidence has accumulated that splicing life tables together in this way is sometimes an unsatisfactory way of combining two 1-parameter models. In particular, if the relative level of mortality in childhood and adulthood is very different from that assumed, the approach introduces an abrupt discontinuity into the survivorship function and produces biased estimates of mortality trends (Brass, 1985). If child mortality is relatively high, the decline in mortality will be underestimated while, if child mortality is relatively low, the trend will be exaggerated.

Brass (1985) suggests a way of dealing with this problem that uses relational model life tables. Because data on child and parental survival measure the relative severity of childhood and adult mortality, they can be used in conjunction to determine a more appropriate standard against which to estimate trends. Thus, if the levels of α yielded by the childhood mortality and orphanhood estimates are different, a new standard can be defined by multiplying the $Y_s(x)$ values by a fixed β and this standard used to calculate new values of α. The objective is to find a value of β that defines a standard that brings the estimates of α for children and adults into line. Brass points out that an appropriate β could be estimated for each date at which it is possible to measure the level of both child and adult mortality, but argues that such a procedure is over-sophisticated. A single value of β applying to the middle of the period during which both the child and adult data are most reliable should be adequate.

In some applications Brass's (1985) procedure has given good results. More frequently it has proved impossible to reconcile the data on child and adult mortality over the entire period about which they both provide information. In some cases this may be because of errors in the data on child and parental survival but, in others, successive surveys confirm the accuracy of the estimates. The problem appears to be in the method. There are two possible explanations. Either the 2-parameter models are unable to represent the age pattern of mortality in some populations accurately and a more flexible model is needed, or the relative level of child and adult mortality has changed over time.

Kamara (1988) has investigated both possibilities. Using 4-parameter models, he demonstrates that, in many applications, the age patterns of mortality that have to be posited to reconcile indirect estimates of child and adult mortality over the entire period for which they can be compared are implausible. In such populations the age pattern of mortality must have changed. He therefore fits a series of 2-parameter models to pairs of estimates of child and adult mortality that refer to the same date, allowing both α and β to vary freely. Again, the β parameter in some of the fitted life tables is implausibly extreme. The argument developed here is that this results from deficiencies in the method used to fit the models. With more robust techniques, a series of 2-parameter relational model life tables can be fitted to indirect estimates of mortality in countries where the relative level of child and adult mortality is changing. The parameter estimates remain plausible and the mortality rates in the fitted life tables usually agree well with those indicated by other sources of data.

7.2 A new procedure

The issues involved in fitting 2-parameter life tables to child survival and orphanhood data when the level and pattern of mortality are changing over time are clarified if the problem is stated in a rather different way. The basic estimates comprise a series of probabilities of dying, $q(x)$, for children and a series of conditional survivorship ratios, ${}_np_b$, for adults where *x* and *n* vary with the age group of respondents and, therefore, the time reference of the indices. These two series can only be compared if they are translated into equivalent measures. By definition, measures referring to the same period are drawn from the same life table but most of this life table is unknown. Thus, extrapolation from known measures to other ages is *inherent* in the measurement of mortality trends from information on the survival of relatives. In essence, what is estimated by indirect methods for any date are two values of α that define 1-parameter models referring to children and adults respectively. With an appropriate standard, these two estimates of α should be the same at any date. What is required is a robust procedure for determining the value of β that defines a standard which brings these points into line.

The value of β that brings the estimates of child and adult mortality into agreement is:

$$\beta = \{Y(x_a) - Y(x_c)\}/\{Y_s(x_a)-Y_s(x_c)\} \qquad (7.2.1)$$

Thus, the estimated value of β will depend on the two ages in childhood and adulthood, x_c and x_a, to which the model is fitted. These ages need not equal *x* and b+n, the ages defining the mortality estimates, q(x) and ${}_np_b$. The importance of the choice of x_c and x_a becomes clear if one considers that the relationship between Y(x) and $Y_s(x)$ is only approximately linear. Two parameters are unable to model human mortality patterns exactly and typically a plot of the relationship between Y(x) and $Y_s(x)$ will reveal a degree of curvature in the line. Four-parameter systems of model life tables have been developed in order to reproduce empirical survivorship functions more accurately (Ewbank *et al*., 1983; Zaba, 1979). The limited information obtained from a single enquiry that asks about child and adult survival retrospectively is insufficient to permit their use. Instead a procedure to estimate α and β from such data should use values of x_c and x_a that yield a robust and consistent series of estimates of mortality in the population under study. To produce consistent estimates for different dates, it is essential that the values of x_c and x_a are held constant over the entire period under investigation and not allowed to vary with the age group of respondents.

Reconsideration of Brass's (1975) method of fitting a life table (see section 7.1) may make the gains from this approach clearer. In this method, x_c is set to age two and x_a to a variable range of ages equal to b+n. It is unlikely that these are optimal choices. Most of the curvature in the logit survivorship function at young ages stems from variation in the level of infant, compared with later childhood, mortality. Use of an age in late childhood for x_c would reduce the average errors in the fitted values of *l*(x). It would also reduce the degree of extrapolation, on average, from the estimates of q(x). Examination of the t(x) function (Zaba, 1979) and the κ transformation (Ewbank *et al*., 1983), that have been used to model variation in the relative severity of infant mortality, suggests that setting x_c to exact age 10 or 15 might be optimal. Equally, use of the upper age limits of the survivorship ratios estimated from orphanhood, b+n, for x_a means that any curvature in the logit survivorship function will have a large impact on the estimates of β obtained from data supplied by younger respondents and thus on recent estimates of mortality in infancy and old age. The form of the functions used by Zaba (1979) and Ewbank *et al*. (1983) to model variation in old age mortality, suggests that, to minimize errors in the fitted life table, exact age 65 or 70 should be used for x_a.[2] Models fitted to the data on younger respondents should also be constrained so that the probability of surviving from x_c to x_a equals that in the 1-parameter model identified by ${}_np_b$.

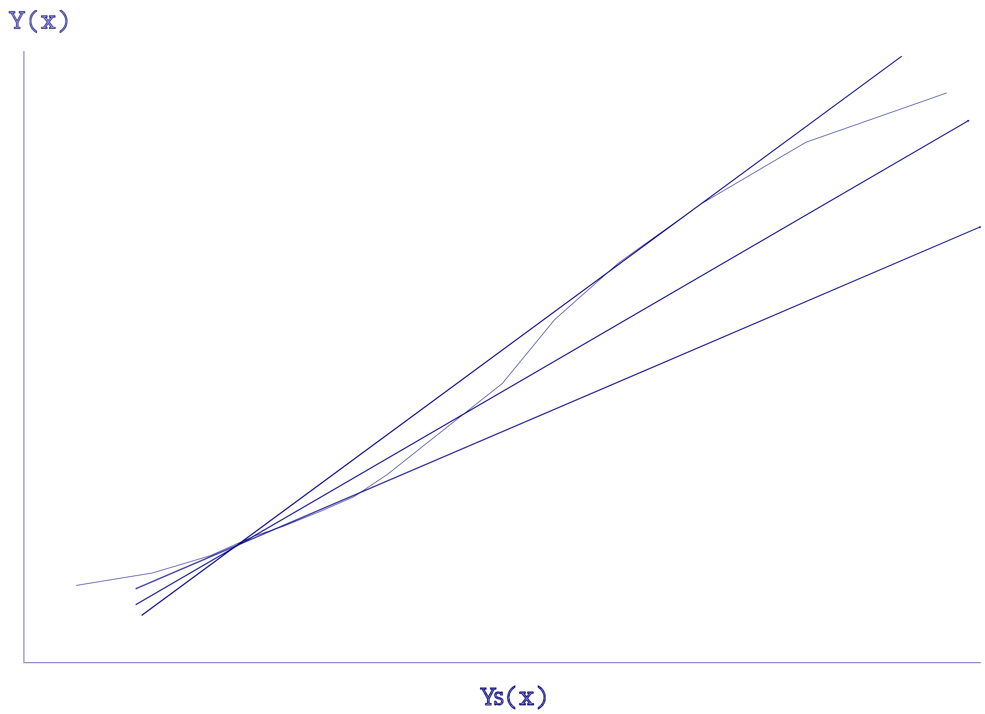

Figure 7.1: Fitting a 2-parameter model life table to different ages.

A schematic representation of the inconsistencies that might be obtained, even if mortality is constant, in a series of model life tables fitted using Brass's (1975) procedure is shown in figure 7.1. It shows that, if the level of infant mortality introduces any curvature in the first half of the logit survivorship function relative to the standard, allowing x_a to vary according to the age of the respondents supplying data will lead to a series of increasingly low estimates for β and high ones for α or *vice versa*. If the lower age at which the life table is fitted (x_c) is also shifted, as in Brass's (1985) procedure, the effect is even greater.

[2] To determine definitively what ages should be used to fit the models would require extensive study of the empirical life tables against which Zaba and Ewbank *et al*. validate their models rather than consideration of the functions that they propose. This is not attempted here. In addition, what defines the best fitting procedure depends not only on the data available but also on the purpose of the analysis. One would need to fit the model to different ages to obtain the best estimate of life expectancy at birth, the best representation of the entire *l*(x) function and the best estimates of mortality over the age range on which indirect methods provide useful data. In the following discussion it is assumed that the latter is the objective.

