Estimating adult mortality in South Africa using information on the year-of-death of parents from the 2016 Community Survey

Chibwe Mambwe
University of Cape Town

A dissertation submitted to the Faculty of Commerce of the
University of Cape Town in partial fulfilment of the requirements
for the degree of Master of Philosophy in Demography

Centre for Actuarial Research (CARe)
2019

~~PLAGIARISM DECLARATION~~

I confirm that this dissertation is my own work, produced with the contributions and support from my supervisor, and has not been submitted to any other university for examination. Exceptional care has been taken to acknowledge, cite and reference all the work from others that have contributed to this dissertation.

Signature:

Signed by candidate

Date: 18/12/2019

ABSTRACT

In developing countries, systems that collect vital statistics are usually inadequate to facilitate the direct estimation of adult mortality. This has necessitated the development of indirect methods such as the orphanhood method. These methods are however limited, i.e., the single-survey approach produces out of date estimates of mortality and the two-survey approach is affected by the differential reporting of orphanhood between two surveys. To avoid these limitations, this research considers an extension of the orphanhood approach pioneered by Chackiel and Orellana (1985) to estimate adult mortality using year-of-death data rather than the conventional form of the orphanhood data. This is because the year-of-death data can be used to produce accurate time locations to which estimates of mortality apply but more important, one can create a synthetic survey from a single survey and hence obtain more recent and accurate estimates of mortality.

The single-survey orphanhood method is applied to survey data to obtain estimates of adult mortality and time location. A variation of the two-survey orphanhood method (Timæus 1991b) is also applied to survey data and the synthetic survey that is created from year-of-death data in order to derive estimates of adult mortality. In addition, the age range of respondents is extended down to age 0 to include year-of-death data from younger respondents on the assumption that underestimating orphanhood due to the adoption effect is minimal. This is done to investigate if the estimates derived from the two-survey method can be improved. Further, a cohort survival method that involves the calculation of a survival ratio for each age group at the first survey and the equivalent older ages groups at the second survey is applied to investigate the possibility of producing useful estimates of adult mortality based on cohort survival.

The level and trend in mortality estimates calculated from the single-survey, two - survey and the cohort survival approaches are discussed and compared to the estimates from the Rapid Mortality Surveillance (RMS) which are used as a benchmark for the trend and level of adult mortality in South Africa. The estimates produced using the single-survey method appear too low, while those from the two-survey method appear to be reasonable for the conventional from of the orphanhood data. Extending the two-survey method to include younger respondents produces estimates that are too low indicating that both the conventional form of the orphanhood data and the year-of-death data suffer from the adoption effect. The cohort survival approach produces reasonable estimates that are consistent with the RMS benchmark for both the conventional form of the orphanhood data and year-of-death data.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank God for his grace. I would like to express my utmost gratitude to my supervisor, Professor Rob Dorrington, for his tutelage, patience, and contribution throughout the completion of this research. I would like to thank my family and friends for their support during my studies. Finally, I would like to express my gratitude to the Hewlett and Mellon Foundation for their financial assistance towards my degree.

1 INTRODUCTION

1.1 Background

The inadequacy of data collection systems in most developing countries has resulted in the development of indirect methods to estimate demographic parameters. Such methods are called indirect to distinguish them from methods that produce demographic estimates directly from data on the occurrence of events and person-years of exposure (Zlotnik and Hill 1981). According to Zlotnik and Hill (1981), the basic idea underlying indirect estimation is to get away from the use of data of uncertain quality from which demographic parameters could be directly calculated (for example, births or deaths in a particular year), and to move towards the use of information which could be collected more reliably but which is only indirectly related to the required demographic parameters.

Hill (2001) classifies methods of estimating adult mortality into three broad groups: methods based on intercensal survival, methods that assess the completeness of death recording relative to census enumeration, and methods that convert indicators of mortality levels based on survival of close relatives into standard life table functions. This dissertation will focus on the third of these approaches.

Brass is credited for developing the first formal method for converting indicators of mortality based on survival of close relatives into standard life table measures such as ${}_{45}q_{15}$ (Hill, Choi and Timæus 2005). Brass and Hill (1973) proposed methods for estimating probabilities of female survival from age 25 to $25+x$ and male survival from age 35 to $35+x$ from proportions of respondents with surviving mother/father for five-year age groups. The calculation of time locations to which these estimates refer is detailed by Brass and Bamgboye (1981).

Zlotnik and Hill (1981) present a procedure that makes possible the estimation of adult mortality levels for a specific period even when vital rates have been changing rapidly. The procedure combines orphanhood data from two surveys conducted five or ten years apart to construct a hypothetical cohort that reflects the mortality experience during the intersurvey period. This procedure requires that the interval length between the surveys be a multiple of five years so that survivors of a five-year age group in the first survey can be identified as belonging to a standard five-year age group in the subsequent survey. When the interval between surveys or censuses is other than a multiple of five years, Timæus (2013a) suggests the calculation of synthetic cohort measures of orphanhood using a 'variable *r*' method involving the multiplication of the average proportions of respondents with living parents between the two inquiries by the exponential of the inter-survey growth rates in those proportions in the period accumulated from age 20.

A more flexible approach to estimating period mortality that addresses limitations of the earlier variants has been developed (Timæus 1990;1991b;2013a) that involves using data on parental survival at two dates to construct a cohort based at 20 years of age. The age range of respondents is confined to ages 15 and over to reduce the possibility of underestimating orphanhood due to the adoption effect. However, this approach will be affected by differential reporting of orphanhood by two surveys.

Chackiel and Orellana (1985) suggest collecting additional data on the year-of-death of the parent as part of the orphanhood question to address some of the limitations of the orphanhood method. Based on the additional information, proportions of parents alive five and ten years before the survey are estimated by adding to the reported numbers of respondents with surviving parents, those cases in which the parents were reported as having died during the five years or ten years before the survey. These proportions are then used to obtain recent period estimates of mortality. With the conventional two-survey orphanhood method, this would require the orphanhood question to have been asked in a previous survey which may not be the case in many countries. With the year-of-death data, this can be achieved with a single-survey. This approach has previously been applied to the year-of-death data collected in the National Demographic Survey from Honduras 1983-84 and the Vietnam Life History Survey (VLHS) conducted in 1991, the results from these studies were both internally consistent and indicated continuity with existing trends. In addition, more accurate time references to which estimates of mortality apply were obtained in the study conducted in Honduras, these time references were determined to provide a more plausible trend over time than those theoretically obtained using the method suggested by Brass and Bamgboye (1981).

Given the challenges associated with data derived from independent surveys, using the synthetic cohort measures of parental survival generated from supplementary information of years of parents' death provides information that is not affected by differential reporting of orphanhood and sampling errors between two surveys. Since the number of respondents whose parents were alive five and ten years before the survey can be determined from a single survey, this extension to the orphanhood method can be used as an alternative to the 'conventional' two-survey approach of estimating adult mortality. There have been limited number of attempts to apply this method presumably since these data are not routinely captured, and in contexts where data are available, applications of the method have been inconclusive, therefore, it is worth exploring whether the year-of-death data collected as part of the orphanhood question can be used to produce plausible estimates of adult male and female mortality in South Africa.

1.2 Objectives of the research

This focus of this study is to assess whether the year-of-death data collected as part of the orphanhood question can be used to produce estimates of adult mortality in South Africa at least as plausible as conventional data when the single and two-survey orphanhood method is applied to these data. This research will also explore whether it is possible to improve the estimates derived from the two-survey method by including responses of younger children or by using a intercensal survival approach to producing estimates from the data.

1.2.1 Research objectives

The research aims to achieve the following specific objectives:

- To identify all relevant research on the method suggested by Chackiel and Orellana (1985).
- To derive estimates of adult mortality using the traditional orphanhood approach for both single and multiple censuses or surveys.
- To derive estimates of adult mortality using year-of-death data after considering the unknown and implausible years of death data.
- To consider the alternative approach of estimating mortality between two censuses/pseudo-surveys directly from the survival probabilities.
- To identify lessons from the comparison of estimates from previous censuses/surveys with estimates from pseudo-surveys.

1.3 Significance of the research

This research will enable us to assess the usefulness of year-of-death data to estimate adult male and female mortality using the approach developed by Timæus (1990). This research will also consider if better estimates can be derived from these data assuming there is no adoption effect. This research will further explore if better estimates can be derived from the year-of-death data if the census-survival approach is used. The outcome of this study will allow us to form a conclusion about whether year of death data produces estimates of mortality that are plausible and are consistent with the RMS benchmark estimates. If the use of year-of-data is found to be suitable for estimating adult mortality, this approach can be used in countries and contexts where collection of accurate data directly from surveys is difficult. This approach will then represent a vital source of adult mortality estimates. If the use of year-of-death data does not produce reasonable estimates of adult mortality, it is important to document the challenges encountered in using these data i.e. how to deal with data with unknown year of death for the benefit of other studies wanting to use a similar approach.

1.4 Structure of the research

The dissertation consists of five chapter which are organized as follows; Chapter 1 will provide the background and significance of the research undertaking, research question and structure of research dissertation. Chapter 2 reviews literature that is relevant and related to this research. Chapter 3 will explore the sources and the type of data that will be used for the analysis. This is very important section of the thesis because the subsequent findings are reliant on the quality of the data used. This chapter will also describe the analysis and methods used to generate results obtained from applying specific methods to the data. Chapter 4 presents the findings from the analysis. Chapter 5 which is the final chapter will discuss how well the research question are answered, importance of the findings and possible areas of future research.

2 LITERATURE REVIEW

2.1 Introduction

This chapter gives an overview of the strengths and limitations of the approaches designed to generate estimates of adult mortality. In addition, it explores how the year-of-death data has been used to produce estimates of adult mortality. Further, the chapter looks at the level and trend of adult mortality in South Africa.

The chapter begins by reviewing the methods that have been proposed for estimating adult mortality. This starts with a detailed review of the orphanhood method and proposed variations aimed at improving the method. This is followed by a review of the work produced by Chackiel and Orellana (1985) on using the year-of-death data to produce estimates of adult mortality and a review of the intercensal survival approach of estimating mortality.

Thereafter, the chapter is focussed on reviewing the trend and level of adult mortality in South Africa. The chapter also covers the level of mortality estimated from using the conventional orphanhood approach[1] and that estimated using the year-of-death data.

2.2 Single-survey orphanhood approach

In contexts where registration of deaths is inadequate or lacking, information on the survival of parents can be used through the orphanhood approach to estimate indirectly levels of adult mortality. The orphanhood method is based on simple questions about the survival of parents of the respondent that can be asked in multi-purpose, single-round surveys, or censuses. Information on orphanhood is collected by asking the questions: “is your mother alive?” and “is your father alive?”. Neither the ages at death of the dead parents nor the year of death are asked. The length of time the parents were subjected to risk of death can be estimated from the age of the respondents (Timæus 1991c). This relationship can be modelled using different patterns of fertility, mortality and age composition to allow for the conversion of proportions with parents surviving into life table survivorship ratios (Hill, Choi and Timæus 2005).

The derivation of simple and robust methods for estimating adult mortality using information on the survival of parents is largely associated with Brass and Hill (1973)

---

[1] Conventional in this research refers to the orphanhood approach that does not make use of the year-of-death information.

who established an equation relating the female probability of survival from the age 25 to age $25+n$ to the proportions of respondents in two adjacent five-year age groups whose mother was still alive at the time of the interview. This equation has the form $\frac{l(25+N)}{l(25)} = W(N)\,{}_5S_{N-5} + (1-W(N))\,{}_5S_N$ where ${}_5S_{N-5}$ is the proportion aged between $N-5$ and N having a surviving mother, ${}_5S_N$ is the proportion the aged between N and $N+5$ having a surviving mother and $W(N)$ is a weighting factor used to make allowance for typical patterns of fertility and mortality (United Nations 1983).

To estimate adult male mortality a similar approach is used but allowing for the fact that fathers must have been alive, not at birth, but at the conception of the child. Therefore, three quarters of a year should be added to the period of exposure when estimating the period of exposure to risk. The various estimates of survivorship derived need to be converted into a common indicator of the level of mortality. This requires fitting a relationship (e.g. Brass logit relationship) using the life table, the logit life table system provides a solution (Brass and Hill 1973).

Hill and Trussell (1977) extend the approach previously proposed by Brass and Hill (1973) by using a regression equation for estimating female mortality from maternal orphanhood. Using 45 fertility schedules, derived from the Coale-Trussell model of fertility and simulations from 900 sets of proportions orphaned which are based on two age patterns and five levels of mortality, derived from Brass's General Standard (Coale and Trussell 1974), regression coefficients are estimated using linear regression.

Hill and Trussell (1977) concluded that survivorship could be estimated with sufficient accuracy by the following relationship $\frac{l(25+n)}{l(25)} = a(n) + b(n)M + c(n)\,{}_5S_{N-5}$, where M is the mean age of mothers(fathers) at the birth of their child and $a(n)$, $b(n)$ and $c(n)$ are regression coefficients. An additional source of bias noted when this variant of the orphanhood method is used is that survivorship estimates based on respondents below age 20 can be affected by mis-reporting because of the "adoption effect", where young orphaned children are often adopted by relatives who are then reported as biological parents (Hill and Trussell 1977).

Improved and modified methods of estimating adult mortality from information on the survival of parents have been suggested (Timæus 1991a;b;1992;Timæus and Nunn 1997) and these have contributed to the presentation of the method as is now applied by Timæus (2013b) and many other demographers. Building on research into the orphanhood method, regression models for both males and females are proposed that make use of coefficients drawn from data on parental survival in a set of simulated populations (Timæus 1991b). These coefficients are then used to estimate mortality from lifetime data on both paternal and maternal orphanhood.

In arriving at the equation that makes use of the regression coefficients, the proportion of individuals aged a with surviving mothers, $S(a)$, is calculated as the average of the probabilities of surviving among mothers who gave birth at each age y, weighting by the proportion of births that occur at y (Brass and Hill 1973;Timæus 1991b) and takes the form $S(a)=\frac{\int_s^w e^{-ry} l(y+a) f(y) dy}{\int_s^w e^{-ry} l(y) f(y) dy}$ where $l(y)$ is the probability of surviving from birth to age y, $f(y)$ is the probability of having a child at age y, w and s represent the child-bearing age range and r is rate of natural increase. Allowing for deaths between conception and birth of the respondents, the equivalent expression for living fathers is $S(a)=\frac{\int_s^w e^{-ry} l(y) f(y) \frac{l(y+a)}{l(y-0.75)} dy}{\int_s^w e^{-ry} l(y) f(y) dy}$.

The possibility of underestimating adult mortality due to the adoption effect is reduced significantly by estimating the proportion of a five-year age group with living mothers or fathers among those who had a living parent at exact age 20 (Timæus 1991b). This is estimated by the equation $\frac{{}_5S_x}{S(20)}=\frac{\int_x^{x+5} e^{-ra} l(a) S(a) da}{\int_x^{x+5} e^{-ra} l(a) S(20) da}$.

Using an approach based on regression produces more precise estimates, especially of men's mortality, as the intercept term provides added flexibility to more precisely model the association between survival and life table survivorship at varying levels of mortality (Timæus 2013b). Women's survivorship over the age range 25 to $25+x$ is more closely correlated with the proportions of mothers alive in age groups x to $x+5$ than any other survivorship ratio, thus maternal survivorship is estimated from the proportions surviving through the regression equation ${}_nP_{25}=a(n)+b(n)\bar{M}^f+c(n)\,{}_5S_{n-5}$ (Timæus 2013b).

Similarly, men's survivorship over the age range 35 to $35+x$ is more closely corelated to the proportions of fathers alive in age groups x to $x+10$ than any other survivorship ratio at most ages. Thus, the regression equation for paternal mortality is $_{n}P_{25} = a(n) + b(n)\bar{M}^{m} + c(n)\,{}_{5}S_{n-5} + d(n)\,{}_{5}S_{n}$ where ${}_{5}S_{n-5}$ is the proportion of respondents in the age group from exact age $n-5$ to exact age n with living fathers, $a(n)$, $b(n)$, $c(n)$ and $d(n)$ are regression coefficients estimated from simulated data created from relational logit life tables (Timæus 1992).

The series of survival ratios, $_{n}P_{b}$, obtained from different age groups of respondents and referring to different dates can be used to infer mortality trends by converting them into an index of mortality. The survivorship ratios are converted into the α parameter of a 1-parameter system of relational logit model life tables, and then into a common indicator of the level of mortality such as estimates of the probability of a 15-year old dying between exact ages 15 and 60, ${}_{45}q_{15}$ (Timæus 2013b). The parameter of the 1-parameter models is calculated from the estimates of $_{n}P_{b}$ as

$$\alpha = -\frac{1}{2}\ln\left(1 + \frac{\frac{{}_{n}P_{b}}{l^{s}_{b+n}} - \frac{1}{l^{s}_{b}}}{1 - {}_{n}P_{b}}\right)$$ where $b = 25$ for the estimates of women's survivorship and $b = 35$ for men's survivorship and l^{s}_{x} is survival of individuals to age x drawn from a standard life-table with a radix of 1. Each value of α corresponding to the measure of conditional survivorship made from data on the different age groups of respondents is converted into probability of dying ${}_{45}q_{15}$ using the equation

$$ {}_{45}q_{15} = 1 - \frac{1 + e^{2\left(\alpha + Y^{s}(15)\right)}}{1 + e^{2\left(\alpha + Y^{s}(60)\right)}}.$$

2.2.1 Limitations of the single-survey orphanhood approach

The application of the method reveals sources of error in the estimates of adult mortality derived (Timæus 1990). In many applications of the method, mortality estimates obtained from the younger respondents relative to those of older respondents were low and this is attributed to the "adoption effect" which refers to the situation where respondents whose parents have died are likely to have been adopted by another adult and may not even know that this person is not their biological parent. If true parents are substituted for foster parents and foster parents are subsequently reported on, mortality will be underestimated (Timæus and Graham 1989).