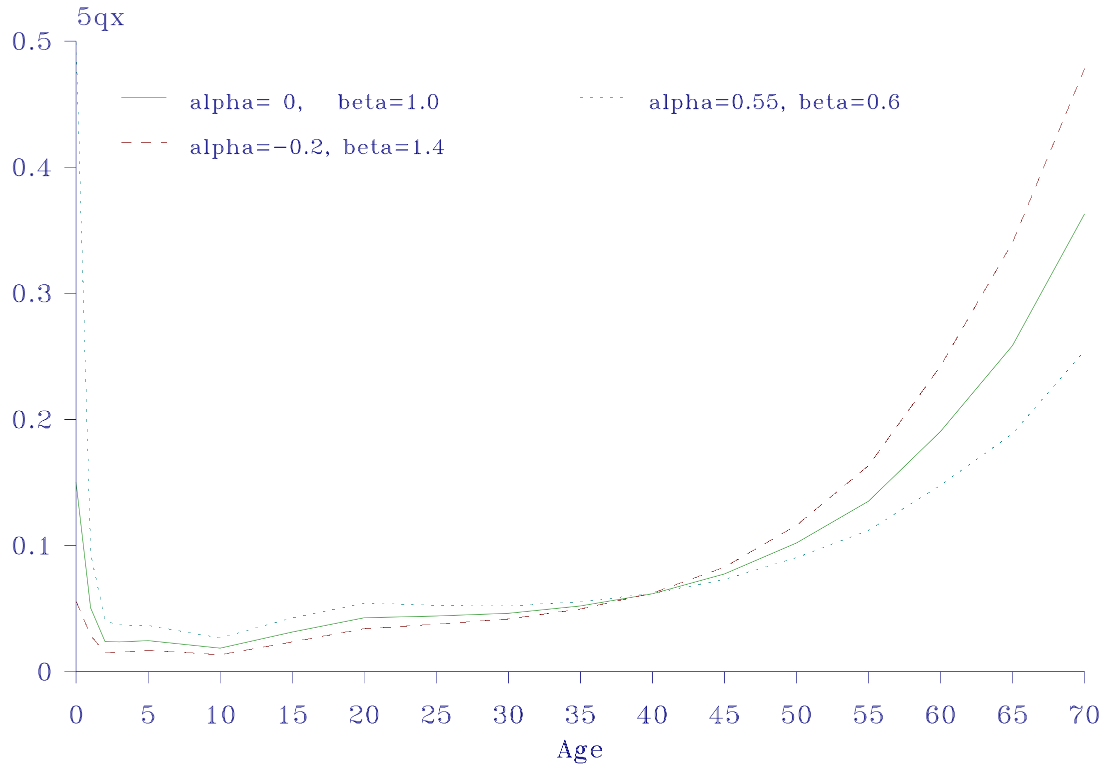

Figure 7.2: Mortality rates by age in selected relational model life tables.

Another consideration supports this argument. Between the ages of 2 and 55 it is rather difficult to distinguish the effects of changes in α and β on mortality. In other words, any pair of estimates of survivorship within this age range can be fitted approximately by a high α and low β or a low α and high β (see figure 7.2). The parameter estimates are rather unstable. Minor errors in the estimates or modest curvature in the survivorship function can have a major impact on the fitted life table. By extrapolating from the survivorship of young adults up to a common age x_a in late middle age, one averages across any such curvature. In other words, this approach ascribes all changes by age group in the relationship between the ${}_np_b$ ratios and the standard to changes in the level of mortality between ages x_c and x_a, not to changes in the pattern of mortality. This eliminates the possibility of obtaining a very extreme value for β and stabilizes the estimation process. While the choice of x_a remains to some extent arbitrary, using a *consistent* assumption to estimate trends in mortality from the reports made by respondents of different ages is essential. Failure to do so explains many of the implausible trends obtained by those trying to estimate the level and age pattern of mortality by indirect means.

Because this procedure involves extrapolating (or occasionally interpolating) from mortality over one age range to mortality over another, it can only be fitted iteratively. Initial parameter estimates (assuming that β = 1) are used to calculate survivorship at ages x_c and x_a;

these measures are then used to derive improved parameter estimates. To put this another way, the counterparts of equation 7.2.1 are equations in which α depends on β. For childhood mortality:

$$\alpha = Y(x) - \beta Y_s(x) \qquad (7.2.2)$$

and

$$Y(x_c) = \alpha + \beta Y_s(x_c) \qquad (7.2.3)$$

For adult mortality, if one assumes that $\beta = 1$ (and uses an asterisk to denote this assumption), Brass and Bamgboye (1981) show that:

$$\alpha^* = -½\ln[1 + \{{}_np_b/l_s(b+n) - 1/l_s(b)\}/\{1-{}_np_b\}]$$

As $Y^*(z) = \alpha^* + Y^*_s(z)$, one obtains:

$$Y^*(z) = ½\ln[\{1-l_s(z)\}/l_s(z)] - ½\ln[1 + \{{}_np_b/l_s(b+n) - 1/l_s(b\}/\{1-{}_np_b\}] \qquad (7.2.4)$$

Denoting $l(x_a)/l(x_c)$, the probability of surviving from x_c to x_a, as p^*:

$$p^* = \{1 + e^{2Y^*(x_c)}\}/\{1 + e^{2Y^*(x_a)}\} \qquad (7.2.5)$$

Substituting equation 7.2.4 into both the numerator and denominator of equation 7.2.5 and rearranging:

$$p^* = \frac{1/l_s(x_c) + \{{}_np_b/l_s(b+n) - 1/l_s(b\}/\{1-{}_np_b\}}{1/l_s(x_a) + \{{}_np_b/l_s(b+n) - 1/l_s(b\}/\{1-{}_np_b\}} \qquad (7.2.6)$$

Combining the estimate of $l(x_c)$ obtained from equation 7.2.3 with the estimate of survivorship from x_c to x_a yielded by equation 7.2.6, one can obtain absolute measures of survivorship in adulthood that allow for variation in the relative level of child and adult mortality:

$$\begin{aligned} Y(x_a) &= ½\ln[\{1-l(x_a)\}/l(x_a)] \\ &= ½\ln[1/\{p^*l(x_c)\} - 1] \\ &= ½\ln[\{1 + e^{2(\alpha+\beta Y_s(x_c))}\}/p^* - 1] \qquad (7.2.7) \end{aligned}$$

By evaluating equation 7.2.1, using the estimates of $Y(x_c)$ and $Y(x_a)$ yielded by equations 7.2.3 and 7.2.7, an improved estimate can be made of β. Finally, a revised estimate of α can be calculated from equation 7.2.2 and the process repeated.

Equations 7.2.1 to 7.2.7 provide the basis for fitting 2-parameter life tables iteratively to single pairs of estimates of child and adult mortality referring to the same date. Typically the child mortality and orphanhood methods yield two series of estimates covering a period of up to 15 years. In the following section a convenient procedure is outlined for fitting life tables to such series. To smooth errors affecting the estimates obtained from particular age groups of respondents, it is assumed that changes in the underlying mortality schedule can be represented adequately by linear trends in both α and β.

7.3 Applications

An application of the new procedure to data on female mortality collected in the Lesotho Fertility Survey is shown in table 7.1. The data were published in Timæus (1984) and are of fairly high quality (see section 3.2). The estimates for adults have been calculated from cohort data on orphanhood, using the coefficients described in section 4.5. The table shows the final results of the iterative fitting procedure. First, the α values corresponding to the q(x)'s were calculated, using equation 7.2.2 and assuming that β = 1. Then the points for the six older ages were regressed on the times to which they apply. Estimates of $Y(x_a)$ were derived for adult women using equation 7.2.7, again assuming that β = 1. Next the regression coefficients derived from the data on children were used to obtain values of $Y(x_c)$ referring to comparable times to the estimates of $Y(x_a)$. Values of β that define a fitted 2-parameter model life table were then estimated using equation 7.2.1, with $x_c = 10$ and $x_a = 70$. The resulting estimates of β for central ages *n* of 20 to 40 years were also regressed on the times to which they apply and new estimates of α derived

Table 7.1 Parameters of relational model life tables fitted iteratively to orphanhood and child mortality data, Lesotho, females.

(a) Estimates of child mortality.

Age (x)	q(x)	α (β=1)	Time reference	Fitted β	Equivalent α	Fitted α
1	0.120	-0.161	1.14	0.524	-0.559	-0.531
2	0.136	-0.207	2.75	0.543	-0.536	-0.501
3	0.162	-0.157	4.66	0.567	-0.446	-0.466
5	0.182	-0.140	6.87	0.593	-0.389	-0.425
10	0.185	-0.197	9.28	0.623	-0.403	-0.381
15	0.213	-0.142	12.00	0.655	-0.319	-0.331
20	0.232	-0.143	15.32	0.696	-0.282	-0.270

(b) Estimates of adult mortality.

Age (n)	${}_nP_{25}$	${}_{60}P_{10}$ (β=1)	Time reference	Fitted α	Equivalent β	Fitted β
20	0.937	0.633	8.02	-0.404	0.605	0.607
25	0.905	0.614	9.99	-0.368	0.634	0.631
30	0.864	0.606	11.72	-0.336	0.648	0.652
35	0.797	0.583	13.31	-0.307	0.682	0.671
40	0.717	0.584	14.72	-0.281	0.682	0.688

from the q(x) measures using the appropriate value of β for the date in question. Then the regression of these points on the times to which they refer was recalculated and the fitted estimates of α and β used to obtain new values of $Y(x_c)$ corresponding to estimates of the time reference of the adult mortality measures. (These were adjusted using equation 2.2.7 and an average value of β). Thus, the fitted values of β and then α can be re-estimated and the process allowed to iterate to a solution.

The results suggest that female mortality has declined over time and that women have benefited slightly more than children. This accords with what is suggested by the original estimates of child and adult mortality and by estimates based on data collected in other surveys (Timæus, 1984). The estimated values of α and β are plausible and generate a consistent series of life tables. These suggest that life expectancy at birth rose from 53.4 years to 58.2 years during the 1960's. In contrast, fitting a model to the estimates for 8 years before the survey, using the procedure suggested by Brass (1975), without obtaining the solution for $x_c = 10$ and $x_a = 70$, produces an estimate for β of 0.485 and for α of -0.447. The estimate of β is implausibly low and the life table generated from these parameters has a life expectancy at birth of 61 years. On the other hand, splicing together life tables generated from the original estimates of child and adult mortality suggests that life expectancy at birth only rose from 54.6 to 56.6 years. The parameters estimated using the procedure proposed here indicate that female infant mortality declined steadily from about 153 to 139 per thousand during the 1960's. Although these estimates are a little higher than direct measures from the birth histories collected in the Lesotho Fertility Survey (Timæus, 1984), they indicate the correct trend. Fitting procedures that allow x_a to change, erroneously suggest that the infant mortality rate *rose* during the 1960's from 147 to 154 per thousand.

Table 7.2 Effect of changing the ages to which a life table is fitted on the parameter estimates, Lesotho, females.

Ages fitted		1962 α	1962 β	1969 α	1969 β
$x_a = 70, x_c =$	1	-0.27	0.73	-0.37	0.67
	2	-0.27	0.71	-0.38	0.65
	5	-0.27	0.70	-0.39	0.62
	10	-0.28	0.69	-0.40	0.61
	15	-0.28	0.68	-0.41	0.60
	20	-0.28	0.67	-0.42	0.58
	25	-0.29	0.65	-0.44	0.55
$x_c = 10, x_a =$	60	-0.30	0.63	-0.46	0.51
	65	-0.29	0.66	-0.43	0.56
	70	-0.28	0.69	-0.40	0.61
	75	-0.27	0.73	-0.37	0.66
	80	-0.25	0.77	-0.34	0.72

It is possible to repeat the estimation procedure suggested here, using a variety of ages for x_c and x_a, and to examine the sensitivity of the final life table parameters to changes in these ages. Thereby one can identify the crucial assumptions involved in fitting a 2-parameter logit model life table to indirect estimates of mortality. Table 7.2 shows the effects of using different ages x_a and x_c on the estimates of female mortality in Lesotho. It demonstrates that the choice of the lower age at which to fit the model has a small impact on the final life table estimated for any date. Use of any age between 5 and 20 yields rather similar results. The choice of the upper age at which to fit the model is more important. The higher x_a, the nearer β is to one.[3] Over the period as a whole, the 2-parameter models that correspond most closely to those generated by splicing together two 1-parameter models, while smoothing out the discontinuity at age 15, are those fitted to ages 10 and 70. They are adopted here. Whatever assumptions are made, however, it is clear that β is less than one and that it has declined slightly over time. One should not fit a life table to these data without taking both facts into account. Fitting separate 1-parameter models to the data on children and adults, leads to underestimation of the decline in mortality and overestimation of recent mortality.

Table 7.3 Parameters of relational model life tables fitted iteratively to orphanhood and child mortality data, Lesotho, males.

(a) Estimates of child mortality.

Age (x)	q(x)	α (β=1)	Time reference	Fitted β	Equivalent α	Fitted α
1	0.129	-0.118	1.14	1.048	-0.078	-0.106
2	0.148	-0.156	2.75	1.039	-0.128	-0.105
3	0.175	-0.110	4.66	1.028	-0.091	-0.103
5	0.202	-0.077	6.87	1.016	-0.067	-0.102
10	0.204	-0.135	9.28	1.003	-0.133	-0.100
15	0.236	-0.074	12.00	0.988	-0.081	-0.099
20	0.251	-0.092	15.32	0.969	-0.106	-0.096

(b) Estimates of adult mortality.

Age (n)	${}_np_{35}$	${}_{60}p_{10}$ (β=1)	Time reference	Fitted α	Equivalent β	Fitted β
20	0.761	0.350	9.21	-0.100	1.006	1.003
25	0.671	0.358	11.05	-0.099	0.989	0.993
30	0.555	0.357	12.77	-0.098	0.985	0.983
35	0.423	0.363	14.39	-0.097	0.972	0.974
40	0.277	0.363	15.98	-0.096	0.968	0.966

[3] Note that if x_a was set to the highest age at which there are survivors in the standard, say 95, or x_c in the first few days of life, β would be necessarily 1.

An application of the procedure to the Lesotho Fertility Survey data on males, published by Timæus (1984), is shown in table 7.3. The final parameter estimates suggest that overall mortality was more or less constant during the 1960's, with adult men suffering worsening mortality, while boys benefited from some mortality decline. Once again this accords with earlier analyses (Timæus, 1984) and the estimates derived from the procedure would yield a plausible series of life tables. The infant mortality rate in the fitted life tables declines from 139 to 133 per thousand over the period considered. These estimates agree well with those obtained from the Lesotho Fertility Survey birth histories. In this case, as one might expect given the similarity between the age pattern of mortality amongst Basotho men and that in the General Standard, all the fitting procedures give very similar results. The simplest procedure, which is to splice together two 1-parameter models at age 15, gives satisfactory results.

A third application to data considered in chapters 5 and 6, from a survey conducted recently in part of the Buganda region of Uganda, is shown in table 7.4. The estimates suggest that there was a rapid decline in female mortality, concentrated almost entirely among adults, during the 1980's. The values of β that result for recent years are implausibly low. Either lifetime

Table 7.4 Parameters of relational model life tables fitted iteratively to orphanhood and child mortality data, Buganda, females.

(a) Estimates of child mortality.

Age (x)	q(x)	α (β=1)	Time reference	Fitted β	Equivalent α	Fitted α
1	0.118	-0.169	1.37	0.271	-0.779	-0.654
2	0.134	-0.213	2.72	0.318	-0.704	-0.612
3	0.197	-0.037	4.63	0.385	-0.447	-0.552
5	0.176	-0.162	6.84	0.462	-0.490	-0.483
10	0.205	-0.132	9.24	0.546	-0.380	-0.408
15	0.206	-0.162	11.96	0.641	-0.346	-0.323
20	0.240	-0.121	15.27	0.757	-0.232	-0.220

(b) Estimates of adult mortality.

Age (n)	${}_{n}p_{25}$	${}_{60}p_{10}$ (β=1)	Time reference	Fitted α	Equivalent β	Fitted β
10	0.976	0.661	3.48	-0.588	0.530	0.344
15	0.970	0.715	5.82	-0.515	0.458	0.426
20	0.954	0.707	8.02	-0.446	0.479	0.503
25	0.905	0.614	10.06	-0.383	0.620	0.575
30	0.871	0.621	11.81	-0.328	0.615	0.636
35	0.782	0.561	13.56	-0.273	0.705	0.697
40	0.676	0.535	15.25	-0.220	0.748	0.756
45	0.616	0.597	16.23	-0.190	0.663	0.790

orphanhood is being under-reported by respondents aged under 20 years or there is extremely sharp disjuncture between the level of mortality experienced by older children and young adults that cannot be modelled adequately by a 2-parameter relational system of life tables based on the General Standard. The analyses in chapters 5 and 6 strongly support the first of these explanations but the attempt to fit a complete life table to these data reveals more clearly, than would consideration in isolation of the cohort data on orphanhood, that the estimates of adult mortality in the recent past cannot be relied upon.