In addition, the method is not suitable as a source of timely estimates because the estimates obtained refer to dates before the orphanhood data were collected. The time location to which estimates of mortality apply are prone to distortions by any rapid changes in the rate of change of mortality because the time span in which the deaths occur is larger. Apart from this, the approach proposed by Brass and Hill (1973) is subject to a number of assumptions. First, it is assumed that age-specific fertility and mortality rates have remained constant. Second, it is assumed that the mortality of children is independent of that of their parents. Third, there should be no relation between mortality experience and the number of surviving children since those with no children have no weight while those with larger families are given greater weight. Fourth, the fertility model used may not be suitable and this is a more serious problem for males because in most cases the shape and characteristics of the male age-specific fertility patterns are less well known (Brass and Hill 1973). The violation of these assumptions may produce inaccurate estimates of adult mortality, so they are discussed in more detail below.

2.2.1.1 The adoption effect

A large body of literature (Blacker 1977;Hill and Trussell 1977;Timæus and Graham 1989) on the orphanhood method refer to a bias called the "adoption effect". This bias was first identified by Hill and Trussell (1977) and refers to the situation where respondents whose parents have died are likely to have been adopted by another adult and may not even know that this person is not their biological parent. If true parents are substituted by foster parents because of death of the parent and foster parents are subsequently reported on, mortality will be underestimated (Timæus and Graham 1989).

The magnitude of the bias due to the adoption effect is likely to be more serious in contexts like Africa, where the terms "father" and "mother" are often used loosely to refer to both a person's biological and foster parents (Blacker 1984). This bias is more apparent for young children, whose adopted parent may answer the question on their behalf or be assumed by the interviewer to be the real parent. As a result, orphanhood information from children below age 20, particularly the youngest children, may in practice be suspect (Hill and Trussell 1977).

2.2.1.2 Time location

Questions on orphanhood refer to deaths that occurred in the past, the resulting estimates represent some average of mortality over a period before a census or survey and the older the respondent, the longer the period (Chackiel and Orellana 1985). Procedures for estimating the time location to which mortality estimates obtained from retrospective questions refer are extensively described by Brass and Bamgboye (1981) who further indicate that estimates of adult mortality from orphanhood are prone to distortions due to rapid changes in the rate of change of mortality because the time span in which the deaths occur is larger.

Brass and Bamgboye (1981) suggest determining for each measure of survivorship, the time point where the period survivorship was equal to the cohort survivorship derived from the data in order to derive time location estimates which are robust to changes in the patterns of mortality. Time location is therefore estimated as: $t(N) = (N - 2.5)\dfrac{1 - u(N)}{2}$, where

$u(N) = \dfrac{1}{3}\ln {}_5S_{N-5} + Z\left(M + N - 2.5\right) + 0.0037\left(27 - M\right)$, Z is a standard function of age calculated from a standard life table, N is the mid-point of the age group of respondents and for paternal orphanhood, three quarters of year is added. M is the mean age of childbearing for mothers and for fathers, M is the mean age at child conception.

Timæus (1990), suggests a different approach of calculating time location which considers a standard life table with high mortality, as

$T = \dfrac{N}{2}\left(1 - \dfrac{1}{3}\ln\left({}_5S_{n-5}\right) + \dfrac{1}{3}\ln\left(\dfrac{80 - \overline{M}^{f} - N}{80 - \overline{M}^{f}}\right)\right)$. For men, an adjustment to the formula is made to account for the difference in the average exposure and age at initial exposure. For mothers, exposure starts at the mean age of childbearing while for men it starts at the mean age of conception which is nine months before the respondent was born or 0.75 of a year. Therefore, T for men is calculated as

$T = \dfrac{\left(N + 0.75\right)}{2}\left(1 - \dfrac{1}{3}\ln\left(\sqrt{{}_5S_{n-5} \cdot {}_5S_{n}}\right) + \dfrac{1}{3}\ln\left(\dfrac{80 - \overline{M}^{m} - N}{80 - \left(\overline{M}^{m} - 0.75\right)}\right)\right)$ (Timæus 2013b).

2.2.1.3 The effect of HIV on adult mortality estimates

One of the underlying assumptions of the orphanhood method is that the mortality experience of mothers should be independent of that of children. In populations affected by AIDS, that assumption is likely to be violated because HIV-positive women can transmit the virus to their children (and parents to one another). Thus, the children of women who are at risk of dying from HIV/AIDS are also likely to suffer higher mortality than the population in general. Because of the possibility of HIV transmission between parents, orphanhood-based estimates of mortality in populations affected by

HIV may also be biased downwards, more for women than men because there is a less direct relationship between infection of the father and infection of their children (Timæus and Nunn 1997).

Through the application of adjustment factors to orphanhood data, the proportion of respondents living with their mothers is adjusted to account for the bias of HIV on estimates of adult mortality. The proportions that would be reported if HIV-positive women had as many living children as other women $\left({}_5S_x \right)$ is obtained as follows: ${}_5S_x = \dfrac{1-hP}{1+\dfrac{1-F}{F}P}\,{}_5S_x^*$, where ${}_5S_x^*$ are proportions of respondents with living mothers, P is an estimate of the prevalence of HIV infection among women attending antenatal clinics, h is the vertical transmission rate of HIV from mother to child, and F is the relative level of fertility among HIV-positive women compared with HIV-negative women (Timæus 2013b).

Thus, for the reports of respondents born before PMTCT interventions and widespread provision of ARVs became widespread, a suitable adjustment might be ${}_5S_x = \dfrac{1-\dfrac{P}{3}}{1+\dfrac{P}{3}}\,{}_5S_x^*$. Adjusting the proportions of respondents with living fathers is more difficult because it is dependent on knowing the proportion of infected men with infected partners, information that is usually unknown. Timæus (2013b) suggests an adjustment of the form ${}_5S_x = (1-(1-(1-h)F)(1-w)P^*)\,{}_5S_x^*$ is used where w denotes the proportion of men with infected partners and P^* is an estimate of HIV prevalence among men in contrast to the antenatal based estimate for women.

Estimates of adult mortality in populations experiencing HIV are likely to be biased because the coefficients that are used to convert probabilities of survival into life table indices are based on populations experiencing different age patterns of mortality in adulthood compared to populations experiencing an AIDS epidemic. Timæus and Nunn (1997) propose a new set of regression coefficients to estimate life table survivorship in populations experiencing a severe HIV epidemic from the proportions of respondents with living mothers. Corresponding coefficients were not given to estimate men's mortality as little was understood about how age patterns of HIV incidence and fertility of men interrelate.

According to Timæus (2013b), these are reasonable assumptions to make about the impact of HIV on mortality estimates before the spread of large-scale treatment initiatives. If such initiatives have reduced the mortality if all HIV-positive adults to levels close to that of the general population, no adjustment would be required.

2.2.1.4 Out of date estimates of mortality

A further important limitation of the basic orphanhood approach is that it produces estimates of mortality referring to dates quite a few years before the date of the survey at which orphanhood data were collected. The deaths of parents can occur over the lifespan of the respondents from birth for mothers and from conception for fathers. Although more recent estimates could be obtained using the responses of the youngest respondents, these estimates are likely to be biased downward by the adoption effect. Using orphanhood information for all age groups provides estimates for 4 to 5 years before the survey and if the information on respondents aged under 20 is not used, the most recent estimates will refer to a time at least 8 years prior to date of the survey (Timæus and Graham 1989).

However, if there are two sets of orphanhood data available from successive surveys conducted in the same population, a hypothetical cohort can be derived that reflects the level of adult mortality in the intersurvey period. This hypothetical cohort based on the experience over the inter-survey period is analysed using the same methods as those used to calculate measures of conditional survivorship from data collected in a single enquiry (Timæus 1986;Timæus and Graham 1989;Zlotnik and Hill 1981). It is advantageous to use this variant of the orphanhood method because more up-to-date estimates of mortality that do not require the estimation of time references are derived compared to the single-survey approach. Another advantage, particularly if rates of mortality are changing in a non-linear fashion over time, is that it is easier to identify a standard table that is appropriate for the conversion of conditional survivorship ratios to ${}_{45}q_{15}$.

2.2.1.5 Choice of life table

The orphanhood method is further limited by using standard lifetables which assume age patterns of mortality that do not reflect the study population. The conversion of conditional survivorship measures to a single comparable index of mortality (such as ${}_{45}q_{15}$) used to assess a trend in the level of mortality over time is more severely affected by the use of an inappropriate model life table (Timæus and Graham 1989). This limitation is worsened in populations experiencing an AIDS epidemic as such populations have very different age patterns of mortality in adulthood from those modelled by standard life tables.

2.2.1.6 Age misstatement

The method is also affected by poor reporting of age by the respondents, since the proportions of persons whose parents are alive decline rapidly with age. Even a modest amount of overstatement of age can introduce a serious bias into the proportion orphaned. Evidence of the effect of age mis-reporting would be seen in low mortality levels implied by the proportions orphaned reported by older age groups of respondents, where age exaggeration on the part of the respondent becomes increasingly prevalent (Blacker 1984).

2.3 Two-survey orphanhood approach

Despite the wide application of the single-survey orphanhood approach, its usefulness has been questioned due to its limitations. In contexts where mortality rates may have changed rapidly, there is need for a method that can estimate mortality even under conditions of changing mortality rates. Zlotnik and Hill (1981) present a procedure that makes possible the estimation of adult mortality levels for the inter-survey period even when vital rates have been changing. It is also possible to obtain more up-to-date estimates which are usually more useful than estimates applicable to a time 8-10 years prior to the survey. The procedure combines orphanhood data from two surveys conducted five or ten years apart to construct a hypothetical cohort that reflects the mortality experience during the intersurvey period. Given information of the same population from two surveys, the proportion of respondents with parents' alive in each five-year age group at the first survey can be compared with proportion five or ten years later at the second survey. The resultant ratio of each pair of proportions, i.e. the proportion from the first survey and the proportion from second survey, will reflect the parents' mortality during the intersurvey period (Timæus 1986).

The procedure described by Zlotnik and Hill (1981) requires the interval between the surveys is a multiple of five years so that survivors of a five-year age group in the first inquiry can be identified as belonging to a standard five-year age group in the second inquiry. When the interval between surveys or censuses is other a multiple of five years, Timæus (2013a) suggests calculating synthetic cohort measures of orphanhood using a 'variable *r*' method which involves multiplying the average proportions of respondents with living parents between two inquiries by the exponential growth rates in those proportions.

Generalisations of the approach include using changes in orphanhood experienced by each age group to estimate stationary measures instead of changes experienced by each age cohort (Timæus 1986). This approach produces similar results as the approach suggested by Zlotnik and Hill but is also applicable when data are sourced from two inquiries that are not separated by either five or ten years exactly. Zlotnik and Hill (1981) propose chaining together a series of cohort changes to obtain a set of variables of interest that reflects the mortality rates in effect during the intersurvey period. However, doing so produces estimates of mortality that are biased downwards especially in older age groups. Timæus (1986) suggests using the ratios of the proportions of young adults with living mothers to resolve the downward bias observed in older age groups.

The estimation procedure suggested by Timæus (1986) is however inconvenient to apply because it involves working backwards from a life table and from an estimate of the mean age of childbearing to the proportions in each age group with living parents. To obtain estimates of mortality, this procedure applies multiplication factors or coefficients that are less reliable in comparison to those developed by Timæus (1990). In addition, like Zlotnik and Hill's approach, this variant involves extrapolation or interpolation between two sets of orphanhood data not exactly five or ten years apart.

A more flexible approach to estimating period mortality that addresses limitations of the earlier variants has been developed (Timæus 1990;1991b;2013a) that involves constructing a cohort based on data of respondents 20 years and above. The age range of respondents is confined to ages 20 and over to reduce the possibility of underestimating orphanhood due to the adoption effect. Based on the generalization of stable population to all populations the relationships between age structure, increase, and mortality, synthetic cohorts can be conveniently constructed. Stationary synthetic cohort measures of parental survival are then derived by adjusting the proportions identified with living parents using the age-specific growth rates in these proportions to eliminate the effect of trends in mortality (Timæus 2013a).

In order to apply this variant of the two-survey orphanhood method, the proportions of respondents with living mothers (or fathers) in an age group over the period between the two inquiries are calculated by applying the following formula:

$_5S_x(\overline{t}) = \sqrt{_5S_x(t) \times {_5S_x(t+h)}}$, where t indicates the time of the first inquiry, $t+h$ the second inquiry occurring h years later, and $\overline{t}$ indicates that the measure is applicable to the five-year age cohorts over an interval t to $t+h$.

Once these measures have been calculated, the average proportion of the parents of individuals aged exactly 20 that are alive during the intervening period can be estimated as $S(20,\overline{t}) = \sqrt{_5S_{15}(\overline{t}) \times {_5S_{20}(\overline{t})}}$.

The growth rates of the proportions of parents that are alive by age group of respondents between the first and second inquiry are calculated as $_5r_x(\overline{t}) = \dfrac{\ln(_5S_x(t+h)) - \ln(_5S_x(t))}{h}$, and the synthetic cohort proportions that have living parents among those who had a living parent at age 20 in the synthetic cohort can be calculated as

$$\frac{_5S_x(\tau)}{S(20,\overline{t})} = \frac{_5S_x(\overline{t})}{S(20,\overline{t})} \cdot e^{5\left(\sum_{a=20,5}^{x-5} {_5r_a(\overline{t})}\right) + 2.5\,{_5r_x(\overline{t})}}$$

where τ indicates adjusted synthetic cohort measures for time $\overline{t}$.

Using the regression coefficients generated by Timæus (1991b), which are estimated from simulated data based on relational logit life tables, the conditional life table survivorship ratio for women are calculated as $_{n-20}p_{45} = a(n) + b(n)\bar{M}^f + c(n)\,{_5S_{n-5}(\tau)} / S(20,\overline{t})$, and for men $_{n-20}p_{55} = a(n) + b(n)\bar{M}^m + c(n)\,{_5S_{n-5}(\tau)} / S(20,\overline{t}) + d(n)\,{_5S_n(\tau)} / S(20,\overline{t})$.

Estimates of $_np_b$ (conditional survivorship ratios), derived from different age groups of respondents refer to mortality experienced over the inter-survey period. As a result of errors in reporting of age, sampling, and other errors, the conditional survival ratios generated are inconsistent and fitting a 2-parameter logit model life table to the ratios removes these inconsistencies. This process involves first calculating the logits of $_np_b$ as $Y_x = \dfrac{1}{2}\ln\left(\dfrac{1 - {_np_b}}{_np_b}\right)$ and the corresponding logits of the standard life table $Y_x^s = \dfrac{1}{2}\ln\left(\dfrac{l_b - l_{b+n}}{l_{b+n}}\right)$, where l_b is life table function for number of survivors at exact age b and l_{b+n} represents the number of survivors at exact age $b+n$.

The parameters (α and β) are the slope and the intercept of the fitted line, which are derived by regressing values of Y_x on Y_x^s. Estimates ${}_n p_b$ from older ages are less likely to be affected by sampling errors than those from younger ages because they are derived from more parental deaths, but they can, however, indicate lower mortality than younger ages due to age exaggeration (Timæus 2013a). Following the calculation of α and β, smoothed estimates of conditional survivorship are calculated as

$${}_n\hat{p}_b = \frac{1}{1+\exp\left(2\left(\alpha+\beta Y_x^s\right)\right)}.$$

The estimates of conditional survivorship can be converted to a common index of mortality such as ${}_{45}q_{15}$ to allow for direct comparison with rates derived from other sources. This is achieved by fitting a 1-parameter model life table to each conditional survivorship ratio. The parameters of the 1-parameter model are derived from estimates of ${}_{n-20}p_b$ as

$$\alpha = -\frac{1}{2}\ln\left(1+\frac{\frac{{}_{n-20}p_b}{l_{b+n}^s}-\frac{1}{l_b^s}}{1-{}_{n-20}p_b}\right)$$, where $b=45$ for the estimates of women's survivorship and $b=55$ for men's survivorship and l_x^s is probability of survival from birth to age x from a standard life table. Each value of α corresponding to the measures of conditional survivorship made from data on the different age groups of respondents is converted into the probability of dying ${}_{45}q_{15}$ using the equation

$${}_{45}q_{15} = 1-\frac{1+e^{2\left(\alpha+Y^s(15)\right)}}{1+e^{2\left(\alpha+Y^s(60)\right)}}.$$

The synthetic cohort estimates measure adult mortality between the intersurvey period and the time location of the estimates is derived as the geometic average of the dates of the two inquiries.

2.3.1 Limitations of the two-census orphanhood approach

A major limitation of the synthetic cohort approach to the analysis of orphanhood is that it is affected by differential reporting of orphanhood between two surveys. Any change in error from one survey to another will be exaggerated in the synthetic data set (Zlotnik and Hill 1981). Crudely estimated, the sampling error for the difference between two proportions is about 41 per cent greater than the sampling error for either of the two proportions considered individually (Timæus and Graham 1989). The two-survey approach does provide a recent period estimate, but it would require the orphanhood question to have been asked in a previous survey which is not the case (at least for older children) in many countries. With the year-of-death data, this can be achieved with a single survey.