In conclusion, it appears that most analysts have been too pessimistic about the scope for estimating the age pattern of mortality using indirect methods. With data of reasonable quality, one need not discard information on mortality patterns and splice together separate sections of 1-parameter model life tables nor assume that the age pattern of mortality has remained constant. Such methods produce implausible results and misrepresent the trend in mortality if the age pattern of mortality differs sharply from that in the models being used. An iterative procedure can be used instead, that yields a series of consistent and plausible fits of 2-parameter relational logit model life tables. Rather than assuming that $\beta = 1$ or some other fixed value, the method allows this parameter to change linearly over time. The procedure also averages across any curvature in the relationship between the observed logit survivorship function, and that in the standard, by making the reasonable assumption that all of the orphanhood estimates indicate the level of survivorship between two fixed and widely spaced ages. The results are better than those of approaches, based on fits to the lower half of the life table, that fail to take into account the instability of parameter estimates when there are complex divergences in the pattern of mortality from that in the standard or errors in the data. Equally, the method helps to clarify the plausibility of the basic estimates obtained from indirect methods and to indicate when they should not be used as the basis of a full life table.

Chapter 8

CONCLUSIONS

It is clear, at the beginning of the 1990's, that the health problems of adults in developing countries have become an area of mounting concern. This has been expressed in many ways. One is the Safe Motherhood Initiative launched in 1987 (Mahler, 1987). A second is the reviews and other activities being mounted in the field of adult health by the World Bank (Anon., 1989). On an academic front, one can cite the decision by the International Union for the Scientific Study of Population in 1989 to replace its Scientific Committee on Comparative Mortality with a Committee on Adult Mortality. The renewed interest of agencies, researchers and, increasingly, national governments in adult health issues has many roots. Perhaps the most profound is the onset of demographic aging in the wide range of developing countries in which fertility began to decline around 1970 and the rapid growth in the number of older people in developing country populations. A second factor is the incipient epidemics, in many developing countries, of particular diseases that predominantly affect adults. While AIDS is certainly the example of which there is most widespread awareness, other health problems, such as smoking related disease, are also of grave concern. Thirdly, new concern about adult health has arisen, in part, as a reaction to the enormous advances made during the last two decades in our understanding of the health problems of children, their causes and determinants and of ways to intervene to prevent childhood deaths. Related to this is the significant, if patchy and incomplete, progress made by many developing countries in applying this knowledge to reduce child mortality. These trends suggest that adult health and mortality in developing countries are set to become problems of increasing absolute and relative magnitude.

Against this background, the limited success that demography and related disciplines have achieved in their attempts to devise reliable and affordable ways of measuring adult mortality must also be an issue of major concern. Considerable ingenuity has been deployed to good effect, to devise ways of rendering incomplete vital statistics usable. As is discussed in chapter 1, such methods are only of value in countries in which the majority of deaths are registered. Moreover, while registration data are of incomparable value for the measurement of the level and trend in mortality nationally and regionally, they are inherently a blunt tool for the study of mortality differentials and for identification of risk factors that are a priority for detailed epidemiological investigation.

Few would now deny that the Brass method of measuring child mortality from data collected in single-round surveys has played a major role in the growth of knowledge about child mortality in poor countries. It is complemented by the, largely unanticipated, value of birth histories, such as those collected by the WFS, for mortality studies. Unfortunately, the development of survey methods for measuring adult mortality in developing countries has proved a more difficult challenge. The poor performance of simple questions about recent deaths in the household is discussed in chapter 1. Their record in censuses is appalling. There is certainly a case to be made that better questionnaires, interviewing procedures and training could yield improved results. Even so, very large studies are required to identify an adequate sample of deaths to permit analysis, limiting the value of such methods for the study of mortality differentials. The same problem affects multi-round surveys which, though expensive and logistically demanding operations, have yielded some of the more reliable estimates of adult mortality available for countries that lack effective death registration systems.

Given the limitations of other sources of information, this thesis focuses on a further way of measuring adult mortality in developing countries. It concentrates on the orphanhood method. At the present time, this is the most widely applied and successful of the methods that measure adult mortality indirectly from data on the deaths of particular relatives. One advantage of such indirect methods is that they are based on straightforward questions that can be included in single-round enquiries, including censuses. They are a valuable way of obtaining estimates of adult mortality for small areas and can be used to study other aspects of differential mortality. If the data are reasonably reliable, they can also be used to establish the overall trend in adult mortality over a number of years.

Like the statistics generated by other approaches to the measurement of adult mortality in developing countries, orphanhood estimates suffer from distinctive limitations. Inherently, they measure only the overall level and trend in mortality and cannot detect unusual age patterns of mortality or short-term fluctuations in mortality. For many purposes this does not matter. More seriously, in a number of applications in East Africa and elsewhere, the orphanhood method has yielded results that indicate unbelievably light mortality, implausibly rapid declines in mortality and gross inconsistencies between the estimates from successive enquiries (Blacker, 1984; Hill, 1984b). While the orphanhood method has performed much better in other applications, these findings have thrown its validity into doubt. The reservations of some demographers about the method, of which those expressed by one of its originators, Ken Hill, have probably been most influential, combined with the scepticism that naturally exists about a method whose theoretical basis is difficult for someone without a good training in demography to grasp, have had a major impact on those responsible for collecting demographic data. For example, some African countries that used the orphanhood questions in the 1970 round of censuses dropped them from their census schedule a decade later. Few of these countries collect information on adult mortality from other sources. Moreover, while questions about orphanhood were asked in a number of WFS surveys and most of those conducted in the first round of DHS enquiries, they have not been included in the core questionnaire for the second round of the DHS. This is unfortunate. The orphanhood method is no panacea, but it is premature to consign it to the waste bin reserved for good ideas that did not work.

This thesis aims to further understanding of the problems that affect orphanhood based estimates of adult mortality and to develop improved methods that circumvent these problems and yield more reliable estimates. Unfortunately, in countries where the orphanhood method is of most value, few reliable alternative sources of information exist against which the method can be validated or which can be used to establish the causes of errors in the estimates. Moreover, as is argued in chapter 3, assessments of the orphanhood method where such alternative sources do exist can be difficult to generalize to other countries. It has proved possible, however, to show that several of the possible candidates are not, in general, major sources of errors in orphanhood estimates. When you have eliminated the possible, only the probable remains.

The hypothesis that under-reporting of orphanhood, concentrated among young respondents due to the adoption effect, is the major source of errors in the estimates was posited soon after the method was first developed (Brass, 1975). It has always been the most likely explanation of the poor performance of the method in some populations and studies, that has done so much to discredit the approach. However, alternative hypotheses about why such errors occur, also exist in the literature. Concern has persisted about various selection biases and about the robustness of the estimation procedures, particularly for paternal orphanhood (Hill, 1984b; McDonald, 1987). Recently, worries have been expressed about the techniques and models used to infer complete mortality schedules from the estimates of survival over limited and varying age spans that result from the orphanhood method (Blacker and Mukiza-Gapere, 1988; Kamara, 1988). This thesis considers these issues.

The analyses in chapter 4 strongly suggest that the steep declines in mortality and inconsistencies between successive sets of orphanhood data observed in some countries do not stem from breaches in the assumptions involved in making the estimates. There *is* an appreciable degree of uncertainty attached to orphanhood estimates, even if the reports from which they are made are reliable. Such errors pale into insignificance compared with either the size of the errors that have shaken confidence in the method or the uncertainty that exists about the level of adult mortality in countries where it has to be guessed on the basis of data concerning children. In particular, estimates from paternal orphanhood are rather more reliable than the general tenor of the literature about the orphanhood method would suggest.

The analysis in chapter 4 also reveals that, because of the way the relationship between parental survival and life table indices of survival changes with the level of mortality, a regression-based approach to estimation, or one that combines the advantages of regression methods with those of the weighting factors, will tend to produce more accurate estimates of adult mortality than the original method. This is particularly true for males and for low mortality populations. Therefore, a regression based procedure for estimating adult male mortality from paternal orphanhood is developed for the first time. A set of estimation coefficients for female mortality based on consistent assumptions is also presented. Estimation of the coefficients for males takes advantage of the recent development of a male standard for use with a relational Gompertz model of fertility (Paget, 1988). However, the analysis in chapter 4 indicates that, although less reliable than the maternal orphanhood method, the paternal orphanhood method is less sensitive than some researchers have feared to variation in the distribution of men's ages at childbearing.

The second major possible source of bias, that it has been suggested may afflict orphanhood estimates, lies in the procedures used to fit complete life tables to individual estimates of survival. The discussion in chapter 7 argues that appreciable distortions in the estimates for populations with extreme age patterns of mortality can result either if the data are constrained to fit a 1-parameter model or if 2-parameter models are fitted to indices of survival over a narrow span of ages. An, admittedly complex, iterative procedure is proposed that reduces the impact on the estimates of non-linearities in the logit survivorship function relative to the standard. With extreme data, the procedure yields more plausible mortality schedules and more believable trends in mortality over time than existing methods. In addition, the analysis in chapter 7 offers further support for the argument that, when the orphanhood method performs very poorly, the source of the problem lies in the data, rather than the methods or models. In problematic applications, fitting a life table using the new method emphasizes, rather than reduces, the implausibility of the basic data. It is conceivable that some East African and other countries, where the orphanhood method has yielded poor results, have more extreme age patterns of mortality than any country for which reliable direct measures of mortality exist. It does not seem likely.

Chapter 3 presents estimates from the WFS for countries at a range of levels of development in different regions of the world. The majority of these countries have collected orphanhood data on more than one occasion and have also used other survey based approaches to estimate the level of adult mortality. Most of the orphanhood estimates seem to be of reasonable quality and the findings indicate that the poor performance of the method, which was first detected in East Africa and some poorer Central American countries, may be a more localized problem than was once feared. A review of the existing literature and some new material, suggest that the selection biases that affect the orphanhood method are only moderate in size. They are unlikely candidates as explanations of why the orphanhood method performs so poorly on occasion. Thus, experience with the method lends support to the hypothesis that under-reporting of orphanhood, due to the adoption effect, is the problem of central concern. Some direct evidence in support of this argument is reviewed in chapter 3 and further evidence of it is provided by the innovative analyses presented in chapters 5 and 6.

On the basis of this assessment of the orphanhood method, new techniques are proposed for estimating adult mortality from orphanhood that can be used when data are available from more than one survey or when supplementary questions are asked in a single survey. Two aims underlie the development of these methods. The first is to produce methods with a more specific time reference than those yielded by the original method. In particular, the value of the orphanhood estimates would be greatly enhanced if they can be made for more recent dates than by the basic method. Secondly, the aim is to develop methods that are less subject to bias due to the under-reporting of orphanhood at young ages.

The new coefficients for estimation of adult mortality from orphanhood after age 20, described in chapter 5, represent a refinement of earlier research (Timæus, 1986). They are intended for the analysis of period data on synthetic cohorts calculated either from the results of two enquiries or using a question on the date of orphanhood. While the coefficients yield similar estimates to those proposed by Timæus (1986), they are much more convenient to apply. In addition, the regression based procedure presented here should yield better results overall than the earlier method which was based on the original multiplying factors for estimating mortality from orphanhood. The latter method both distorts the relationship between mortality and orphanhood, when mortality is light, and lacks the robustness inherent in the averaging process by which the weighting factors estimate mortality.

The applications presented in chapter 5 indicate that period data on orphanhood in adulthood produce more accurate estimates than period data on lifetime orphanhood when reporting of orphanhood at young ages is biased downward by the adoption effect. This is because they prevent the chaining of changes in orphanhood to produce synthetic cohort estimates, transmitting this bias to the proportion of living parents in the population aged 20 and above. On the other hand, because the estimates from the new method are based only on the mortality of older parents, they are more sensitive than period estimates from lifetime orphanhood to other biases in the data. If the data come from two enquiries, such biases may arise from changes in the accuracy of reporting and, in particular, in the degree of age exaggeration by respondents. On the other hand, if the period estimates are calculated using a question on the timing of orphanhood, similar biases may result from errors in the reporting of dates of orphanhood.

The coefficients for orphanhood since marriage, that are presented in chapter 6, provide, for the first time, a way of estimating mortality from a novel form of data. These data have several potential advantages. They are unlikely to be biased significantly by the adoption effect, they can yield recent mortality estimates from data on adults, yet they can be collected in a single enquiry by means of a simple question that only requires respondents to recall the relative timing of major life-cycle events, not calendar dates. Further experience with this approach is needed before a confident assessment of its performance can be made. Like estimates based on orphanhood after age 20, the results are sensitive to the accuracy of age reporting and it seems unlikely that they will prove useful for measuring recent fluctuations in adult mortality. The initial applications presented here, however, indicate that the method is a promising way of estimating adult mortality at a date about six years prior to a single-round survey.

Questions on the timing of orphanhood relative to marriage also make it possible to estimate adult mortality from orphanhood before marriage. Coefficients are also presented for the first time that facilitate the analysis of such data. In some populations, the quality of the data on orphanhood before marriage is evidently very poor. This is almost certainly due to the adoption effect. Elsewhere the results seem very plausible. In such populations, this variant on the orphanhood method has two valuable characteristics. First, it measures mortality over a limited and fairly clearly defined interval of time and range of ages. Secondly, the estimates obtained from different age groups of respondents are directly comparable and can be used to measure mortality over an period extending back to 30 or more years before the data were collected.

The sets of coefficients for making estimates from different forms of orphanhood data presented in this thesis provide a means of obtaining more accurate estimates of male mortality from paternal orphanhood than previously, of readily estimating mortality from period measures of the incidence of orphanhood experienced by young adults and of analyzing data on orphanhood before and after marriage. One can envisage the development in future of more sophisticated estimation procedures, calculated from larger sets of simulated populations, based on more realistic demographic models and using more parameters. The gains from this are unlikely to be great. In 1977, Hill and Trussell argued:

> ‘The methodology of indirect estimation is now likely to follow divergent patterns. On the one hand, the methods pioneered by Brass are certain to be employed as tools in the foreseeable future. (...) They are simple to use and their application can be readily understood by those without much formal training in demography. In addition, they make few demands on the data. On the other hand, it is now a relatively easy task to program a reverse survival technique which iterates to a feasible solution. The straightforwardness of the methodology, the relatively small number of assumptions and the gradual improvement of data quality over time guarantee that this approach will gain increasing favour.’

In 1977 this view was perhaps premature. In the 1990's, with the developments in microcomputer hardware and software that have occurred in the interim and their widespread dissemination to developing countries, one would hope that it is no longer.

The procedures developed here represent a way of estimating adult mortality from various forms of orphanhood data quickly and conveniently, with a minimum of data. Yet, appreciable errors in the results can be expected in particular populations where age patterns of childbearing, mortality and marriage differ from those used to derive the coefficients. The data on respondents' parents needed to apply the technique of reverse survival are seldom available. Nevertheless, given the increasing diversity in demographic terms of developing country populations, it is doubtful that the best way to improve orphanhood estimates further is to produce different sets of estimation coefficients for use in populations with particular characteristics or to attempt to introduce additional parameters that capture, for example, the impact of differences in the variance of ages at childbearing. Instead, estimates can be calculated readily from the basic formulae, using fertility, mortality and nuptiality patterns and age distributions that are appropriate for the population under study. The exact methodology can be tailored to the scope, historical depth and reliability of the data available. Thus, it is hoped that the investigation of the orphanhood method presented here will help to stimulate more imaginative and detailed analyses of adult mortality in poor countries.

REFERENCES

Adlakha, A. and Nizamuddin, M. (1984) Mortality data collections: a review of integrated multi-purpose household surveys and multi-round demographic surveys, in *Data Bases for Mortality Measurement*. New York: United Nations.

Al-Jem, M., Timæus, I. and Aoun, S. (1988) La Mortalité au Maroc d'apres des résultats de l'E.N.P.S, in *The Determinants of Health and Mortality in Africa*, Hill, A.G. (ed). New York: The Population Council & IRD/Macro.

Anon. (1989) Adult mortality in developing countries. *The Lancet*, **1**, 1150, May 20th.

Bennett, N.G. and Horiuchi, S. (1981) Estimating the completeness of death registration in a closed population. *Population Index*, **47**, 207-221.

Blacker, J.G.C. (1974) The estimation of vital rates in Kenya and Uganda, in *Population in African Development*, Cantrelle, P. (ed). Liège: Ordina.

Blacker, J.G.C. (1977) The estimation of adult mortality in Africa from data on orphanhood. *Population Studies,* **31**, 107-128.

Blacker, J.G.C. (1984) Experience in the use of special mortality questions in multi-purpose surveys: the single-round approach, in *Data Bases for Mortality Measurement*. New York: United Nations.

Blacker, J.G.C., Hill, A.G. and Moser, K. (1983) *Mortality Levels and Trends in Jordan from the Results of the 1976 Fertility Survey*. WFS Scientific Reports, 47. Voorburg: International Statistical Institute.

Blacker, J.G.C., Hill, A.G. and Timæus, I. (1985) Age patterns of mortality in Africa: an examination of recent evidence, in *International Population Conference, Florence, 1985*. Liège: International Union for the Scientific Study of Population.

Blacker, J.G.C. and Mukiza-Gapere, J. (1988) The indirect measurement of adult mortality in Africa, in *African Population Conference, Dakar, 1988*. Liège: International Union for the Scientific Study of Population.

Blacker, J.G.C. and Scott, C.S. (1974) *Manual on Demographic Sample Surveys in Africa*. UNECA-UNESCO.

Booth, H. (1984) Transforming Gompertz's function for fertility analysis: the development of a standard for the relational Gompertz function. *Population Studies*, **38**, 495-506.

Brass, W. (1971a) A critique of methods for estimating population growth in countries with limited data. *Bulletin of the International Statistical Institute*, **44**, 397-412.

Brass, W. (1971b) On the scale of mortality, in *Biological Aspects of Demography*, Brass, W. (ed). London: Taylor and Francis.

Brass, W. (1975) *Methods for Estimating Fertility and Mortality from Limited and Defective Data*. Occasional Publication. Chapel Hill: International Program of Laboratories for Population Statistics.

Brass, W. (1979) A procedure for comparing mortality measures calculated from intercensal survival with the corresponding estimates from registered deaths. *Asian and Pacific Census Forum*, **6(2)**, 5-7.