2.4 Data on year of death of parent

Chackiel and Orellana (1985) suggest collecting additionally data on year of death as part of the orphanhood question to address some of the limitations of the traditional orphanhood method and to improve estimates of adult mortality. Referring to the National Demographic Survey for Honduras 1983-84 (EDENH II), where all persons sampled were asked whether their mother/father was live, in cases where the mother/father was reported as dead, the year of death was asked. Using the information on year-of-death, Chackiel and Orellana (1985) suggest a procedure to (a) calculate estimated dates to which mortality estimates refer and (b) estimate mortality for periods 0-4 and 5-9 years before the survey.

Chackiel and Orellana (1985) present a more flexible approach to estimating time location that requires less data requirements to estimate time location than that initially suggested by Brass and Bamgboye (1981) and subsequent approaches suggested by Timæus (1990;2013b). According to Chackiel and Orellana (1985) the time location of the survival probability of each age group of respondents is calculated as the mean date on which the mother/father died for each age group of respondents assuming all deaths occurred in the middle of the year of death. Beyond describing how estimates of time location were calculated, Chackiel and Orellana (1985) do not provide further information. However, Kupamupindi (2010) referring to their study, suggests using year-of-death data to estimate time location as:

$M_k = \frac{\sum_{i=0}^{4}\sum_{j} d_{kij} Y_{kij}}{\sum_{i=0}^{4}\sum_{j} d_{kij}}$, where M_k is the mean year of death of parents of children *k* to *k*+4, where *k* = 5, 10, ..., 60 and d_{kij} represents the number of dead parents of children aged *k*+*i* who died in year Y_{kij}. The denominator can be simplified as the total number of deaths of parents in the child's age group *k* to *k+4*, where *k* = 5, 10, 15, …, 60.

Time location estimates derived by this approach are comparable to those obtained using the method suggested by Brass and Bamgboye (1981) for estimates derived from data from younger respondents but tend to stretch further back in time. However, for older respondents empirical calculations were deemed to be more consistent and not underestimated in comparison to estimates obtained using the traditional orphanhood approach (Chackiel and Orellana 1985).

The supplementary information on year of death can be also used to calculate the proportion of respondents whose parents were alive five or ten years before the survey. They showed through simple calculations how this information can then be used to construct a 'synthetic' survey five or ten years prior to the survey date, which can then be used in the two-survey method to estimate mortality for the intersurvey period (Timæus 2013a).

Chackiel and Orellana (1985) suggest computing the proportions of respondents with mother/father alive five and ten years before the survey by adding to the reported numbers of respondents with surviving mother those cases in which the mother/father was reported as having died during the five or ten years before the survey.

Beginning, for example, from the age group 15-19 at the survey date, the proportion with mother alive for the age group 10-14, five years before the survey, is obtained by dividing the total number of cases in which the mother/father was reported as having died during the five years before the survey by the total number of respondents in the age group 15-19 at the survey date. The proportion with mother/father alive for the group 5-9, ten years before the survey, is obtained by dividing the total number of cases in which the mother/father was reported as having died during the ten years before the survey by the total number of respondents in the age group 15-19 at the survey date.

In this way, the proportions at these two points (5 and 10 years before the survey) can be applied to the conventional orphanhood approach as if data separated by five years was available, giving adult mortality estimates for a much longer period than can usually be obtained (Chackiel and Orellana 1985).

This approach would not be subject to the problem of differential reporting of orphanhood by different respondents between two surveys (Chackiel and Orellana 1985). Further, with the year-of-death data, it is possible to calculate recent period estimates of mortality from a single survey.

2.5 Limitations of using year of death

The approach suggested by Chackiel and Orellana (1985) requires that respondents are able to recall accurately the year their parents died or the interval in which their parents died. Chackiel and Orellana (1985) does not mention the recall bias associated with using this year-of-death data as a limitation in the application of the approach. The findings obtained from applying the method to data from East Africa and South Africa point to the importance of knowledge of and accuracy of year of death. In both cases, the estimates were unreliable because a large proportion of respondents could not give parents' year of death. When previously applied to data from Burundi, the method did not perform very well even after apportioning the cases with unknown year of death of parents. This is because about 24 per cent of respondents with dead mothers and 27.5 per cent of those with dead fathers were unable to report how long ago their parents died (Makinson 1993). In a survey conducted in Vietnam in 1991, Hirschman, Preston and Loi (1995) demonstrate that this might not always be the case, respondents in this survey were able to respond to questions on the survival of relatives with a high degree of accuracy. In this study, only 18 per cent of the respondents were unable to recall the date of death of their mother/father.

2.6 Trends and levels of adult mortality in South Africa

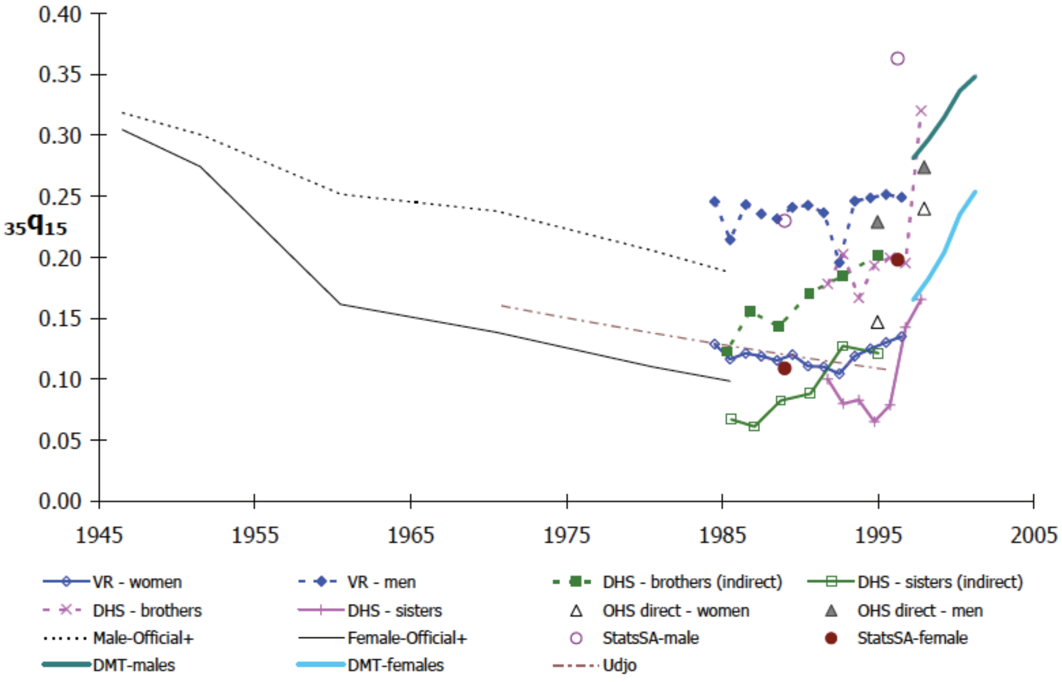
Before 1994, there was uncertainty about the level of mortality in South Africa. Under apartheid, life tables were at first only available for the white population and later for the coloured and Indian population groups, with no official life tables being produced for the African population. From 1979, deaths of Africans were included in the vital registration system, but only for those deaths that occurred in areas which the government at the time designated as part of the Republic of South Africa. Deaths occurring in areas such as Transkei, Bophuthatswana, Venda and Ciskei which were deemed outside the Republic of South Africa were supposed to be registered by these "countries" (Dorrington, Moultrie and Timæus 2004).

The last set of official life tables to be published covered the period 1984-86 and like the earlier South African Life Tables (SALTs) covered only the White, Indian, and Coloured population groups. However, Statistics South Africa (Stats SA) produced abridged life tables covering two periods 1985-94 and 1996 from data on children ever born/surviving and survival of parents. The tables for the first period were produced by population group and not by province. Those for the second period were produced by province rather and not population group (Dorrington, Moultrie and Timæus 2004).

Dorrington, Moultrie and Timæus (2004) analyse a range of data sources from South Africa covering the period 1945 to 2001. This analysis takes into consideration the apartheid and later democratic eras including the emergence of HIV/AIDS and therefore provides a comprehensive picture of adult mortality in South Africa.

Figure 2.1 depicts the probability of both 15-year-old females and males dying between ages 15 to 50, ${}_{35}q_{15}$. According to the "Official+" estimates, which are a weighted average of the official South Africa Life Tables for the White, Indian and Coloured populations and estimates for the African population from Sadie (1970), ${}_{35}q_{15}$ for both females and males decreased between 1945 to 1985 with estimates for females showing a sharper decline from 1950 to 1960. The estimates derived by Udjo indicate a decline between 1970 and 1985.

Figure 2.1 Trends in the probability of dying between ages 15 to 50, ${}_{35}q_{15}$, South Africa



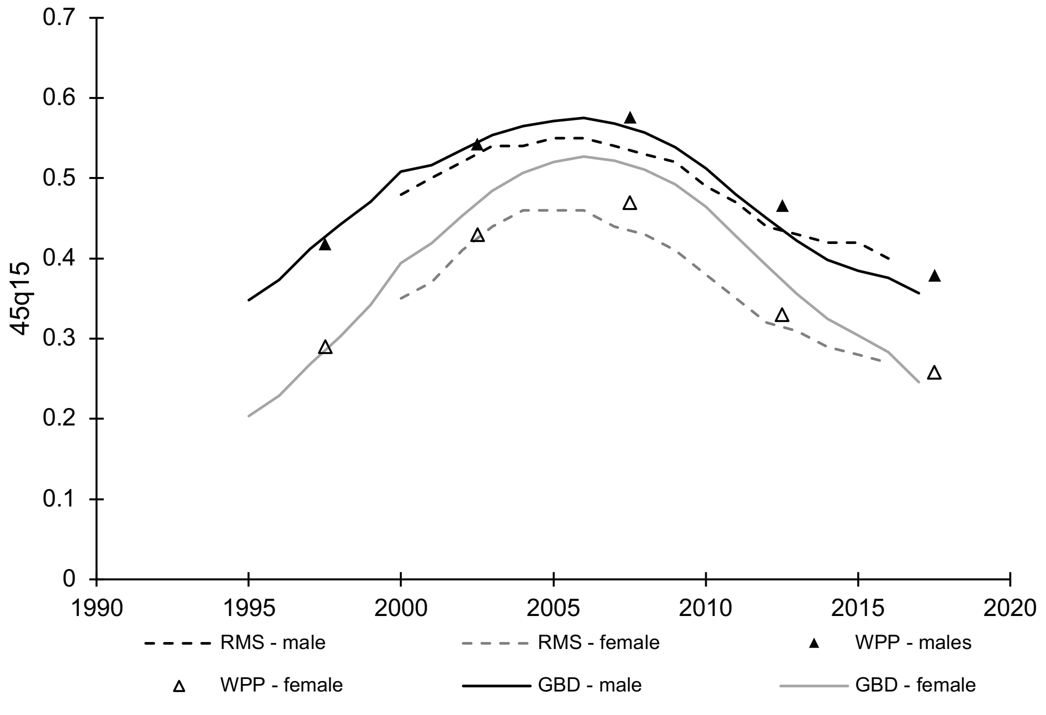
Source: Figure 6.2; page 65; Dorrington, Moultrie and Timæus (2004)
Notes: VR stands for vital registration deaths, DHS for Demographic Health Survey, OHS for October Household Survey and DMT for Dorrington, Moultrie and Timæus.

The estimates from the Demographic Health Survey (DHS) sibling data covering the period 1985-95, are not consistent with any of the estimates available for that period. Estimates from United Nations (2017) World Population Prospects supports the contention that ${}_{35}q_{15}$ decreased between 1970 and 1985.

The narrative on adult mortality in South Africa is incomplete without referring to the start of the impact of HIV/AIDS and this is depicted by Figure 2.1, which indicates an increase in female mortality from 1995.

Data from a surveillance site in Kwazulu-Natal which is one of the provinces to have been severely impacted by HIV/AIDS, show a major increase in adult mortality between 1995 and 2000. This lends credence to the uptick in the level of mortality observed after 1995. The South Africa Demographic and Health Survey (SADHS) 2016 estimates ${}_{35}q_{15}$ at 21 per cent for females and 23 per cent for males.

Figure 2.2 Trends in the probability of dying between ages 15 to 60, ${}_{45}q_{15}$, South Africa



Notes: RMS stands for Rapid Mortality Surveillance, WPP for World Population Prospects (2017 revision) (United Nations 2017), GBD for Global Burden of Disease 2017 (IHME 2018).

Figure 2.2 illustrates estimates of adult mortality presented in terms of the probability of a 15-year old dying between ages 15 and 60, ${}_{45}q_{15}$, as this measure has become a preferred summary indicator of the mortality of young and middle-aged adults (Timæus 2013b). In addition, more recent estimates of adult mortality for South Africa, (e.g. Dorrington, Bradshaw, Laubscher *et al.* (2018), make use of ${}_{45}q_{15}$ as a summary indicator). As can be seen from Figure 2.2, estimates of ${}_{45}q_{15}$ for both females and males indicate a steady rise from 1995 to 2005 and start to decrease thereafter. The rise in adult mortality observed from 1995 is largely attributed to the impact of HIV/AIDS.

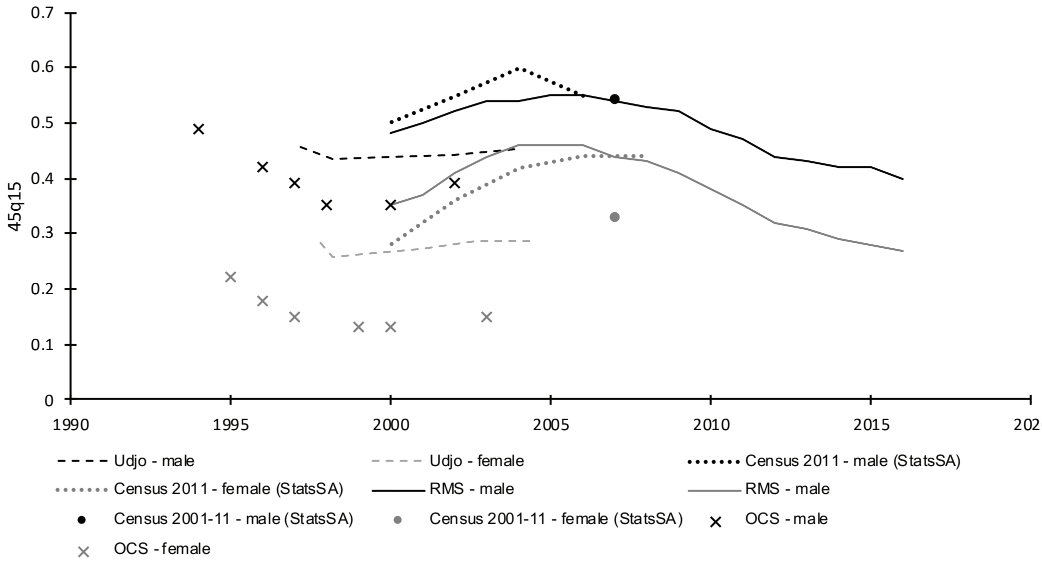
It is reasonable to attribute the increase in young and adult mortality to HIV/AIDS and this is corroborated by projections from models of the AIDS epidemic such as the ASSA model, developed by the Actuarial Society of South Africa (Dorrington, Bourne, Bradshaw *et al.* 2001). The estimates derived from the Global Burden of Disease (GBD) Study of 2017 show consistency with the other estimates with respect to overall trend, but they are generally inconsistent with respect to level especially between 2000 and 2013.

Recent estimates of adult mortality (Dorrington, Bradshaw, Laubscher *et al.* 2018) spanning the period 2000 to 2016 are fairly consistent with the 2017 series of WPP estimates especially for male mortality both in terms of level and trend.

2.7 Trends and levels of mortality produced using the orphanhood method

South Africa conducted its first post-Apartheid census in 1996 which included questions to assess adult mortality using the orphanhood method. The questions on orphanhood have become a permanent feature in subsequent censuses and community surveys. Using orphanhood data from the 2011 census, Udjo (2014) estimated values of alpha that indicate the level and trend of adult mortality. The values of ${}_{45}q_{15}$ in Figure 2.3 attributed to Udjo are derived from translating the alpha values using the INDEPTH network model life tables (IMLT) for Sub-Saharan Africa (SSA) by applying the formula ${}_{45}q_{15} = 1 - \dfrac{1+e^{2(\alpha+Y^s(15))}}{1+e^{2(\alpha+Y^s(60))}}$. IMLT were developed to account for HIV and other characteristics of SSA mortality. Unfortunately, Udjo (2014) did not publish his actual estimates of alpha, hence Figure 2.3 depicts ${}_{45}q_{15}$, obtained by translating values of alpha estimated from Figure 5 of his paper. These estimates are at odds with the other estimates.

Figure 2.3 Trends in the probability of dying between ages 15 to 60, $_{45}q_{15}$, South Africa (estimated using the orphanhood method)



Notes: RMS (Rapid Mortality Survey), OCS (Odimegwu, Chisumpa and Somefun), StatsSA (Statistics South Africa)

Estimates of $_{45}q_{15}$ derived by Statistics South Africa (2015a) from applying the orphanhood method to data from the 2011 census indicate an increase in female and male adult mortality from 2000 to 2005 and a gradual decline thereafter.

Statistics South Africa (2015a) use orphanhood data from the 2001 and 2011 censuses to estimate $_{45}q_{15}$ for the intercensal period (estimated to be around 2007). The estimates obtained for males are more consistent with the RMS estimates than those for females, which are implausibly low. It is unclear what standard was used to produce these estimates.