Brass, W. (1980) Paths and problems in the analysis of mortality differentials, in *Proceedings of the UN/WHO Meeting on Socio-Economic Determinants and Consequences of Mortality, Mexico City, 19-25 June, 1979*. New York: United Nations.

Brass, W. (1985) *Advances in Methods for Estimating Fertility and Mortality from Limited and Defective Data*. Centre for Population Studies Occasional Paper. London: London School of Hygiene and Tropical Medicine.

Brass, W. (1986) `The estimation of adult mortality from survey reports on survivorship of relatives'. Unpublished manuscript, London School of Hygiene and Tropical Medicine.

Brass, W. and Bamgboye, E. (1981) *The Time Location of Reports of Survivorship: Estimates for Maternal and Paternal Orphanhood and the Ever-Widowed*. Centre for Population Studies Research Papers, 81-1. London: London School of Hygiene and Tropical Medicine.

Brass, W. and Hill, K. (1973) Estimating adult mortality from orphanhood, in *International Population Conference, Liège, 1973*. Liège: International Union for the Scientific Study of Population.

Carrier, N. and Hobcraft, J. (1971) *Demographic Estimation for Developing Countries*. London: Population Investigation Committee.

Chackiel, J. and Orellana, H. (1985) Adult female mortality trends from retrospective questions about maternal orphanhood included in censuses and surveys, in *International Population Conference, Florence, 1985*. Liège: International Union for the Scientific Study of Population.

Clairin, R. (1974) The estimation of adult mortality from data on the survival of parents, in *Population in African Development*, Cantrelle, P. (ed). Liège: Ordina.

Coale, A.J. and McNeil, D.R. (1972) The distribution by age of the frequency of first marriage in a female cohort. *Journal of the American Statistical Association*, **67**, 743-749.

Coale, A.J. and Trussell, T.J. (1974) Model fertility schedules: variations in the age structure of childbearing in human populations. *Population Index*, **40**, 185-258.

Condé, J., Fleury-Brousse, M. and Waltisperger, D. (1980) *Mortality in Developing Countries*. Paris: OECD.

Courbage, Y. and Fargues, P. (1979) A method for deriving mortality estimates from incomplete vital statistics. *Population Studies*, **33**, 165-180.

Ewbank, D.C., Gomez de Leon, J.C. and Stoto, M.A. (1983) A reducible four-parameter system of model life tables. *Population Studies*, **37**, 105-127.

Feeney, G. (1980) Estimating infant mortality trends from child survivorship data. *Population Studies*, **34**, 109-128.

Goldman, N., Rutstein, S.O. and Singh, S. (1985) *Assessment of the Quality of Data in 41 WFS Surveys: A Comparative Approach*. WFS Comparative Studies, 44. Voorburg: International Statistical Institute.

Graham, W.J. (1986) *Health Status Indicators in Developing Countries: a Selective Review*. London: The Commonwealth Secretariat.

Graham, W., Brass, W. and Snow, R.W. (1989) Estimating maternal mortality: the sisterhood method. *Studies in Family Planning*, **20**, 125-135.

Henry, L. (1960) Mesure indirecte de la mortalité des adultes. *Population*, **15**, 457-466.

Hill, A.G. and Graham, W.J. (1988) *West African Sources of Health and Mortality Information: a Comparative Review*. Technical Study 58e. Ottawa: International Development Research Centre.

Hill, K. (1977) Estimating adult mortality levels from data on widowhood, *Population Studies*, **31**, 75-84.

Hill, K. (1981) Comments by the discussant on the session on advances in methods for estimating demographic parameters in populations with limited data, in *International Population Conference, Manila, 1981*. Liège: International Union for the Scientific Study of Population.

Hill, K. (1984a) Overview, in *Data Bases for Mortality Measurement*. New York: United Nations.

Hill, K. (1984b) An evaluation of indirect methods for estimating mortality, in *Methodologies for the Collection and Analysis of Mortality Data*, Vallin, J., Pollard, J.H. and Heligman, L. (eds). Liège: Ordina.

Hill, K., Behm, H. and Soliz, A. (1976) *La Situación de la Mortalidad en Bolivia*. La Paz: Instituto Nacional de Estadística.

Hill, K. and Trussell, T.J. (1977) Further developments in indirect mortality estimation. *Population Studies*, **31**, 313-333.

Kamara, J.A.L. (1988) The use of model life tables in the estimation of mortality trends for African countries, in *African Population Conference, Dakar, 1988*. Liège: International Union for the Scientific Study of Population.

Kim, O-K. (1986) Estimation of adult mortality in Korea: levels, trends and socioeconomic differentials. *Journal of Biosocial Science*, **18**, 347-356.

Ledermann, S. and Breas, J. (1959) Les dimensions de la mortalité. *Population*, **14**, 637-682.

Lotka, A.J. (1939) *Théorie Analytique des Associations Bioloqiques*. Deuxième partie. Paris: Hermann et C[ie].

McDonald, P. (1987) An evaluation of the parental survival approach to estimating adult mortality, in *The Survey under Difficult Conditions: Population Data Collection and Analysis in Papua New Guinea*, McDevitt, T.M. (ed). New Haven: Human Relations Area Files.

Mahler, H. (1987) The Safe Motherhood initiative: a call to action. *The Lancet*, **1**, 668-670, March 21st.

Makinson, C. (1988) `Indirect estimation of adult mortality'. Unpublished manuscript, Institute for Resource Development

Mammo, A. (1988) `Mortality in Ethiopia: Levels, Trends and Differentials'. Unpublished PhD thesis, University of Pennsylvania.

Moser, K. (1985) *Levels and Trends in Child and Adult Mortality in Peru*. WFS Scientific Reports, 77. Voorburg: International Statistical Institute.

Murray, C.J.L. and Feachem, R.G. (1990) Adult mortality in developing countries. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **84**, 1-2.

National Statistical Office, Malaŵi (1973) *Malaŵi Population Change Survey, February 1970 - January 1972*. Zomba: The Government Printer.

National Statistical Office, Malaŵi (1987) *Malaŵi Demographic Survey, 1982*. Zomba: The Government Printer.

Nault, F., Boleda, M. and Légaré, J. (1986) Estimation de la mortalité des adultes à partir des proportions d'orphelins. *Population*, **41**, 749-762.

Padmanabha, P. (1984) Use of sample registration systems for studying levels, trends and differentials in mortality: the experience of India, in *Data Bases for Mortality Measurement*. New York: United Nations.

Page, H.J. and Wunsch, G. (1976) Parental survival data: some results of the application of Ledermann's model life tables. *Population Studies*, **30**, 59-76.

Paget, W.J. (1988) `The Development of a Male Fertility Standard and its Applications to the Paternal Orphanhood Time Location Method'. Unpublished MSc thesis, London School of Hygiene and Tropical Medicine.

Palloni, A. and Heligman, L. (1986) Re-estimation of the structural parameters to obtain estimates of mortality in developing countries. *Population Bulletin of the United Nations*, **18**, 10-33.

Palloni, A., Massagli, M. and Marcotte, J. (1984) Estimating adult mortality with maternal orphanhood data: analysis of sensitivity of the techniques. *Population Studies*, **38**, 255-279.

Pison, G. and Langaney, A. (1985) The level and age pattern of mortality in Bandafassi (Eastern Senegal): results from a small-scale and intensive multi-round survey. *Population Studies*, **39**, 387-405.

Pison, G. and Langaney, A. (1988) Age patterns of mortality in Eastern Senegal: a comparison of micro and survey approaches, in *Micro-Approaches to Demographic Research*, Caldwell, J.C., Hill, A.G. and Hull V.J. (eds). London: Kegan Paul International.

Preston, S.H. (1983) An integrated system for demographic estimation from two age distributions. *Demography*, **20**, 213-226.

Preston, S.H. (1984) Use of direct and indirect techniques for estimating the completeness of death registration systems, in *Data Bases for Mortality Measurement*. New York: United Nations.

Preston, S.H. and Bennett, N.G. (1983) A census-based method for estimating adult mortality. *Population Studies*, **37**, 91-104.

Preston, S.H. and Chen, N. (1984) `Two census orphanhood methods for estimating adult mortality, with applications to Latin America'. Unpublished manuscript, University of Pennsylvania.

Preston, S.H., Coale, A.J., Trussell, J. and Weinstein, M. (1980) Estimating the completeness of reporting of adult deaths in populations that are approximately stable. *Population Index*, **46**, 179-202.

Preston, S.H. and Coale, A.J. (1982) Age structure, growth, attrition and accession: a new synthesis. *Population Index*, **48**, 217-259.

Preston, S.H. and Hill, K. (1980) Estimating the completeness of death registration. *Population Studies*, **34**, 349-366.

Scott, C. (1973) *Technical Problems of Multiround Demographic Surveys*. Reprint Series 11. Chapel Hill: International Program of Laboratories for Population Statistics.

Seltzer, W. (1969) Some results from Asian population growth studies. *Population Studies*, **33**, 395-406.

Somoza, J. (1981) An evaluation of the performance of indirect estimation techniques in the analysis of defective data, in *International Population Conference, Manila, 1981*. Liège: International Union for the Scientific Study of Population.

Sullivan, J. (1972) Models for the estimation of the probability of dying between birth and exact ages in early childhood. *Population Studies*, **22**, 79-97.

Tabutin, D. (1984) Comparison of single- and multi-round surveys for measuring mortality in developing countries, in *Methodologies for the Collection and Analysis of Mortality Data*, Vallin, J., Pollard, J.H. and Heligman, L. (eds). Liège: Ordina.

Timæus, I. (1984) *Mortality in Lesotho: a Study of Levels, Trends and Differentials Based on Retrospective Survey Data*. WFS Scientific Reports, 59. Voorburg: International Statistical Institute.

Timæus, I. (1986) An assessment of methods for estimating adult mortality from two sets of data on maternal orphanhood. *Demography*, **23**, 435-450.

Timæus, I. (1987a) Estimation of fertility and mortality from WFS household surveys, in *The World Fertility Survey: An Assessment*, Cleland, J. and Scott, C. (eds). Oxford: Clarendon Press.

Timæus, I. (1987b) Adult mortality in Mauritania, in *Evaluation de l'Enquête Nationale Mauritanienne sur la Fecondité*. WFS Scientific Reports, 83. Voorburg: International Statistical Institute.

Timæus, I.M. (1988) Regional differentials in adult mortality in sub-Saharan Africa, in *African Population Conference, Dakar, 1988*. Liège: International Union for the Scientific Study of Population.

Timæus, I.M. (1990, in press) Adult mortality: levels, trends and data sources, in *Disease and Mortality in sub-Saharan Africa*, Feachem, R.G. and Jamison, D.T. (eds). New York: Oxford University Press for The World Bank.

Timæus, I.M. and Graham, W.J. (1989) *Measuring Adult Mortality in Developing Countries: A Review and Assessment of Methods*. Planning Policy and Research Working Papers, WPS 155. Washington DC: The World Bank.

Timæus, I., Harpham, T., Price, M. and Gilson, L. (1988) Health surveys in developing countries: the objectives and design of an international programme. *Social Science and Medicine*, **24**, 359-368.

Trussell, T.J. (1975) A re-estimation of the multiplying factors for the Brass technique for determining childhood survivorship rates. *Population Studies*, **24**, 97-108.

UN (1955) *Age and Sex Patterns of Mortality: Model Life Tables for Underdeveloped Countries*. New York: United Nations.

UN (1967) *Methods of Estimating Basic Demographic Measures from Incomplete Data*. New York: United Nations.

UN (1982) *Model Life Tables for Developing Countries*. New York: United Nations.

UN (1983) *Indirect Techniques for Demographic Estimation*. New York: United Nations.

UN (1984) *Data Bases for Mortality Measurement*. New York: United Nations.

van Poppel, F. and Bartlema, J. (1985) *Levels of Orphanhood and Measurement of Adult Mortality in Populations of the Past: the Case of The Netherlands (The Hague, 1850-1880)*. NIDI Working Papers, 52. Voorburg: Netherlands Interuniversity Demographic Institute.

Zaba, B. (1979) The four-parameter logit life table system. *Population Studies*, **33**, 79-100.

Zlotnik, H. and Hill, K. (1981) The use of hypothetical cohorts in estimating demographic parameters under conditions of changing fertility and mortality. *Demography*, **18**, 103-122.

Appendix 1

NOTATIONAL CONVENTIONS

a	age in years (of children exposed to the risk of orphanhood) in continuous formulations
B	exact base age from which life table survival is estimated by orphanhood methods
b	rounded base age from which life table survival is estimated by orphanhood methods
C(a,y)	number of individuals aged *a* born to mothers (or fathers) aged *y* at the time of the birth
c	scaling factor for the level of fertility in Brass's cubic model
f(y)	birth rate at exact age *y*
f_a	a function of age, *a*, derived from a standard life table
${}_ag_y$	mean time since death of those who died between ages *y* and y+a and would have been aged y+a at the reference date
h	interval in years between two demographic enquiries
i	used as an index, usually to refer to the *i*th age group
K	rate of mortality decline, measured as a linear trend in the α parameter of a relational logit system of model life tables
k_s	a constant term derived from a standard life table
$l_s(a)$	probability of surviving to age *a* in a standard life table
$l(a,t)$	life table probability of surviving to age *a* at time *t*
$\overline{M}$	mean age of mothers (or fathers) at the birth of their children
$\overline{m}$	mean age at first marriage
m(y)	rate of first marriage at exact age *y*
N	central age of a five-year age group
N(a)	number of individuals of exact age *a*
NO(a)	number of individuals with a living mother (or father) of exact age *a*
n	central age, lying between two five-year age groups
O(a)	the proportion of the population whose mothers (or fathers) have died by exact age *a*
P_i	mean parity of women in age group *i*, where 1 = 15-19 years *etc*.
${}_ap_y$	life table probability of surviving from age *y* to age y+a, $l(y+a)/l(y)$
${}_ap_{y,s}$	probability of surviving from age *y* to age y+a in a standard life table
${}_ap_y(t)$	life table probability of surviving from age *y* to age y+a at time *t*
r(a)	rate of population increase at exact age *a*

$r^{NO}(a)$ rate of increase in the number of individuals with living mothers (or fathers) at exact age *a*

$r^{S}(a)$ rate of increase in the proportion of the population with living mothers (or fathers) at exact age *a*

${}_5r_x(t)$ rate of increase of the age group *x* to x+5 at time *t*

${}_5r_x^S(t)$ rate of increase in the proportion of the population with living mothers (or fathers) in the age group *x* to x+5 at time *t*

$S(a)$ proportion of the population of exact age *a* whose mothers (or fathers) are alive

$S^m(a)$ proportion of the ever-married population of exact age *a* whose mothers (or fathers) were alive at the time that they first married

${}_5S_x$ proportion of the population aged *x* to x+5 whose mothers (or fathers) are alive

${}_5S_x^m$ proportion of the ever-married population aged *x* to x+5 whose mothers (or fathers) were alive at the time that they first married

${}_5S_x(t)$ proportion of the population aged *x* to x+5 whose mothers (or fathers) are alive at time *t*

${}_5S_x(\tau)$ proportion of the population aged *x* to x+5 whose mothers (or fathers) are alive in a synthetic cohort referring to time $\bar{t}$

s earliest age at which childbearing occurs

T interval in years since life table measures estimated from orphanhood equalled the equivalent period measure

t time at which a survey was conducted or to which a mortality measure refers

$\bar{t}$ time reference of average measures of experience over an interval *t* to t+h

${}_av_y$ the exact weights relating life table survival to the proportion orphaned for individuals aged *a* whose mothers (or fathers) were aged *y* when they were born, $C(a,y)/\int_s^w C(a,y)$

W_n weighting factor applied to the proportion of respondents with surviving parents in the two age groups surrounding *n* to estimate life table survivorship

w the oldest age at which childbearing occurs

x age in years (of children exposed to the risk of orphanhood) in discrete formulations

x_a a fixed age in adulthood

x_c a fixed age in childhood

$Y(a)$ the logit of $l(a)$, $\frac{1}{2}\ln[\{1-l(a)\}/l(a)]$

$Y_s(a)$ the logit of $l_s(a)$

y age in years of a mother (or father) at the time of the birth of a child

z age in years, used in expressions involving integration or summation to the current age

α the level parameter in a relational logit system of model life tables

α_f	the timing parameter in the relational Gompertz fertility model
β	the slope of mortality parameter in a relational logit system of model life tables
β_f	the parameter determining the spread of the distribution in the relational Gompertz fertility model
$\gamma(a)$	instantaneous rate of orphanhood at exact age *a*
$\lambda_i(x)$	*i*th coefficient, referring to age group *x* to x+5
$\mu(a)$	force of mortality at exact age *a*
$\mu(a,t)$	force of mortality at exact age *a* and time *t*
τ	time reference equivalent to $\overline{t}$ for synthetic cohort measures.