Odimegwu, Chisumpa and Somefun (2018) apply the orphanhood method to data from the 2001 census. They used the INDEPTH network model life tables for Sub-Saharan Africa to translate conditional probabilities of surviving into $_{45}q_{15}$. The results reflected in Figure 2.3 are inconsistent and implausible in comparison to RMS estimates and are lower than Udjo (2014) which were already too low relative to the other estimates. For instance, adult mortality for both males and females is indicated as falling between 1995 and 2000, meanwhile, the evidence presented in the previous section about adult mortality during this period is that it was rising since 1995 due to the impact of HIV/AIDS (Dorrington, Bourne, Bradshaw *et al.* 2001).

2.8 Estimating adult mortality using data on the year-of -death of parents

There is a dearth of research focussing on the use of information on the year-of-death of parents collected as part of the orphanhood question to estimate adult mortality. The only research discovered, is that documented by Chackiel and Orellana (1985), Timæus (1990), Makinson (1993), Hirschman, Preston and Loi (1995) and Kupamupindi (2010).

Chackiel and Orellana (1985) present results from the application of the traditional orphanhood method as well as the approach based on the supplementary information on the year-of-death based on data collected in Honduras. The results obtained are encouraging and show that the time location of estimates obtained using the year-of-death of parents are like those obtained using the traditional orphanhood approach. In addition, the time location estimates are more consistent over time than those obtained using the traditional orphanhood method. The results also show that adult mortality estimates derived from applying information about year of death of parents as noted by Chackiel and Orellana (1985) refer to more recent periods and are consistent with secondary sources.

Supplementary questions on the timing of parental deaths were also collected in the Burundi Demographic and Health Survey conducted in 1987. The results also show that time location estimates calculated using the traditional orphanhood method span a shorter period than those calculated using data on year of death of parents but they are biased towards the recent times (Makinson 1993). In addition, the conditional survival probability derived from the youngest age group seemed implausibly high mainly because of the adoption effect. Data on the year of death of parents was applied to estimate the proportions with mother/father alive five and ten years before the survey. The estimates of adult mortality derived were compared with those from widowhood and lifetime orphanhood data. A major problem encountered in applying this approach to the data from Burundi is that about 25 per cent of the respondents were unable to recall the date of death of their mother or father and this was more common for older than younger respondents, potentially biasing estimates derived from older respondents (Makinson 1993).

Hirschman, Preston and Loi (1995) investigate the possibility of gathering retrospective data on adult mortality by analysing information on birthdates of parents and siblings collected by the Vietnam Life History Survey (VLHS) conducted in 1991. If parents and siblings had died, further questions on the year of death or age at death were asked.

The results indicate that adjusted rates that were obtained after allocating deaths with missing data on age at death were closer to independent estimates of mortality based on kin survival and implied mortality levels in the West model life table system. Further, Hirschman, Preston and Loi (1995) indicate how this methodology could provide more robust estimates of Vietnamese mortality. The results, as demonstrated by Hirschman, Preston and Loi (1995), are encouraging and demonstrate the broader potential of using year-of-death data to describe mortality levels and trends especially in contexts where vital records and censuses are deficient.

The use of year-of-death data to determine levels and trends in adult mortality is significantly limited by the inability of respondents to recall the death dates of parents. However, in this study, Hirschman, Preston and Loi (1995) demonstrate that this might not always be the case. They noted that in the 1991 VLHS cultural traditions among the Vietnamese population such ceremonies to mark anniversaries of death may create a predisposition for "calendar literacy" that enables survey respondents to answer questions on the survival of relatives with a high degree of accuracy. In the 1991 VLHS 82 per cent of the respondents were able to recall the date of death of their mother/father. To determine whether this capacity exists in other countries and in other cultural settings to a sufficient degree to make the methodology applicable will require more data-collection experiments like the VLHS that use year-of-death data to estimate adult mortality.

There is only one documented application of using year of death data to estimate adult mortality from South African orphanhood data. Kupamupindi (2010) has applied the approach suggested by Chackiel and Orellana (1985) to NIDS 2008 data. Kupamupindi (2010) investigated whether time location estimates derived from supplementary questions on timing of parental deaths are more accurate than those derived using the (Brass and Bamgboye 1981) method. He also explores the consistency and level of adult mortality estimates produced using the two-survey orphanhood approach when compared to standard estimates from the Actuarial Society of South Africa (ASSA) 2003 model.

Kupamupindi (2010) notes that estimates of female mortality derived from the 2008 NIDS survey when plotted with time reference points calculated from year of death data are located in the more recent past and are inconsistent when compared to estimates from the ASSA 2003 model. Further, mortality estimates derived from using synthetic surveys going back 5 and 10 years before the 2018 NIDS survey are not consistent with the ASSA 2003 model, as estimates for both male and female mortality are higher. This is mainly attributed to the missing data on the timing of parental deaths and the subsequent methods used to deal with the missing data. For instance, of the respondents indicating not having a living father/mother, 36 per cent did not state when their father died while 32 per cent did not state when their mother died (Kupamupindi 2010).

Synthetic cohort estimates of adult mortality for both males and females derived from using year of death of parent's data applied to the NIDS 2008 are unrealistically high owing to how deaths with unknown year of death were apportioned (Kupamupindi 2010). Despite being able to produce estimates of adult mortality using supplementary information on year of death of parents, in the absence of quality retrospective reporting on year of orphanhood, the estimates produced did not match those from the traditional orphanhood approach and were not consistent with best estimates of adult mortality (Kupamupindi 2010).

The estimates of adult mortality derived using information on year of death of parents are distorted by the quality of reporting on year of orphanhood and this is well documented by Timæus (1990), Makinson (1993) and Kupamupindi (2010). Hirschman, Preston and Loi (1995) also document the limits of using year-of-data due to the inability of respondents to recall year of orphanhood. However, this was not the case with the study conducted in Vietnam. To determine whether this might be the case in other countries will require more collection of year-of-death data. In instances where the approach has produced inconsistent estimates, the data used has had a high degree of subject non-response. These studies for the most part, however, lack evidence of applying the approach to a wider context and to different sources of data.

Applying the approach to recent data from the South African Community Survey of 2016, provides an opportunity to supplement the existing research on the approach and to evaluate the reliability of estimates of adult mortality derived from it. The limitations associated with both the traditional one-census and two-census orphanhood approach, such as, time location, out of date estimates and estimates affected by differential reporting of orphanhood between surveys underscore the need to consider an approach that can potentially address these limitations.

3 DATA SOURCES AND METHODOLOGY

This chapter describes the data sources from which data on orphanhood are derived and the methods applied to these data to obtain estimates of adult mortality. These methods focus on the applying the variants of the single-survey and two-survey orphanhood approach of estimating adult mortality (Timæus 1990) to South African survey data. In addition, the chapter describes the extension of the two-census approach to include responses from younger respondents and the application of the intercensal survival method to the proportions of mothers/fathers alive at the two survey points as alternatives for producing estimates of adult mortality.

3.1 Data source and quality

Orphanhood data are extracted from the 2001 and 2011 censuses, the 2017 National Income Dynamics Study (NIDS) and the 2016 Community Survey (CS). The levels and trends of adult mortality in South Africa are derived from applying the variant of the orphanhood approach proposed by Timæus (1990) and proposed innovations proposed by Chackiel and Orellana (1985) that make use of year-of-death data to estimate adult mortality.

3.1.1 The 2001 and 2011 censuses

The 2001 and 2011 censuses are the second and third comprehensive census undertaken by Statistics South Africa. The reference date for both is midnight 9/10 October. They collected information on the demographic, social, economic, and housing characteristics of the population from all de jure household members and residents of institutions. The only source of the orphanhood data is in the form of publicly available unit record of the data from two censuses. These data are in the form of a 10 per cent sample that is representative of all individuals, households and institutions and includes a weight variable which adjusts the sample numbers for sample size and undercount (Statistics South Africa 2003;2015b). Questions on the survival of parents were asked in both censuses. The missing data in the 2001 census had valid responses imputed. With reference to the 2011 census, 98 per cent of the respondents responded to the question on survival of mothers and 96 per cent responded to the question on survival of fathers.

3.1.2 The 2017 National Income Dynamics Study

The 2017 National Income dynamics Study (NIDS) represent the fifth since the initial wave carried out in 2008. NIDS is a face-to-face rolling longitudinal survey of individuals and their households in South Africa. The period of the fieldwork was 6 February 2017 to 31 December 2017 (SALDRU 2018). The survey was designed to track the well-being of South Africans over time. Respondents below the age of 14 at the date of the survey are considered as children and those aged 15 and above at the date of the survey are considered as adults. The 2017 NIDS include questions about the survival of mothers/fathers. The respondents indicating their mother/father had died, were asked an additional question on the year when the mother/father had died.

Table 3.1 Per cent responses to questions on parental survival and year of death of parents, 2017 NIDS

	% response
Mother alive	99
Mother year of death – Don’t know/unspecified	17
Father alive	97
Father year of death – don’t know/unspecified	22

3.1.3 The 2016 Community Survey

The 2016 Community Survey (CS) is the second intercensal survey undertaken by Statistics South Africa (Stats SA) following the one conducted in 2007. The primary purpose of the Community Survey is to provide economic, social, and demographic data in between censuses. The survey covered 1,370,809 dwelling units. The reference date of the 2016 CS was midnight of 6/7 March (Statistics South Africa 2017). In addition to questions on the survival of parents, the 2016 CS asked for the year of death of parents. About 97 per cent of the respondents responded to the question on the survival of mothers and 5 per cent of the respondents whose mother had died did not specify the year of death of their mother. About 92 per cent of the respondents responded to the question on the survival of fathers and 8 per cent of the respondents whose father had died did not specify the year of death of their father.

As the 2016 CS is the focus of the research, the data on orphanhood and year of death as collected in the 2016 CS are explored in more detail. It is to be expected that responses of "Don't Know" and "Unspecified" will be much higher for respondents’ knowledge of paternal survival compared to that of maternal survival which is attributed to the absence of these fathers from the lives of their children (Dorrington, Moultrie and Timæus 2004). This is indeed the case as seen in Table 3.2.

Table 3.2 Per cent responses to parental survival and year of death of parents, 2016 Community Survey

	% response
Mother alive – Don't know	2.68
Mother Year of Death (MYOD) - Unspecified	4.83
Father alive – Don't know	8.27
Father Year of Death (FYOD) – Unspecified	8.07

It is worth pointing out that the respondents interviewed in the 2016 Community Survey are three times more likely not to know the survival status of their father relative to that of their mother.

Figure 3.1 Proportion of respondents not knowing father's/mother's year of death by age, Community Survey 2016

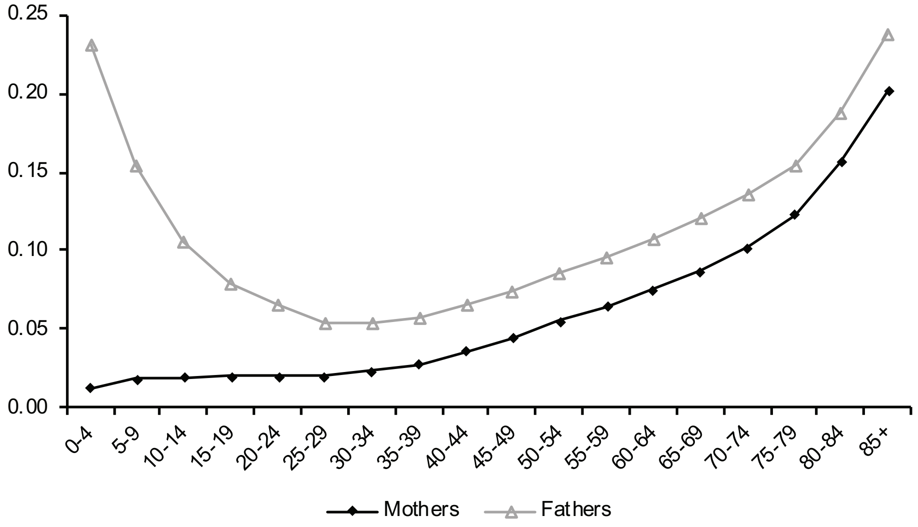


Figure 3.1 shows the proportion of respondents that do not know the year of death of their mother/father by the age of the respondent. There are more respondents not knowing the year when their fathers died, particularly at the younger ages, as this data is reported on behalf of children and mothers might refuse to say whether the father is alive, however, the proportion rapidly declines to minimum at ages 25-29 before steadily increasing thereafter. The proportion of respondents not knowing the year of death of mothers is relatively low and consistent at younger ages before steadily increasing from about age 30-34. These proportions are higher for older respondents because the deaths recorded by older respondents are likely to have occurred in the distant past and therefore, they are unlikely to recall the year when these deaths

occurred. It is clear from Figure 3.1 that knowledge of the survival of mothers was better at all ages than that of fathers.

Figure 3.2 Proportion of respondents whose mothers have died not knowing mother's year of death, Community Survey 2016

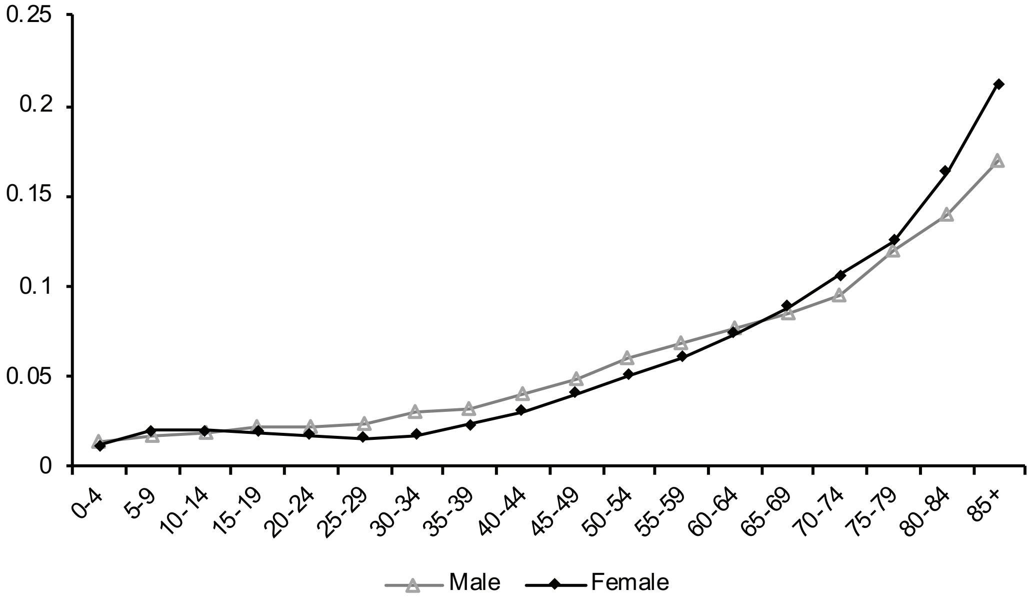
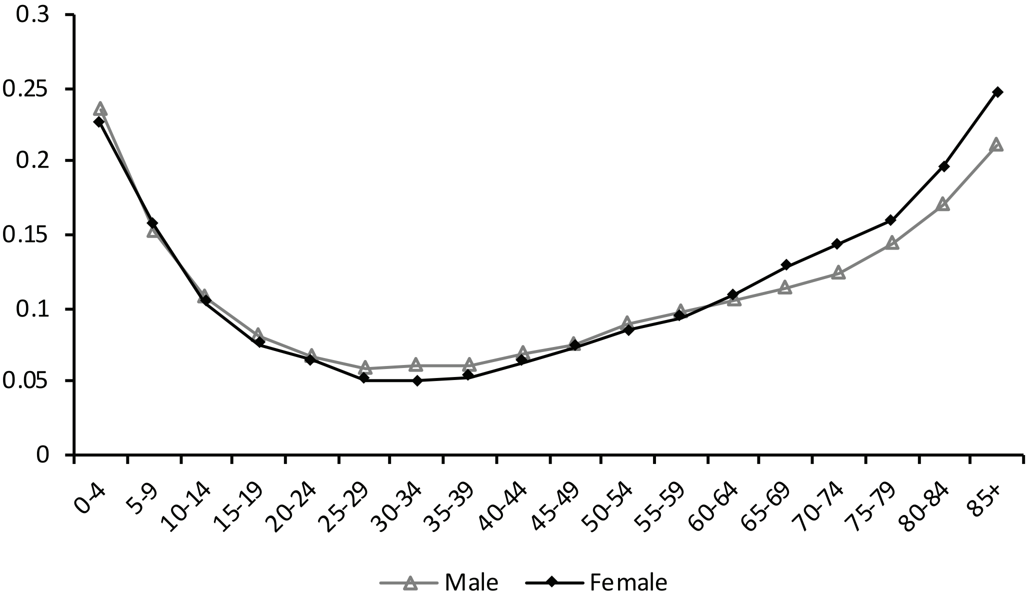


Figure 3.2 and 3.3 show the proportion of respondents that do not know the year of death of their mothers and fathers, respectively, by sex of respondent. The proportions are similar between female and male respondents and almost level for the younger ages. The proportions rise with age and are slightly higher for older female respondents than male respondents.

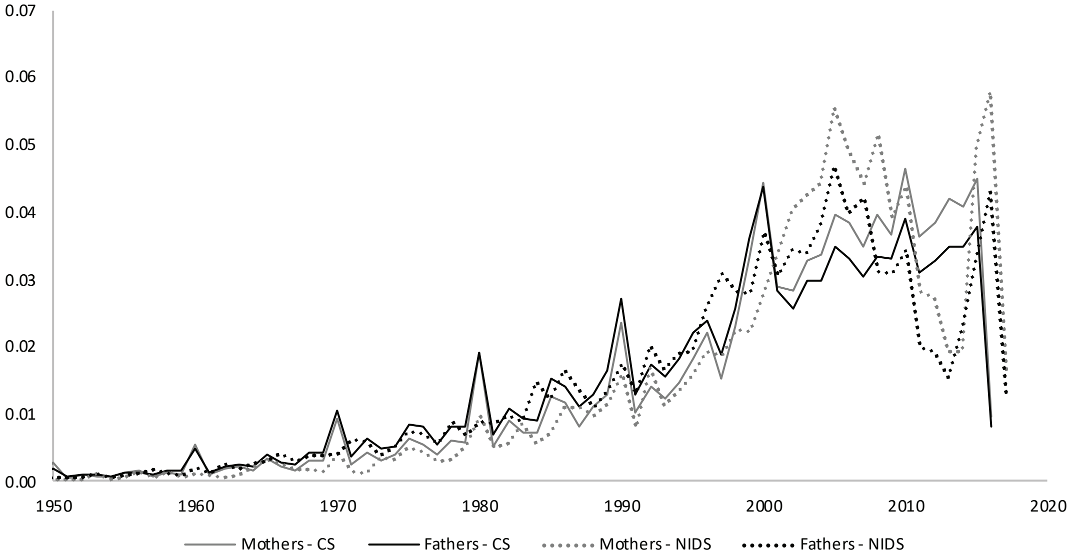
Figure 3.3 Proportion of respondents whose fathers have died not knowing father's year of death, Community Survey 2016



3.1.4 Proportion of deaths by year of death

Figure 3.4 below shows the distribution of reported deaths by year of death for both fathers and mothers from the 2016 Community Survey and the 2017 NIDS. For ease of presentation and interpretation, the deaths that occurred before 1950 have been excluded in Figure 3.4, but they have been included in the analysis. There is a steady increase in reported deaths over time, increasing rapidly from about 1990. These data also reveal heaping in the reporting of deaths on years of death ending in zero particularly for the Community Survey data. The reporting of more recent deaths in the 2017 NIDS suggests implausible patterns, indicating a decline after 2005 and increasing after 2013 before falling further back in time. This is likely to be attributable to the data rather than an actual depiction of mortality patterns. The proportion of reported deaths occurring further in the past are comparable between the two surveys, which is a bit surprising since one would expect the more recent data to be better. There are discernible differences between the two surveys for deaths reported closer to the dates of the surveys. In both surveys, more deaths for mothers than fathers are reported closer to the dates of the surveys. This is likely because the year-of-death of more of these deaths of fathers could have been reported as missing or unknown.

Figure 3.4 Proportion of deaths by year of death, Community Survey 2016 and NIDS 2017



3.2 Methodology

Orphanhood data from the 2001 and 2011 censuses, 2016 Community Survey and 2017 NIDS are entered into the orphanhood excel workbook (Timæus 2013a) obtained from the *Tools for Demographic Estimation* website[2]. The orphanhood excel workbook simplifies the application of the orphanhood method and is used to obtain period and cohort estimates of adult mortality. Where two sets of proportions of respondents whose parents were alive at the time of the survey are available, it is possible to estimate mortality applying the synthetic cohort method to orphanhood data in adulthood for the intervening period. In this research, the two-survey workbook is used to produce estimates of adult mortality for single as well as multiple surveys.

3.2.1 The single survey orphanhood method

The variant of the orphanhood method proposed by Timæus (2013b) in *Tools for Demographic Estimation* is applied to survival data from a single survey. To estimate mortality of adult women and men, the total respondents and respondents with mother/father alive by five-year age group of respondents aged 0-49 for adult female mortality and ages 0-44 for adult male mortality are extracted from the survey data. The respondents not knowing or not declaring the survival status of a parent are excluded from the calculations, thus effectively assuming the survival of these parents is the same as those that provided information on survival of parents.

Apart from the responses to questions about the survival of parents, the input data that are required for the orphanhood method include: the mean age of childbearing of the mothers and the mean age at child conception of the fathers.

To calculate the mean age of childbearing of women, one needs to estimate the number of births occurring in the twelve months before the reference date of the survey to women aged 15-49 years at the time of the interview.

Estimating the age of childbearing of fathers applicable to each survey is more challenging. The approach suggested by *Tools for Demographic Estimation* is to calculate the difference between the median ages of currently married men and women and add this to the age of childbearing for women. However, the data on currently married men and women is compromised by high proportions of unmarried cohabiting with many births occurring outside wedlock. In this research, the approach employed is to estimate the difference between the ages of mothers and fathers of registered births. Registered

---

[2] http://demographicestimation.iussp.org/content/indirect-estimation-orphanhood-multiple-inquiries

births by age of mother/father are obtained from Statistics South Africa (Stats SA)[3] for the years 2001, 2011, 2016 and 2017. Data on recorded births are not ideal because only a small proportion of recorded births contain information on the age of fathers. Therefore, the calculation of age of father at conception is confined to recorded births were information for both mothers and fathers is available.

The conditional survivorship ratios derived from the single-survey orphanhood method are converted into measures of ${}_{45}q_{15}$ using the one-paremeter system of the relational logit model life table. The choice of a model life table to be used should reflect the age pattern of mortality in the population being studied. However, it is difficult to find a life table that is able to reflect the mortality level in the study propulation and makes allowances, where suitable for the level of excess mortality owing to HIV/AIDS (Dorrington, Moultrie and Timæus 2004). This research makes use of Rapid Mortality Surveillance (RMS) standard life tables[4] derived from vital registration of deaths in South Africa. The tables used for each survey are those for the calendar year of the reference date of the survey. Ideally for the single-survey approach, a different life table should be used for each age-cohort and one needs to assume that mortality changed linearly over time. Applying a different life table to every age cohort is computationally inconvenient and the assumption that mortality changed linearly over time is not valid.

In addition, this research investigates the possibility of producing useful estimates by using life tables from World Population Prospects (WPP) (United Nations 2017), Institute for Health Metrics and Evaluations (IHME) Global Burden of Disease (GBD) (IHME 2018) and the default AIDS life tables contained in the orphanhood excel workbook.

To estimate the dates to which each of the adult mortality estimates refer, the reference date of the inquiry that asked about the survival of mothers and fathers was conducted must be known. If the questionnaire does not specify a reference date for the survey it is taken to be the average of the dates on which the interviews took place or the mid-point of the period of the fieldwork if the exact dates of the interview are not available. The reference date for the 2016 Community Survey is given midnight as 6/7 March 2016. The reference date used in this research for the 2016 CS is 7 March 2016. Since no reference date is specified for the 2017 NIDS, it is calculated as the mid-point

---

[3] http://nesstar.statssa.gov.za.

[4] Obtained from personal communication with Professor R E Dorrington.

of the period of the fieldwork (6 February 2017 – 31 December 2017) which is 20 July 2017.

3.2.2 The two-survey orphanhood method

To estimate adult mortality for the inter-survey period, this research uses the variant of the two-survey orphanhood method proposed by Timæus (2013a) that involves constructing a synthetic cohort based at 20 years of age, from data on parental survival at two dates. The age range of respondents is confined to ages 15 and over to reduce the possibility of underestimating orphanhood due to the adoption effect. The mean age of childbearing for women and men is calculated as described in section 3.2.1. The information is used with the synthetic cohort proportion of respondents that have living parents among those who had a living parent at age 20 and regression coefficients estimated from simulated data based on relational logit life tables (Timæus 1992) to calculate the conditional probabilities of survival for women and men.

To translate the conditional probabilities of survival into a common index of mortality, a suitable standard life table which reflects the mortality pattern in the inter-survey period should be selected. The RMS standard described in section 3.2.1 is used to translate the various conditional probabilities of survival into ${}_{45}q_{15}$. The exact date to which the period estimate of mortality refer is calculated as the geometric average of the two dates of the surveys.

3.2.3 Extension of the two-survey orphanhood method

To investigate the possibility of producing useful estimates by incorporating data from younger respondents, the age range of respondents is extended down to age 0, on the assumption that underestimating orphanhood due to the adoption effect is minimal (Kupamupindi 2010). Data on parental survival at two dates are used to calculate the synthetic cohort proportion of respondents that have living parents among those who had a living parent at age 5 (instead of age 20 as is used by the standard two-survey method). The synthetic cohort proportion of respondents, the mean age of childbearing for men or the mean age of conception for men and the regression coefficients for the single-survey method are used to calculate conditional probabilities of survival. The regression coefficients for the single-survey method are used as they are suitable for estimating survivorship from the proportion of adult respondents with living parents among those with living parents at age 5 (instead of age 20 as is used by the standard two-survey approach). The conditional probabilities are translated to ${}_{45}q_{15}$ using the standard two-survey orphanhood approach.

3.2.4 Year of death data

The year of death data are extracted from the 2016 Community Survey and the 2017 NIDS. The use of year of death data is affected by cases with unknown or missing year of death. Previous studies that have used year of death data with a significant proportion of cases with unknown year of death have not produced plausible estimates of ${}_{45}q_{15}$. In their paper, Chackiel and Orellana (1985) did not specify how they dealt with cases of unknown year of death, presumably because the proportion was insignificant. In the absence of any data to the contrary the approach used in this research is to assume the distribution of deaths with missing year of death is the same as those with reported year of death. Therefore, the cases with unspecified year of death in each five-year age group are apportioned proportionally based on the distribution of the dates of deaths for which information was available

After distributing the deaths with unknown year of death proportionally, these data on year of death are used to estimate the mean year of death of parents for each five-year age group of respondents. This is calculated as the mean date on which the mother/father died for each age group of the respondent and since the exact point in the year when the mother/father died is unknown, we assume that deaths occurred midway in the calendar year they are reported to have occurred.

Based on the additional information on year of death as suggested by Chackiel and Orellana (1985), the proportions of mothers/fathers alive for five and ten years before the respective survey are calculated by adding reported numbers of respondents with surviving mother/father to those cases in which the mother/father was reported as having died during the five years or ten years before the survey. In this way, one creates a synthetic survey which allows for the two-survey orphanhood method to be applied as described in section 3.2.2 as if orphanhood data was available from two surveys.

3.2.5 Intercensal survival approach

In another effort to improve the estimates from the data, use was made of the age-specific survival ratios. This involves calculating a survival ratio for each age group at the first census or survey and the equivalent older age groups at the second census or survey. This method is applied to the 2001 and 2011 censuses to estimate survival over a ten-year intercensal period and to the year-of-death data from the 2016 CS and 2017 NIDS, which are used to create synthetic surveys applicable five and ten years before the survey.

To estimate the indicator of adult mortality ($_{45}q_{15}$) over a five-year inter-survey period, we use the orphanhood reports of respondents starting from 0-4 years to estimate the survival ratios of mothers and fathers. Assuming the respondents aged 0-4 years at the first survey are concentrated at the mid-point of the age group, such that they are all aged exactly 2.5 years, we can calculate survival ratios of mothers starting from age 29.9[5] years and fathers starting from age 35.7[6] years. These respondents will be five years older at the second survey and so will their mothers and fathers.

Dividing the proportion of respondents aged 5-9 years with surviving parents at the second survey with those aged 0-4 years at the first survey gives an estimate of mother's survival ratios from age 29.9-34.9 years and from age 35.7-40.7 years for fathers. Similar quotients for subsequent age groups of respondents will estimate the rest of the parent's survival ratios over five-year intercensal period. Cumulative multiplication of these survival ratios gives a schedule of the probability of survival.

When surveys are exactly ten years apart, intercensal survival ratios for parents can be calculated by dividing the proportion of respondents with surviving parents aged 10-14 years at the second survey by the number aged 0-4 years at the first survey; and so on. Assuming, as in the case of surveys five years apart, that the respondents are concentrated at the mid-points of age groups, the intercensal survival ratios for the proportion of respondent's aged 0-4 ,10-14, etc., with surviving parents are estimates of survival ratios of mothers aged 29.9 years, 39.9 years and fathers aged 35.7 years, 45.7 years, etc. The intercensal survival ratios for the proportion of respondents aged groups 5-9, 15-1, etc with surviving parents are estimates of survival ratios of mothers aged 34.9 years, 44.9 years, etc and fathers aged 40.7 years, 50.7 years, etc.

This results in two series of probabilities of survival of mothers/fathers that are computed from the survival ratios. If we carried out subsequent calculations on these series, we would have two sets of estimates and we would assume we are working with ten-year age intervals rather than five-year intervals. Therefore, we need to merge the two series, thus giving probabilities of survival at five-year intervals.

Interpolation is then used to covert the nonstandard ages i.e., ages 29.9, 34.9, …to standard ages 30, 35, …for mothers and 35.7, 40.7, … to ages 35, 40, … for fathers. Linear interpolation is used to convert the probabilities of survival so that they are

---

[5] This is calculated as 2.5 years plus the mean age of mothers at childbearing.

[6] This is calculated as 2.5 years plus the mean age of fathers at the conception of the child.

applicable to standard ages of 30 years and onwards for mothers and 35 years and onwards for fathers.

The probabilities of survival of mothers/fathers from either the five or ten-year intercensal approach are then converted to a series of probabilities of dying (${}_nq_x$) in the age range 30-70 years for mothers and 35-70 years for fathers. In addition, a series of ${}_nq_x$ values in the age range 15-70 years are obtained from the RMS standard. The observed and RMS standard ${}_nq_x$ values in the age range 40-70 years for mothers and 45-70 years for fathers are used to calculate the intercept (α) and the slope (β) and through a simple linear regression model, these constants are applied to the RMS standard ${}_nq_x$values to obtain a series of fitted ${}_nq_x$ values ranging from 15-70 years. These fitted ${}_nq_x$ values are then used to calculate estimaes of ${}_{45}q_{15}$.

4 RESULTS

This chapter presents estimates of female and male adult mortality obtained after applying the orphanhood approach to the conventional form of the orphanhood data and to the year-of-death data. The first section presents results obtained after applying the one-survey orphanhood approach to survey data. The second section presents results obtained after applying the variant of the two-survey orphanhood approach proposed by Timæus (1990) to the year-of-death data and conventional orphanhood data. In addition, this section presents estimates of female and male adult mortality obtained from adapting the two-survey orphanhood approach proposed by Timæus (1990) to include the year-of-death data from younger respondents. The final section presents estimates obtained from applying the intercensal survival approach to produce alternative estimates of adult mortality between two surveys (this includes the synthetic surveys created from the year-of-death data). All the estimates of female and male adult mortality presented are compared with the RMS benchmark estimates to assess reasonableness.

4.1 One-survey estimates of adult male and female mortality

Estimates of ${}_{45}q_{15}$ obtained from applying the single-survey orphanhood method to survey data are presented and compared with the RMS benchmark estimates. Figure 4.1 below present estimates of adult male and female mortality. The estimates derived from the 2016 Community Survey and 2017 NIDS survey data are plotted against time reference points calculated using the Brass and Bamgboye method and those calculated using year of death data with unknown year of death data proportionally distributed to the years where the year of death is known. The estimates from the 2011 Census are plotted against reference points calculated using the Brass and Bamgboye method.

The reference points of the estimates of adult male and female mortality derived from the year-of-death data and those derived from the Brass and Bamgboye method have a similar trend and level, consistent with the findings of Chackiel and Orellana (1985). The reference points for adult male mortality which are derived from the year-of-death data are, however, shifted towards the date of the survey by about a year. This is likely the result of the assumption that unknown year of death are proportionally distributed to cases where year of death is known, and the year of death is more likely to be known for more recent deaths.