Appendix 2

COEFFICIENTS FOR ORPHANHOOD ESTIMATION

2.1 Coefficients for maternal orphanhood

Estimation of l(25+n)/l(25) from proportions of mothers alive

$$l(25+n)/l(25) = \lambda_0(n) + \lambda_1(n)\overline{M} + \lambda_2(n)\,{}_5S_{n-5}$$

n	$\lambda_0(n)$	$\lambda_1(n)$	$\lambda_2(n)$	R^2
10	-0.3012	0.00129	1.2668	0.997
15	-0.1769	0.00227	1.1165	0.996
20	-0.1556	0.00376	1.0562	0.995
25	-0.1850	0.00591	1.0305	0.993
30	-0.2556	0.00890	1.0264	0.992
35	-0.3696	0.01292	1.0438	0.992
40	-0.5236	0.01799	1.0828	0.992
45	-0.6924	0.02342	1.1364	0.993
50	-0.8077	0.02712	1.1763	0.992

2.2 Coefficients for paternal orphanhood

Estimation of l(35+n)/l(35) from proportions of fathers alive

$$l(35+n)/l(35) = \lambda_0(n) + \lambda_1(n)\overline{M} + \lambda_2(n)\,{}_5S_{n-5} + \lambda_3(n)\,{}_5S_n$$

n	$\lambda_0(n)$	$\lambda_1(n)$	$\lambda_2(n)$	$\lambda_3(n)$	R^2
10	-0.7391	0.00279	2.5478	-0.9079	0.975
15	-0.3475	0.00613	1.3964	-0.2523	0.972
20	-0.3079	0.01075	0.5758	0.3966	0.974
25	-0.4730	0.01593	0.2645	0.7348	0.980
30	-0.7140	0.02101	0.3332	0.7767	0.988
35	-0.9231	0.02512	0.5203	0.6948	0.992
40	-0.9991	0.02647	0.6072	0.6476	0.987

2.3 Coefficients for maternal orphanhood after age 20

Estimation of l(25+n)/l(45) from proportions of mothers alive among those with living mothers at age 20

$$l(25+n)/l(45) = \lambda_0(n) + \lambda_1(n)\overline{M} + \lambda_2(n)\,{}_5S_{n-5}(\tau)/S(20,\overline{t})$$

n	$\lambda_0(n)$	$\lambda_1(n)$	$\lambda_2(n)$	R^2
25	-0.9062	0.00299	1.8292	0.981
30	-0.3964	0.00685	1.2203	0.983
35	-0.4471	0.01204	1.1432	0.984
40	-0.6107	0.01854	1.1550	0.987
45	-0.8081	0.02550	1.2007	0.991
50	-0.9428	0.03036	1.2361	0.992

2.4 Coefficients for paternal orphanhood after age 20

Estimation of l(35+n)/l(55) from proportions of fathers alive among those with living fathers at age 20

$$l(35+n)/l(55) = \lambda_0(n) + \lambda_1(n)\overline{M} + \lambda_2(n)\,{}_5S_{n-5}(\tau)/S(20,\overline{t}) + \lambda_3(n)\,{}_5S_n(\tau)/S(20,\overline{t})$$

n	$\lambda_0(n)$	$\lambda_1(n)$	$\lambda_2(n)$	$\lambda_3(n)$	R^2
25	-0.1135	0.00751	0.2610	0.6304	0.984
30	-0.7005	0.01552	0.7219	0.5227	0.990
35	-1.0421	0.02265	1.0299	0.3450	0.991
40	-1.1462	0.02637	1.0586	0.2963	0.979

2.5 Coefficients for maternal orphanhood after marriage

Estimation of l(25+n)/l(45) from proportions of mothers alive among women with living mothers when they married

$$l(25+n)/l(45) = \lambda_0(n) + \lambda_1(n)\overline{M} + \lambda_2(n)\overline{m} + \lambda_3(n)\,{}_5S_{n-5}/{}_5S^m_{n-5} + \lambda_4(n)\,{}_5S_n/{}_5S^m_n$$

n	$\lambda_0(n)$	$\lambda_1(n)$	$\lambda_2(n)$	$\lambda_3(n)$	$\lambda_4(n)$	R^2
30	0.5617	0.00836	-0.00261	-1.1231	1.4199	0.964
35	0.0476	0.01396	-0.00536	-0.3916	1.1354	0.966
40	-0.3715	0.01966	-0.00744	0.5394	0.5286	0.976
45	-0.6562	0.02587	-0.00716	1.0208	0.1789	0.987
50	-0.8341	0.03045	-0.00561	1.1898	0.0541	0.990

2.6 Coefficients for paternal orphanhood after marriage

Estimation of l(35+n)/l(55) from proportions of fathers alive among women with living fathers when they married

$$l(35+n)/l(55) = \lambda_0(n) + \lambda_1(n)\overline{M} + \lambda_2(n)\overline{m} + \lambda_3(n)\,{}_5S_{n-5}/{}_5S^m_{n-5} + \lambda_4(n)\,{}_5S_n/{}_5S^m_n$$

n	$\lambda_0(n)$	$\lambda_1(n)$	$\lambda_2(n)$	$\lambda_3(n)$	$\lambda_4(n)$	R^2
30	0.0795	0.01588	-0.00633	-1.2070	1.8284	0.979
35	-0.5289	0.02273	-0.01083	-0.2509	1.3867	0.980
40	-0.8477	0.02622	-0.01135	0.6057	0.7198	0.974

2.7 Coefficients for maternal orphanhood before marriage

Estimation of l(45)/l(25) from proportions of
women with living mothers when they married

$$l(45)/l(25) = \lambda_0(n) + \lambda_1(n)\overline{M} + \lambda_2(n)\overline{m} + \lambda_3(n)_5S_n^m + \lambda_4(n)\overline{m}_5S_n^m$$

n	$\lambda_0(n)$	$\lambda_1(n)$	$\lambda_2(n)$	$\lambda_3(n)$	$\lambda_4(n)$	R^2
25	-0.9607	0.00418	0.04466	1.8178	-0.04291	0.988
30	-0.9921	0.00429	0.04700	1.8428	-0.04501	0.988
35	-1.0129	0.00433	0.04822	1.8607	-0.04611	0.988
40+	-1.0206	0.00434	0.04861	1.8680	-0.04648	0.988

2.8 Coefficients for paternal orphanhood before marriage

Estimation of l(55)/l(35) from proportions of
women with living fathers when they married

$$l(55)/l(35) = \lambda_0(n) + \lambda_1(n)\overline{M} + \lambda_2(n)\overline{m} + \lambda_3(n)_5S_n^m + \lambda_4(n)\overline{m}_5S_n^m$$

n	$\lambda_0(n)$	$\lambda_1(n)$	$\lambda_2(n)$	$\lambda_3(n)$	$\lambda_4(n)$	R^2
25	-1.2640	0.01060	0.04480	1.8383	-0.04007	0.969
30	-1.2897	0.01068	0.04652	1.8530	-0.04124	0.969
35	-1.3123	0.01070	0.04769	1.8726	-0.04225	0.970
40+	-1.3152	0.01070	0.04783	1.8753	-0.04238	0.970

Appendix 3

ORPHANHOOD DATA, BURUNDI AND UGANDA

3.1 Enquête Démographique et de Santé au Burundi, 1987

	Proportion of parents surviving				
Age group	At interview	Synthetic cohort 0-4 years ago*	Synthetic cohort 5-9 years ago*	At marriage	Since marriage
A) Mothers					
15-19	0.9040	0.9040	0.8803		
20-24	0.8314	0.8606	0.8430		
25-29	0.7852	0.7880	0.7844	0.8786	0.8851
30-34	0.7141	0.7248	0.7107	0.8418	0.8444
35-39	0.5567	0.5935	0.6121	0.8277	0.6701
40-44	0.4764	0.4448	0.5201	0.8388	0.5633
45-49	0.3694	0.3049		0.7849	0.4734
			$\overline{M}$ = 27.3, $\overline{m}$ = 19.6		
B) Fathers					
15-19	0.7751	0.7751	0.8131		
20-24	0.7206	0.6870	0.7351		
25-29	0.6499	0.6114	0.6685	0.7644	0.8494
30-34	0.4868	0.4961	0.5771	0.7065	0.6852
35-39	0.4412	0.3845	0.4892	0.7368	0.6000
40-44	0.2909	0.2329	0.3426	0.7033	0.4167
45-49				0.6377	
			$\overline{M}$ = 32.2, $\overline{m}$ = 19.6		

* Makinson (1988).

3.2 Uganda Trypanosomiasis Study, Demographic Baseline Survey, 1988

Age group	Proportion of parents surviving				
	At interview	5 years ago	10 years ago	At marriage	Since marriage
A) Mothers - female respondents					
0- 4	0.9955	0.9929	0.9940		
5- 9	0.9812	0.9791	0.9771		
10-14	0.9731	0.9688	0.9477		
15-19	0.9563	0.9278	0.9345		
20-24	0.9061	0.8943	0.8742		
25-29	0.8689	0.8097	0.8065	0.9387	0.9231
30-34	0.7742	0.7097	0.7194	0.9115	0.8511
35-39	0.6694	0.6735	0.6302	0.9061	0.7411
40-44	0.6071	0.5052	0.3966	0.8878	0.6839
45-49	0.3958	0.3240	0.3200	0.8594	0.4606
50-54				0.8156	0.3767
			$\overline{M} = 26.4$, $\overline{m} = 17.8$		
B) Fathers - female respondents					
0- 4	0.9919	0.9941	0.9821		
5- 9	0.9812	0.9686	0.9563		
10-14	0.9507	0.9292	0.9422		
15-19	0.8854	0.8953	0.8858		
20-24	0.8267	0.8203	0.7896		
25-29	0.7421	0.6861	0.6948	0.9016	0.8228
30-34	0.6019	0.5301	0.5153	0.8656	0.6966
35-39	0.4418	0.3980	0.3646	0.8204	0.5419
40-44	0.3214	0.2969	0.2123	0.8010	0.4013
			$\overline{M} = 35.6$, $\overline{m} = 17.8$		