Figure 4.1 Estimates of adult male mortality derived using the single-survey orphanhood approach: 2011 census, 2016 Community Survey and 2017 NIDS compared to RMS benchmark estimates

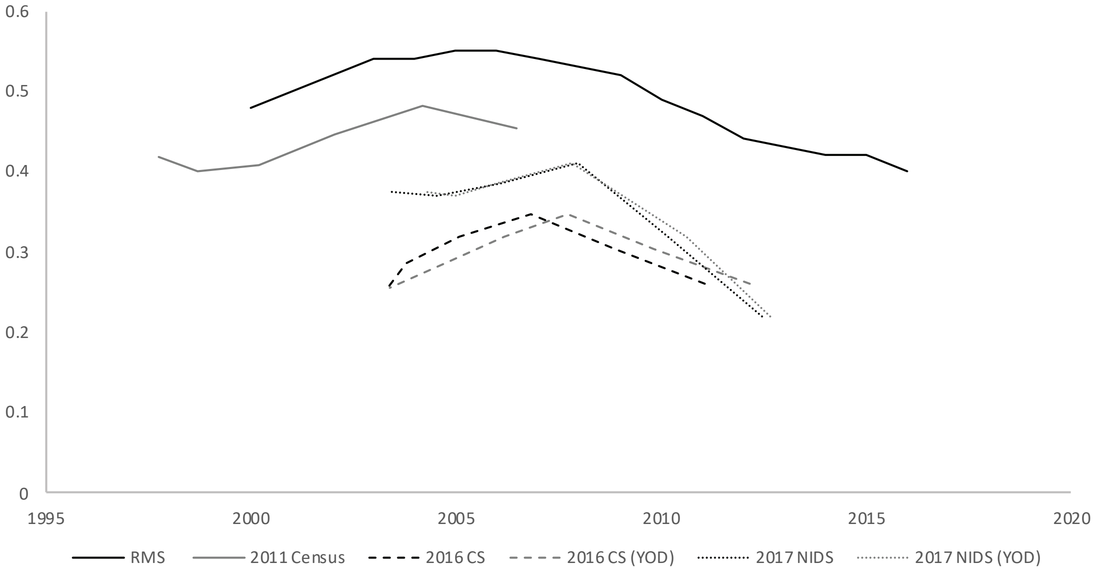
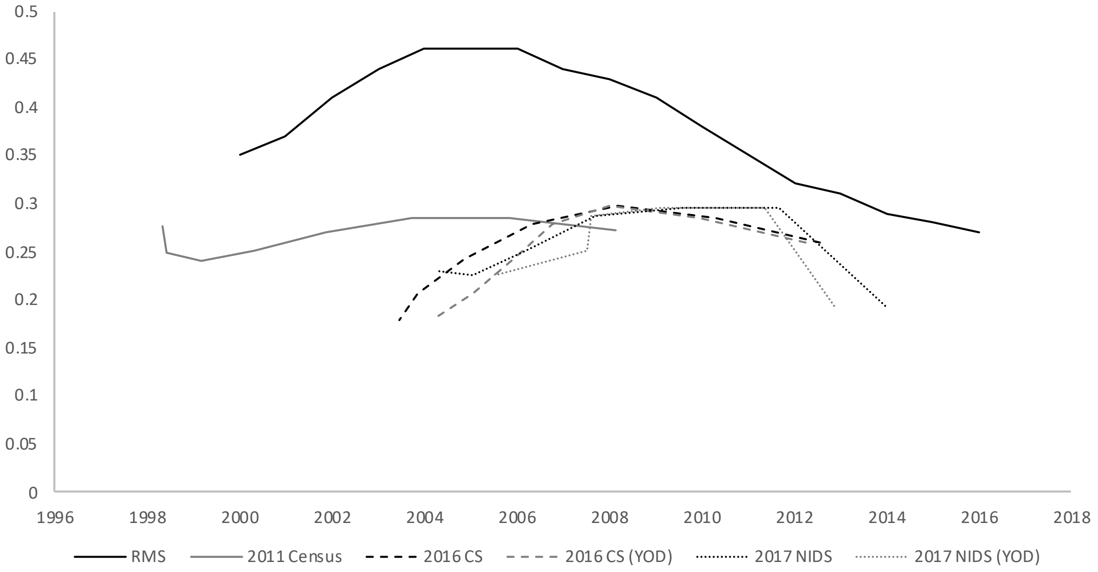


Figure 4.2 Estimates of adult female mortality derived using the single-survey orphanhood approach: 2011 census, 2016 Community Survey and 2017 NIDS compared to RMS benchmark estimates



Note: RMS – Rapid Mortality Surveillance, CS – Community Survey, YOD – Year of death, NIDS – National Income Dynamics Study

As illustrated in Figure 4.1 and 4.2 , ${}_{45}q_{15}$ of both females and males estimated from survey data is considerably lower than the RMS benchmark estimates. The underestimation of ${}_{45}q_{15}$ is greater using orphanhood data from the 2016 Community Survey and 2017 NIDS than the 2011 census. It is likely that orphanhood was better reported in the census compared to the surveys and the census has better coverage compared to the surveys. The estimates indicate a gradual increase, peak, and decline which is what is expected of mortality estimates during the period 2000-2015 that covers the impact of AIDS on mortality and later covers the impact of treatment. However, the peak occurs neither at the same point in time nor level of mortality. There is no discernible trend that is observed in the estimates of adult male mortality, while the trend is more apparent with estimates of adult female mortality. A steep fall in the estimates occurring closer to the survey date is observed for both estimates of adult male and female mortality. In the case of adult male mortality, the steep fall could be a result of underreporting of dead fathers in the younger age groups. The steep fall in adult female mortality reflected in Figure 4.2 is unexpected as data pertaining to mothers is usually better reported on but this does not appear to be the case.

The conversion of conditional survivorship ratios to a single comparable index of mortality is greatly affected by using an inappropriate mortality model to assess trends in mortality over time and this can explain the level of estimates illustrated in Figure 4.1 and 4.2 . Ideally, a different standard life table should be used for each age group, but this is computationally impractical.

4.1.1 Application of different standard life tables in the one-survey orphanhood approach

It should be noted that different standard/model life tables were tested in order to investigate the possibility of producing better estimates. First, standard tables based on 2017 series World Population Prospects (WPP)[7] South African age-specific patterns of mortality for the year of the survey were tested. Second, standard life tables for South Africa from the Global Burden of Disease study (GBD 2017)[8] corresponding to the year of the survey, were tested. Third, the default AIDS life table found in the orphanhood workbook from *Tools for Demographic Estimation*[9] was tested. The choice of different standards did not produce superior results to the Rapid Mortality Survey (RMS) standard shown in the figures. The RMS standard for the year of the survey was chosen and applied to the survey data.

---

[7] https://population.un.org/wpp/Download/Standard/Mortality/

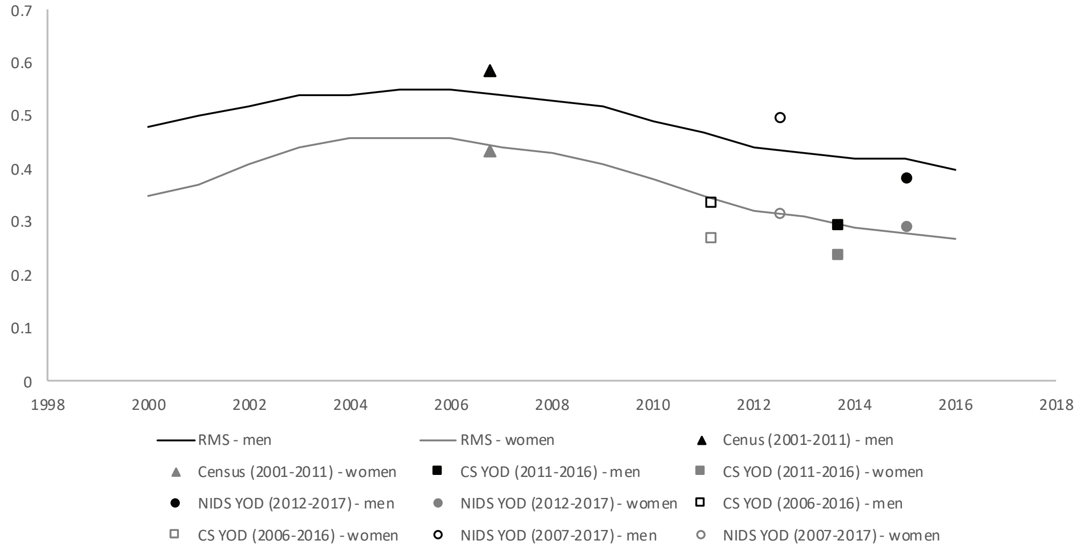
[8] http://ghdx.healthdata.org/record/ihme-data/gbd-2017-life-tables-1950-2017

[9] http://demographicestimation.iussp.org/content/indirect-estimation-orphanhood-multiple-inquiries

4.2 Two-survey estimates of adult male and female mortality

This section presents results of applying the variant of the two-survey method proposed by Timæus (1990) to survey data. Figure 4.3 shows the results of adult male and female mortality derived from applying the conventional approach of this variant to orphanhood data from 2001 and 2011 censuses. The year-of-death data (with cases of unknown year of death data proportionally distributed) from the 2016 Community Survey and the 2017 NIDS which are used to derive the proportions with mother or father alive five and ten years before the survey which are presented in Appendix A – D. The two-survey method is then applied to these data as though they are derived from surveys at different dates, and survival probabilities ($_{45}q_{15}$) are calculated for the intersurvey cohorts, these probabilities are presented in Figure 4.3.

Figure 4.3 Two-survey estimates of adult male and female mortality: comparison of estimates from 2001 and 2011 censuses, 2016 Community Survey and 2017 NIDS to the RMS benchmark estimates



Note: RMS – Rapid Mortality Surveillance, CS – Community Survey, YOD – Year of death, NIDS – National Income Dynamics Study

Intersurvey estimates are plotted at the mid-point of the period namely, 2006.76 for the census, 2013.67 for the Community Survey and 2015.05 for the 2017 NIDS. Figure 4.3 shows a much greater consistency and plausibility (to RMS benchmark estimates) of adult male and female mortality estimates from the application of the conventional orphanhood approach to the census than those from the 2016 CS and 2017 NIDS that use the year-of-death data. Although the estimates from the 2017 NIDS are comparatively better than those from the 2016 CS.

Survival probabilities are calculated for intersurvey cohorts for two periods (five and ten years before the survey) using the year-of-death data from the 2016 CS. The estimated levels of adult male and female mortality within each period (five and ten years before the survey) vary and mortality is lower in the most recent intersurvey period than the preceding one. The estimates of adult male and female mortality are underestimated for each intersurvey period and are implausibly lower than benchmark estimates. It is likely that these estimates are affected by errors in reporting the dates of death, the way data with missing year of death was treated (in this research, these data were distributed proportionally to those where year-of-death is known), sampling errors and possibly quality of the field work.

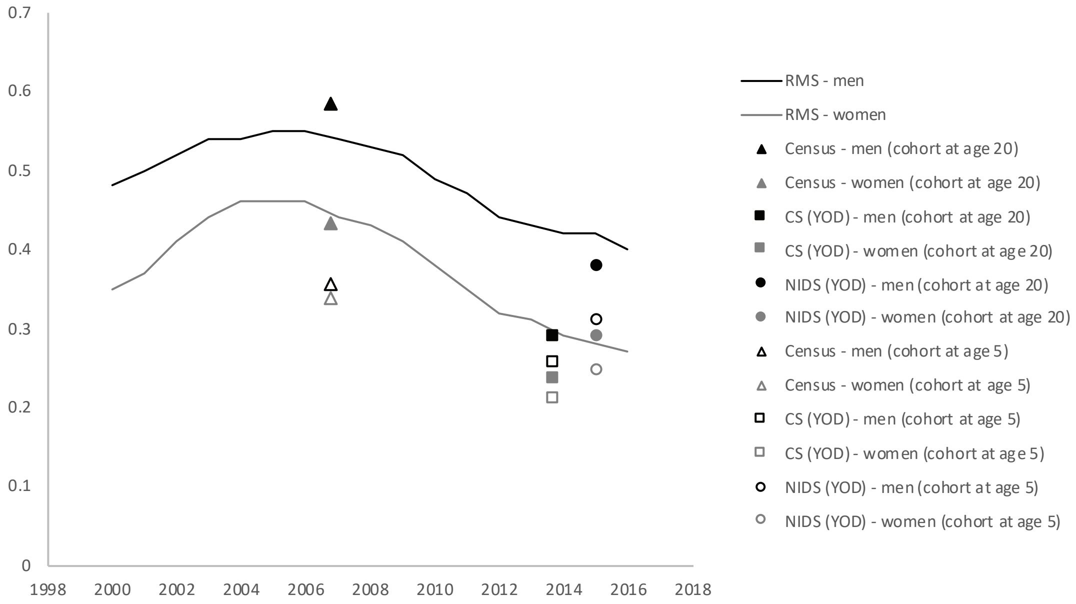
Using the year-of-death data from the 2017 NIDS, the estimated levels of adult male and female mortality also vary for each intersurvey period and adult male mortality is lower in the most recent period than the preceding one. Estimates of adult male mortality are underestimated in the most recent period and overestimated in the preceding period when compared to RMS benchmark estimates. Overall, these estimates are broadly consistent to RMS benchmark estimates given the high proportion of data with missing year of death. These results are also generally more useful than those using the one-survey data alone. They are also comparatively better than those from the 2016 CS.

4.2.1 Estimates of adult male and female mortality derived from orphanhood data from age 0 and over

The results of investigating the possibility of producing useful estimates of adult male and female mortality by incorporating data from younger respondents in the two-survey method are illustrated in Figure 4.4, the age range of respondents is extended down to age 0 assuming the possibility of underestimating orphanhood due to the adoption effect is minimal.

Estimates from the approach that incorporated data from younger respondents indicate lower mortality than those that use orphanhood data for respondents aged 20 years and over. This suggests that the marked decline in the cohort estimates is the result of the adoption effect which has less impact on estimates from respondents aged 20 years and over.

Figure 4.4 Two-survey estimates of adult male and female mortality derived using orphanhood data of respondents aged 0 and over and those aged 20 and over: comparison of estimates from 2001 and 2011 censuses, 2016 Community Survey and 2017 NIDS to the RMS benchmark estimates

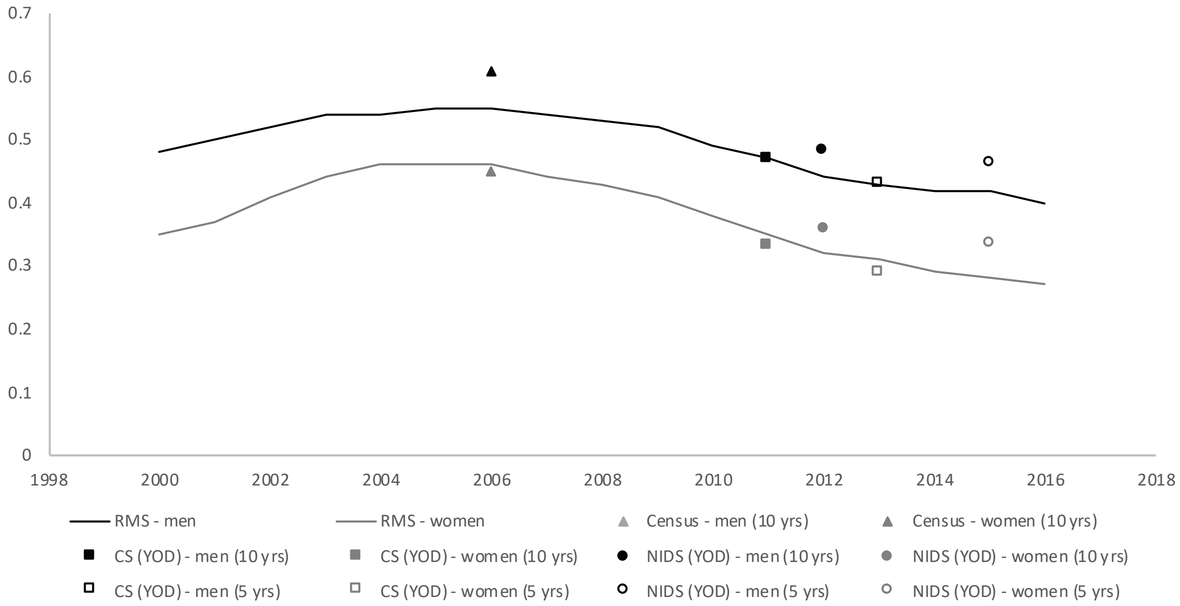


Note: RMS – Rapid Mortality Surveillance, CS – Community Survey, YOD – Year of death, NIDS – National Income Dynamics Study

4.3 The estimates of adult male and female mortality derived over a five and ten-year intersurvey period using the intercensal survival approach

The results of applying the intercensal survival approach to estimate adult male and female mortality are illustrated in Figure 4.5. This approach assumes mortality can be estimated from the survival ratios for each age cohort over an intersurvey period. The approach is used to investigate the possibility of producing better estimates of adult mortality because, under optimal conditions (no differences in the completeness of reporting of survival status of surveys, no migration, no age reporting errors), the intercensal survival methods may produce good results. The synthetic surveys generated from the year-of-death data from the 2016 Community Survey and the 2017 NIDS will not be subject to differential reporting between surveys, age reporting errors or migration and for synthetic cohorts. The impact of the adoption effect is more constrained in that it doesn’t include differential levels between the first and second surveys. Therefore, applying the intercensal survival approach to these data may produce better results of adult mortality.

Figure 4.5 Estimates of adult male and female mortality over a five and ten-year intercensal period derived using the intercensal survival approach: comparison of estimates from 2001 and 2011 censuses, 2016 Community Survey and 2017 NIDS to the RMS standard



Note: RMS – Rapid Mortality Surveillance, CS – Community Survey, YOD – Year of death, NIDS – National Income Dynamics Study

Estimates of adult male and female mortality are calculated for two intersurvey periods (five and ten years before the survey) using the year-of-death data from the 2016 CS. The estimated levels of adult male and female mortality within each period (five and ten years before the survey) are broadly consistent with the RMS benchmark estimates even though adult female mortality is slightly underestimated for the period closer to the survey than the preceding one. These estimates are more consistent than those from the census, where adult male mortality is overestimated and NIDS, where adult male and female mortality is overestimated. The differences in reporting of orphanhood between the two censuses with a higher than expected proportion of fathers not alive could explain the observed results, while for NIDS, estimates could be affected by quality of the data, which has a high proportion of missing information on year of death (17% for women and 22% for men). The comparison of these estimates suggests that mortality has declined in South Africa.

These estimates are calculated by fitting a straight line to the estimates of ${}_5q_x$ or ${}_{10}q_x$ as a function of a standard life table for the calendar year corresponding to the mid-point of the intersurvey period. The observed intercensal and standard estimates of ${}_nq_x$ in the age range 40-70 for women and 45-70 for men are chosen to estimate the parameters of the straight line, thus eliminating the first two data points for the younger ages and to reduce the adoption effect on the estimates obtained. The age range was shortened (dropping more data points for younger ages) when the fit produces implausible values.

4.3.1 Investigating the effect of using different life table standards in the intercensal survival approach

The above makes use of RMS standard life tables to translate survival probabilities to comparable estimates of adult male and female mortality ($_{45}q_{15}$). In order to assess the robustness of these estimates, various standards including the AIDS default life table found in the orphanhood workbook from *Tools for Demographic Estimation*[10] were tested to investigate the effect on estimates of adult male and female mortality.

The estimates presented below are the result of fitting a straight line to the estimates of $_{5}q_{x}$ or $_{10}q_{x}$ derived from the standard tables presented in Table 4.1.

Table 4.1 Estimates of adult male and female mortality for the five-year intercensal period derived using the intercensal survival approach: Comparison of estimates from applying RMS, WPP, GBD and AIDS default standard life tables

		RMS		**WPP**		**GBD**		**AIDS default**	
	Standard	*Survival approach*	*% diff*	*Survival approach*	*% diff*	*Survival approach*	*% diff*	*Survival approach*	*% diff*
Community Survey - 2016									
Female	0.307	0.291	-5%	0.320	4%	0.298	-3%	0.349	14%
Male	0.433	0.431	0%	0.467	8%	0.437	1%	0.515	19%
Census - 2001-2011									
Female	0.460	0.450	-2%	0.457	-1%	0.506	10%	0.490	7%
Male	0.549	0.608	11%	0.648	18%	0.611	11%	0.692	26%
NIDS - 2017									
Female	0.284	0.336	18%	0.371	31%	0.323	14%	0.40	41%
Male	0.415	0.463	12%	0.427	3%	0.534	29%	0.515	24%

Note: RMS – Rapid Mortality Surveillance, WPP – World Population Prospects, GBD – Global Burden of Disease.

---

[10] http://demographicestimation.iussp.org/content/indirect-estimation-orphanhood-multiple-inquiries

There is noticeable variability in the estimates of adult male and female mortality derived from applying different life tables and as illustrated by Table 4.1. The RMS, WPP and GBD standard tables produce similar estimates of adult male and female mortality to the benchmark standard when applied to data from the community survey. The AIDS default standard produces slightly worse results compared to the other standards and overestimates both adult male and female mortality. All the tested standards produce similar estimates of adult female mortality to those from the benchmark standard when applied to census data. A lot more variability is noticeable in the results of adult male mortality with the RMS and GBD standard tables producing results that slightly overestimate adult male mortality and WPP and AIDS default standards producing results that overestimate adult male mortality by a greater percentage. When applied to NIDS data, WPP and AIDS default standards overestimate adult female mortality by a greater percentage than those from RMS and GBD standard tables. GBD and AIDS default standards overestimate adult male mortality by a greater percentage than the RMS standard.

The WPP standard produces similar results to the benchmark estimates. Generally, the default AIDS produces poor results, and this could suggest that the default AIDS life table is considerably different from the mortality experience of the study population. The AIDS default standard table does produce better results if all data points are used in the fit instead of dropping the first two points. This might be because including these points distorts the fit of the straight line towards producing an estimate of ${}_{45}q_{15}$ closer to the benchmark, even though the fit to the data is not good.

5 DISCUSSION

5.1 Introduction

This chapter reflects on the usefulness of using the year-of-death data to estimate adult mortality. It discusses the different approaches applied to these data on the year-of-death and the reasonableness of the estimates of adult mortality that are produced. The chapter concludes by discussing the limitations of the study and identifying possible areas of further inquiry.

5.2 Discussion of results

This research set out to assess whether the year-of-death data collected as part of the orphanhood question can be used to produce plausible estimates of adult mortality in South Africa. The year-of-death data was used to derive proportions of respondents with parents alive five and ten years before the survey. The data with unknown year-of-death was proportionally distributed to the years where the year-of-death is known. If there is a high percentage of cases with unknown year-of-death, the estimates obtained are likely to be biased. An upward bias in the estimates might be observed if the undated deaths are assumed to have occurred longer ago on average than the deaths of which the date of death is known, by proportionally distributing the deaths we are assuming the deaths happened in the same period as the deaths of which the date of death is known.

The single-survey and two-survey orphanhood approaches are the main techniques used to assess the usefulness of the year-of-death data to estimate adult mortality in this research. However, innovations that use response from respondents aged 0 and above and those that use the intercensal survival approach are applied to the year-of-death data to obtain estimates of adult mortality.

5.2.1 Single-survey estimates of adult mortality

To determine whether the year-of-death data can be used to estimate adult male and female mortality, we must look at the mortality estimates and the time reference points that are produced when these data are used in the single-survey orphanhood approach after having proportionally distributed deaths with unknown year-of-death to where the year-of-death is known. The mortality estimates produced from using the year-of-death data are much lower than the RMS benchmark estimates. In addition, these estimates are lower (especially those for adult male mortality) than those from the standard data which are equally poor when compared to the RMS benchmark estimates. The nature of collecting the year-of-death data requires that respondents accurately recall when the deaths occurred, this means that estimates derived from these data are likely biased by recall errors. If there is a high percentage of cases with unknown year-of-death in these data, how such cases are apportioned could also bias the estimates. In such instances, the year-of-death data are not suitable to be used to estimate adult mortality.

However, the year-of-death data cannot be dismissed simply because using the conventional from of the data does not produce better estimates of adult male and female mortality. These results are inconclusive as to whether the year-of-death data can or cannot be used in the single-survey approach to estimate mortality because the conversion of conditional survivorship ratios to a single comparable index of mortality is greatly affected by using an inappropriate mortality model to assess trends in mortality over time and this could also explain the poor estimates obtained when the year-of-death are used with the single-survey approach. Ideally, a different standard life table should be used for each age cohort, but this is computationally impractical and it is difficult to find a life table that makes allowances, where applicable for the level of excess mortality due to HIV/AIDS and has the right shape (Dorrington, Moultrie and Timæus 2004).

The year-of-death data are used to calculate time reference points to which estimates of adult male and female mortality refer. The time reference points for adult male mortality calculated from the year-of-death data are comparable to those calculated using the Brass and Bamgboye method, however, there is an indication of a shift towards the more recent past especially for paternal time allocation estimates calculated using the year-of-death data. The shift is probably because the deaths are being reported as having occurred more recently than they did. The distribution of cases with missing year-of-death data in proportion to those cases where year-of-death is known may also explain the shift.

These findings agree with earlier research findings from Chackiel and Orellana (1985) that found that estimates of time location calculated using the year-of-death data gave similar results to those calculated using the conventional method especially for younger ages, differences were however noted for older ages.

In this research, the estimates of time location calculated from the year-of-death data and those calculated from the conventional approach suggest the year-of-death data performs just as well as the conventional from of the data[11] in estimating time references to which estimates of mortality apply.

5.2.2 Two-survey estimates of adult mortality

It is clear from section 5.2.1 that none of the data used in the single-survey orphanhood approach was able to produce plausible estimates of adult male and female mortality. The year-of-death data are used in the two-survey approach to investigate the possibility of producing better estimates of adult male and female mortality after distributing deaths with unknown year-of-death proportionally to data where the year-of-death is known. In this research, the year-of-death data are available from two surveys; the 2016 Community Survey and 2017 NIDS.

The estimates obtained from using the year-of-death data from the two sources are surprisingly inconsistent. While variation in the estimates is expected due to the differences in data quality between the two sources, the year-of-death data from the 2016 Community Survey produces implausibly light mortality, while 2017 NIDS year-of-death data produces estimates of adult male and female mortality that are surprisingly consistent with the benchmark despite having a higher percentage of reported deaths with unknown year-of-death. It would be expected that proportionally apportioning deaths with unknown year-of-death would have a more noticeable effect on the overall estimates from these data than those from the 2016 Community Survey, but it is not the case in this research. But, it is possible that these data were subject to fewer errors in the reporting of dates of orphanhood, data collection errors and sampling errors, since collection of these data may have improved with repeated surveys. In addition, the unrealistically low estimates derived from the year-of-death data from the 2016 Community Survey cannot be attributed to the adoption effect because these results are derived from applying the two-survey approach proposed by Timæus (1990) which considers orphanhood data for older respondents (15 years and older).

These results pose a challenge in answering the question on whether the year-of-death data can be used to produce plausible estimates of mortality, because the data considered in this research produces conflicting results and thus arriving at a definite conclusion about the year-of-death data in this regard is difficult.

---

[11] Conventional form of the data refers to data that excludes additional information on the year-of-death.

5.2.2.1 Estimates of adult male and female mortality derived from orphanhood data from age 0 and over

Attempts to determine whether the year-of-death data can be used to produce plausible estimates of mortality when used in the two-survey orphanhood approach have so far been inconclusive. The two-survey orphanhood approach proposed by Timæus (1990) is extended to consider the year-of-death data from younger respondents (from age 0 and over rather than from age 15 and above) so as to determine if more plausible estimates of mortality can be produced from these data assuming there is no adoption effect. Earlier research (Kupamupindi 2010), suggested estimates derived from using the year of death data underestimated mortality but evidence did not point to the adoption effect.

The results derived from the year-of-death data and those from the conventional form of the orphanhood data reveal unrealistically low levels of mortality suggesting that the adoption effect is a source of bias. These results suggest that adapting the approach suggested by Timæus (1990) to include data from younger respondents helps us answer the research question as to whether the year-of-death data can produce plausible estimates of adult male and female mortality. These findings further suggest this approach is not suitable in estimating adult mortality because both the year-of-death data and conventional form of the orphanhood data when used in this approach produce unrealistically low estimates of adult male and female mortality.

5.2.3 The estimates of adult male and female mortality derived over a five and ten-year intersurvey period using the intercensal survival approach

The year-of-death data are used to create synthetic surveys occurring five and ten years before the survey date. These surveys are used to measure survival over an inter-survey period using the intercensal survival approach. The estimates of adult male and female mortality for the five and ten-year intersurvey period produced using these synthetic surveys are largely consistent with the RMS benchmark estimates of adult mortality. The results suggest that the year-of-death data may possibly be used to produce plausible estimates of adult male and female mortality for the five and ten-year period before the survey when the intercensal survival approach is applied to these data.

The unchanging characteristics of the synthetic surveys created from the year-of-death data in comparison to data obtained from two separate sources, i.e., two subsequent censuses or surveys may explain why the intercensal survival approach produces results that are consistent with the RMS benchmark estimates of adult male and female mortality. In addition, the results presented in this research show that estimates of mortality derived from synthetic surveys created from the year-of-death data are not affected by age reporting errors and differences in completeness of reporting of survival status, which is a more fundamental drawback of calculating rates for an intercensal period from two sources and frequently introduces large errors.

Therefore, the differences in calculated survival ratios are minimised and resulting estimates are consistent with benchmark estimates.

5.2.3.1 Choice of life table

It could be that year-of-death data produces good results because of the choice of life table used in the conversion of the conditional survivorship ratios to a single index of mortality which is compared to that from the standard. Therefore, the RMS, WPP, GBD and the AIDS default life tables are applied to the conditional survivorship ratios produced from the year-of-death data to ascertain how estimates of mortality produced from these life tables compare with the standard estimates. The estimates of mortality produced by applying the RMS, WPP and GBD life tables are sufficiently close (0-10%) to the standard. This may be because the RMS, WPP and GBD tables are similar in shape by age and probably have the general shape and timing of the AIDS epidemic. Generally, the default AIDS model life table produces poor results, and this is probably because it is considerably different from the mortality experience of the study population. These results suggest that year-of-death data are usable in the derivation of plausible estimates of mortality and that the results are not entirely dependent on the choice of life table.

5.3 Limitations of the study

One of the major limitations encountered in this study was how to deal with data with missing year-of-death. In this research, it was assumed that the distribution of these data with unknown year-of-death is proportional to those data were the year-of-death was known and the results show that doing so might lead to a bias in the estimates of mortality. This research relies on using the year-of-death data which is subject to recall errors that are more apparent for older respondents that must recall timing of deaths that occurred further in the past. It is not certain whether the year-of-death data used in estimating adult mortality are accurate since this was not determined in this study.

Another important limitation of this study is that it does not make any adjustments for HIV-related biases when estimating adult mortality. Adult mortality estimates made using the orphanhood method in populations experiencing a generalized

HIV epidemic are vulnerable to severe bias. According to Timæus and Nunn (1997), if estimates are available of the pertinent characteristics of the HIV epidemic in a population at the time that the respondents were born, it is possible to limit the biases by adjusting the reported proportions of mothers that are alive downward by assuming that most individuals who are already infected with HIV when their children are born will die within a few years. This was a reasonable assumption to make before the recent advent of large-scale treatment initiatives. This research makes use of orphanhood data collected after the advent of large-scale treatment initiatives and therefore assumes that the mortality of all HIV-positive adults has been reduced to levels close to that of the general population and therefore no adjustment would be required.

The intercensal survival approach does not correct for HIV-related bias as suggested by Timæus and Nunn (1997). If we use year-of-death data to create synthetic surveys applicable further in the past (before large-scale treatment initiatives) and apply them to the intercensal survival approach, the proportions of surviving parents and the resulting estimates of adult mortality are prone to HIV-related biases. Therefore, the intercensal approach as proposed in this research has limited application in populations experiencing a generalised HIV epidemic and where access to wide scale treatment initiatives is limited.

The collection of year-of-death data is a relatively recent undertaking in South Africa. In the case of censuses and Community Surveys, the 2016 CS was the first time these data were collected. The quality of data collected may have been compromised by a lack of enumerator training on how to elicit responses to questions on the timing of deaths.

Finally, in their research, Page and Wunsch (1976) mention that the calculation of the mean age of fathers at birth is a potential source of bias in the procedure of estimating mortality using the parental survival approach. A similar challenge was encountered in this study concerning the estimation of the mean age of fathers at the birth of the children. The approach used in this study was to estimate the mean age of fathers based on the age of fathers of the children of women reporting a birth. However, data on recorded births are not without limitations, there is a low proportion of recorded births that contain information on fathers by age and thus the calculation of age of father is confined to births where information for both mother and fathers is available, which may be unrepresentative of ages of all parents at birth.

5.4 Conclusion

Based on the comparison of estimates obtained from applying the single-survey orphanhood approach to year-of-death data, there is inadequate evidence to confirm or dismiss the usefulness of the year-of-death data to produce plausible estimates of adult mortality. This is because, the estimates of mortality produced from using year-of-death data and those produced from the conventional approach are both implausibly lower than the RMS benchmark estimates. The estimates of adult male mortality derived from year-of-death data are however worse than those from the conventional form of the data. The inconsistencies associated with the year-of-death data such as recall errors, recording errors and the potential bias introduced by proportionally distributing data with unknown year-of-death may explain, to some extent, the implausibly low estimates produced by using these data. However, while conventional orphanhood data are also prone to errors which may result in biased estimates of mortality being produced, the results obtained suggest the single-survey approach is of limited usefulness for populations suffering a significant HIV/AIDS epidemic.

It is difficult to draw a general conclusion about whether year-of-death data can be used to produce plausible estimates of mortality when used in the two-survey approach. This is because the two data sets of the year-of-death data used in this study produce conflicting results. One data set produces estimates that are surprisingly consistent to benchmark estimates despite the data being of more questionable quality while the other set produces estimates that are implausibly low. While differences in the results were expected, the extent of the differences reflected in the results was not expected. There is insufficient evidence from these results to draw a definite conclusion about how useful year-of-death data are in estimating adult mortality.

An extension of this approach which makes use of orphanhood data from younger respondents, produces estimates of adult male and female mortality that are implausible low. Based on these results, one can conclude that this approach is inappropriate for our inquiry and does not help us determine whether year-of-death data can be used to produce estimates of mortality that are plausible. Further, these results are an indication that the adoption effect, which is concentrated among younger respondents, is present in these data and including data from younger respondents has a noticeable impact on the estimates.

The year-of-death data when used with the intercensal survival approach produces results that are surprisingly consistent with benchmark estimates. The results suggest that year-of-death data can produce plausible estimates of mortality when used with this approach. The year-of-death data produces plausible estimates of both adult male and female mortality even after proportionally distributing data with unknown year-of-death.

The general conclusion from this research is, that it may be possible to produce plausible estimates of mortality using the year-of-death data. In some instances, the estimates derived from year-of-death data match those estimates derived from the conventional orphanhood approach. However, when the year-of-death data are used with the single-survey and two-survey approach, the results are inconclusive as to the usefulness of these data in producing plausible estimates. The results suggest plausible estimates of adult mortality that are consistent with benchmark estimates may be produced from year-of-death data when the intercensal survival approach is applied to these data.

5.5 Areas of further research

Further research should be conducted into how best to deal with deaths with unknown year of death, which could include investigating the distribution of deaths by the year of death and investigating the levels of non-response that can be deemed acceptable and have minimal effect on estimating mortality.

The intercensal approach used in this study does not adjust for HIV-related biases. Further research into extending this approach to adjust for HIV-related biases should be considered. In addition, the effect of such an adjustment or lack thereof on the overall estimates of adult mortality obtained by using the intercensal approach should be assessed. This will potentially enhance the applicability of this approach even in contexts where HIV treatment initiatives are not widespread.

The year-of-death data used in this study were collected in surveys which are prone to sampling errors. Even though the results from this study imply the data collected in a survey can be used to estimate mortality, there is need to add the question on the year of death in the upcoming round of censuses. In addition, a question asking whether the parent died in the past five years should be included in the data collection. This may counter some of the recall errors encountered when only data on the year of death is collected. Regarding the use of census data, there is need to assess whether censuses would be a better source of these data and if there is any evidence that suggests that the quality of census undertakings has been improving over time.

This research has compared the estimates derived from the intercensal survival approach and those from the traditional two-survey orphanhood method. While the results derived from these approaches are similar, it is important to investigate the relationship between the intercensal survival estimates and the traditional two-survey estimates.

The results from this research suggest that the adoption effect is a source of bias observed in the estimates of adult male and female mortality. This research did not explore the differences in effect of adoption on the year-of-death data in comparison to the conventional data estimates. This is an area needing further investigation.

There is a need for further research into methods and data for estimating the mean age of fathers at conception. This research relied on the information about the age of fathers of the children of women reporting a birth. However, this approach is limited because of the high proportion of births reported with missing father's information which will result in underestimating the mean age of fathers at conception.

The results from this study suggest the year-of-death data collected in South Africa can be used to produce plausible estimates of mortality. However, this does not mean that similar results would be produced from the year-of-death data collected in other countries in Sub-Saharan Africa or even other surveys in South Africa. There is therefore a need to collect the year-of-death data in other countries and investigate whether these data will produce better estimates of adult mortality than the conventional from of the orphanhood data. The results in this study demonstrate the broader potential of this approach and its contribution to the description of mortality trends in other countries where vital records and census are lacking or deficient.

REFERENCES

Blacker, John GC. 1977. "The estimation of adult mortality in Africa from data on orphanhood", *Population Studies* **31**(1):107-128.

Blacker, John GC. 1984. "Experiences in the use of special mortality questions in multi-purpose surveys: the single-round approach," Paper presented at Databases for Mortality Measurement. Bangkok, 20-23 October 1981. United Nations, Department of International Economic and Social Affairs. 79-89.

Brass, William and Elijah A Bamgboye. 1981. *The time location of reports of survivorship : estimates for maternal and paternal orphanhood and the ever-widowed*. London: London School of Hygiene & Tropical Medicine, University of London.

Brass, William and Kenneth H Hill. 1973. "Estimating adult mortality from orphanhood," Paper presented at International Population Conference. Liege, Belgium. International Union for the Scientific Study of Population. Vol. 3:111-123.

Chackiel, J and H Orellana. 1985. "Adult female mortality trends from retrospective questions about maternal orphanhood included in censuses and surveys.," Paper presented at International Population Conference. Florence. International Union for the Scientific Study of Population. Vol. Volume 4:39-51.

Coale, Ansley J. and T. James Trussell. 1974. "Model Fertility Schedules: Variations in The Age Structure of Childbearing in Human Populations", *Population Index* **40**(2):185-258.

Dorrington, RE, D Bradshaw, R Laubscher and N Nannan. 2018. *Rapid Mortality Surveilance Report 2016*. Cape Town: South African Medical Research Council. www.mrc.ac.za/bod/bod.htm

Dorrington, Rob, David Bourne, Debbie Bradshaw, Ria Laubscher *et al.* 2001. *The impact of HIV/AIDS on adult mortality in South Africa*. 1919809147. Cape Town: Medical Research Council. www.mrc.ac.za/bod

Dorrington, Rob, Tom A Moultrie and Ian Timæus. 2004. *Estimation of mortality using the South African Census 2001 data*. Centre for Actuarial Research, University of Cape Town Cape Town.

Hill, K. 2001. "Methods for measuring adult mortality in developing countries: a comparative review," Paper presented at International Union for the Scientific Study of Population 24th General Conference. Salvador, Brazil, 18-24 August, 2001. Unpublished] 2001.

Hill, K. and J. Trussell. 1977. "Further Developments in Indirect Mortality Estimation", *Population Studies* **31**(2):313-334.

Hill, Kenneth, Yoonjoung Choi and Ian M Timæus. 2005. "Unconventional approaches to mortality estimation", *Demographic Research* **13**:281-300.

Hirschman, Charles, Samuel Preston and Vu Manh Loi. 1995. "Vietnamese casualties during the American war: A new estimate", *Population and Development Review* **21**:783-812.

IHME. 2018. *Global Burden of Disease Study 2017 (GBD 2017) Reference Life Table*. Seatle, United States: Institute for Health Metrics and Evaluation (IHME).

Kupamupindi, Takura. 2010. "Estimating adult mortality in South Africa using orphanhood and year of death data from the 2008 National Income Dynamics Study." Unpublished thesis, University of Cape Town.

Makinson, Carolyn. 1993. "Estimates of adult mortality in Burundi", *Journal of Biosocial Science* **25**(2):169-186.

Odimegwu, Clifford, Vesper H. Chisumpa and Oluwaseyi Dolapo Somefun. 2018. "Adult mortality in sub-Saharan Africa using 2001-2009 census data: does estimation method matter?", *Genus* **74**(1):10-10.

Page, H. J. and G. Wunsch. 1976. "Parental Survival Data: Some Results of the Application of Ledermann's Model Life Tables", *Population Studies* **30**(1):59-76.

Sadie, Jan L 1970. "An Evaluation of Demographic Data Pertaining to the Non-White Population of South Africa 1", *South African Journal of Economics* **38**(1):1-24.

SALDRU. 2018. National Income Dynamics Study 2017, Wave 5. 1.0.0. Pretoria: Southern Africa Labour and Development Research Unit. DataFirst. https://doi.org/10.25828/fw3h-v708

Statistics South Africa. 2003. South African Census 2001. 1.1. Pretoria: Statistics South Africa. DataFirst. https://doi.org/10.25828/bp70-m263

Statistics South Africa. 2015a. *Census 2011: Estimation of Mortality in South Africa.* 03-01-62 Pretoria, South Africa:

Statistics South Africa. 2015b. South African Census 2011, 10% sample 2. Pretoria: Africa, Statistics South. DataFirst. https://doi.org/10.25828/vjy1-tz66

Statistics South Africa. 2017. Community Survey 2016. 1. Pretoria: Africa, Statistics South. DataFirst. https://doi.org/10.25828/12sy-yj26

Timæus, Ian M. 1986. "An Assessment of Methods for Estimating Adult Mortality from Two Sets of Data on Maternal Orphanhood", *Demography* **23**(3):435-450.

Timæus, Ian M. 1990. "Advances in the measurement of adult mortality from data on orphanhood." Unpublished thesis, London School of Hygiene & Tropical Medicine.

Timæus, Ian M. 1991a. "Estimation of Adult Mortality from Orphanhood Before and Since Marriage", *Population Studies* **45**(3):455-472.

Timæus, Ian M. 1991b. "Estimation of Mortality from Orphanhood in Adulthood", *Demography* **28**(2):213-227.

Timæus, Ian M. 1991c. "Measurement of Adult Mortality in Less Developed Countries: A Comparative Review", *Population Index* **57**(4):552-568.

Timæus, Ian M. 1992. "Estimation of adult mortality from paternal orphanhood: a reassessment and a new approach", *Population Bulletin of the United Nations* **33**(1992):47-63.

Timæus, Ian M. 2013a. "Indirect estimation from orphanhood in multiple inquiries," in Moultrie, TA, RE Dorrington, AG Hill, K Hill, IM Timæus and B Zaba (eds). *Tools for Demographic Estimation.* International Union for the Scientific Study of Population, Paris, pp. 293-308.

Timæus, Ian M. 2013b. "Indirect estimation of adult mortality from orphanhood," in Moultrie, TA, RE Dorrington, AG Hill, K Hill, IM Timæus and B Zaba (eds). *Tools for Demographic Estimation.* International Union for the Scientific Study of Population, Paris, pp. 222-243.

Timæus, Ian M and Wendy Graham. 1989. *Measuring adult mortality in developing countries: a review and assessment of methods.* WPS155. Washington, D.C: World Bank Group.

Timæus, Ian M and Andrew J Nunn. 1997. "Measurement of adult mortality in populations affected by AIDS: an assessment of the orphanhood method", *Health Transition Review* **7**:23-43.

Udjo, Eric O. 2014. "Estimating demographic parameters from the 2011 South Africa Population Census", *African Population Studies* **28**(1):564-578.

United Nations. 1983. *Manual X. Indirect techniques for demographic estimation.* New York: Department of International Economic and Social Affairs, United States National Academy of Sciences.

United Nations. 2017. *World Population Prospects: The 2017 Revision*. New York: Department of Economic and Social Affairs, Population Division. https://population.un.org/wpp/Download/Standard/Mortality/

Zlotnik, Hania and Kenneth Hill. 1981. "The use of hypothetical cohorts in estimating demographic parameters under conditions of changing fertility and mortality", *Demography* **18**(1):103-122.

APPENDICES

Appendix A: Proportion of respondents with a surviving mother or father, at five and ten years before the survey: 2016 Community Survey, 2017 NIDS

	Proportion Surviving (2016 Community Survey)			*Proportion Surviving (2017 NIDS)*		
	2016	*2011*	*2006*	*2017*	*2012*	*2007*
Age group	*(At interview)*	*(5 years before)*	*(10 years before)*	*(At interview)*	*(5 years before)*	*(10 years before)*
Mothers surviving						
0-4	0.9906	0.9859	0.9852	0.9944	0.9936	0.9997
5-9	0.9609	0.9491	0.9516	0.9754	0.9585	0.9759
10-14	0.9169	0.9048	0.9234	0.9208	0.9187	0.9295
15-19	0.8674	0.8720	0.9071	0.8760	0.8782	0.9136
20-24	0.8309	0.8553	0.8902	0.8337	0.8670	0.8786
25-29	0.8095	0.8385	0.8704	0.8072	0.8234	0.8627
30-34	0.7859	0.8145	0.8219	0.7707	0.7779	0.8641
35-39	0.7526	0.7554	0.7504	0.6985	0.7773	0.7414
40-44	0.6765	0.6728	0.6494	0.6596	0.6323	0.6318
45-49	0.5795	0.5598	0.5371	0.5332	0.5162	0.5434
Fathers Surviving						
0-4	0.9800	0.9761	0.9735	0.98335	0.9851	0.9847
5-9	0.9358	0.9215	0.9170	0.95125	0.9241	0.8998
10-14	0.8718	0.8523	0.8667	0.86392	0.8222	0.8206
15-19	0.7907	0.7922	0.8189	0.75761	0.7358	0.7712
20-24	0.7228	0.7399	0.7718	0.67261	0.6849	0.7386
25-29	0.6649	0.6901	0.7206	0.61110	0.6346	0.6515
30-34	0.6062	0.6369	0.6308	0.52823	0.5385	0.5930
35-39	0.5471	0.5397	0.5340	0.42479	0.4765	0.4332
40-44	0.4427	0.4428	0.4099	0.35593	0.3257	0.3012
45-49	0.3463	0.3204	0.2982	0.22707	0.2084	0.2380

Appendix B: Proportion of respondents with a surviving mother or father: 2001 and 2001 Census

	Proportion Surviving (2001and 2011 Censuses)	
Age group	*2011 Census*	*2001 Census*
Mothers surviving		
0-4	0.9780	0.9857
5-9	0.9424	0.9684
10-14	0.8943	0.9527
15-19	0.8527	0.9353
20-24	0.8221	0.9057
25-29	0.7958	0.8566
30-34	0.7612	0.7902
35-39	0.6991	0.7033
40-44	0.6053	0.6075
45-49	0.4925	0.4936
Fathers Surviving		
0-4	0.9279	0.9404
5-9	0.8610	0.8953
10-14	0.7804	0.8541
15-19	0.7096	0.8021
20-24	0.6503	0.7282
25-29	0.5898	0.6293
30-34	0.5171	0.5229
35-39	0.4226	0.4030
40-44	0.3163	0.2989

Appendix C: Number of respondents with a surviving mother or father, at the survey by sex of respondent: 2016 Community Survey

	Number of respondents with a surviving father			
Age Group	*Yes*	*No*	*Don't know*	*Total*
Male respondents				
0-4	2,861,649.70	57,777.28	97,268.28	3,016,695.26
5-9	2,525,997.70	173,573.50	128,796.64	2,828,367.84
10-14	2,143,618.10	310,423.30	150,699.23	2,604,740.63
15-19	1,889,294.80	496,461.30	169,587.00	2,555,343.10
20-24	1,798,822.30	673,988.10	186,062.99	2,658,873.39
25-29	1,658,790.60	811,186.00	196,281.89	2,666,258.49
30-34	1,238,301.10	784,903.60	163,559.70	2,186,764.40
35-39	972,886.90	771,911.00	159,999.08	1,904,796.98
40-44	661,340.10	810,672.72	149,457.77	1,621,470.59
45-49	437,412.70	780,539.20	141,805.53	1,359,757.43
50-54	238,267.10	740,297.90	136,801.34	1,115,366.34
55-59	130,242.20	662,583.10	120,734.86	913,560.16
60-64	57,514.76	551,456.90	95,284.22	704,255.88
65-69	24,582.65	405,774.80	69,961.15	500,318.60
70-74	8,143.61	263,759.90	48,303.22	320,206.72
75-79	2,714.36	135,202.90	25,277.25	163,194.51
80-84	648.83	61,556.35	11,901.26	74,106.44
85+	715.53	41,640.17	10,793.58	53,149.28
Total	16,650,943.04	8,533,708.01	2,062,574.99	27,247,226.04
Female respondents				
0-4	2,803,096.40	57,881.30	98,846.06	2,959,823.76
5-9	2,484,993.80	170,406.30	136,027.88	2,791,427.98
10-14	2,119,112.60	316,527.80	149,422.23	2,585,062.63
15-19	1,877,367.30	500,347.90	171,423.49	2,549,138.69
20-24	1,764,374.70	692,338.20	186,748.49	2,643,461.39
25-29	1,595,029.70	828,611.50	190,604.86	2,614,246.06
30-34	1,263,320.90	840,182.10	164,420.69	2,267,923.69
35-39	959,664.20	827,758.20	155,741.94	1,943,164.34
40-44	650,321.30	840,739.60	148,052.37	1,639,113.27
45-49	427,497.50	852,250.70	143,424.74	1,423,172.94
50-54	242,910.50	834,736.20	141,496.03	1,219,142.73
55-59	130,986.80	795,724.60	133,924.85	1,060,636.25
60-64	59,737.36	690,309.40	118,614.89	868,661.65
65-69	26,938.94	553,627.70	98,401.96	678,968.60
70-74	10,030.13	410,762.50	83,733.91	504,526.54
75-79	4,340.75	264,997.30	53,804.19	323,142.25
80-84	1,833.72	142,474.00	32,592.49	176,900.21
85+	1,165.60	121,751.80	34,997.68	157,915.08
Total	16,422,722.21	9,741,427.10	2,242,278.75	28,406,428.06

	Number of respondents with a surviving mother			
Age Group	*Yes*	*No*	*Don't know*	*Total*
Male respondents				
0-4	2,981,395.50	28,133.08	7,166.75	3,016,695.33
5-9	2,704,491.10	111,401.00	12,475.70	2,828,367.80
10-14	2,374,582.40	212,402.40	17,755.83	2,604,740.63
15-19	2,200,749.90	331,534.40	23,058.71	2,555,343.01
20-24	2,190,046.70	435,323.40	33,503.40	2,658,873.50
25-29	2,140,945.60	482,832.00	42,480.79	2,666,258.39
30-34	1,701,024.50	444,553.40	41,186.55	2,186,764.45
35-39	1,413,703.90	444,160.30	46,932.87	1,904,797.07
40-44	1,075,127.50	492,596.60	53,746.49	1,621,470.59
45-49	772,622.80	527,102.20	60,032.42	1,359,757.42
50-54	488,738.60	558,282.30	68,345.56	1,115,366.46
55-59	295,670.40	550,164.30	67,725.41	913,560.11
60-64	140,882.70	503,884.60	59,488.52	704,255.82
65-69	55,399.26	396,177.90	48,741.43	500,318.59
70-74	15,897.19	268,021.70	36,287.83	320,206.72
75-79	5,049.03	138,009.30	20,136.23	163,194.56
80-84	931.39	62,742.65	10,432.39	74,106.43
85+	769.85	43,105.99	9,273.44	53,149.28
Total	20,558,028.33	6,030,427.52	658,770.31	27,247,226.16
Female respondents				
0-4	2,924,679.70	27,980.54	7,163.54	2,959,823.78
5-9	2,672,482.40	107,663.50	11,282.13	2,791,428.03
10-14	2,353,586.40	216,178.40	15,297.94	2,585,062.74
15-19	2,188,416.50	339,462.70	21,259.48	2,549,138.68
20-24	2,165,287.20	451,279.05	26,895.24	2,643,461.49
25-29	2,073,101.30	508,635.70	32,509.01	2,614,246.01
30-34	1,742,115.70	493,336.00	32,472.01	2,267,923.71
35-39	1,418,372.80	486,805.90	37,985.58	1,943,164.28
40-44	1,065,094.50	530,672.80	43,346.03	1,639,113.33
45-49	775,892.89	596,598.90	50,681.18	1,423,172.97
50-54	508,173.00	650,337.60	60,632.12	1,219,142.72
55-59	314,107.10	679,927.40	66,601.80	1,060,636.30
60-64	159,538.80	640,318.10	68,804.74	868,661.64
65-69	67,447.93	547,874.20	63,646.47	678,968.60
70-74	25,279.12	419,347.39	59,899.99	504,526.50
75-79	8,748.87	272,919.10	41,474.22	323,142.19
80-84	2,945.96	147,472.90	26,481.37	176,900.23
85+	1,934.43	126,387.40	29,593.30	157,915.13
Total	20,467,204.60	7,243,197.58	696,026.14	28,406,428.32

Appendix D: Number of respondents with a surviving mother or father, at the survey: 2017 NIDS

	Number of respondents with a surviving father			
Age Group	*Yes*	*No*	*Don't Know*	*Total*
0-4	5,475,189.40	92,701.57	109,420.70	5,677,311.67
5-9	4,927,542.50	252,542.90	220,894.30	5,400,979.70
10-14	3,819,600.90	601,647.50	164,065.21	4,585,313.61
15-19	2,757,984.50	882,388.00	99,250.10	3,739,622.60
20-24	2,564,433.80	1,248,222.40	172,055.80	3,984,712.00
25-29	2,648,365.40	1,685,388.30	120,489.09	4,454,242.79
30-34	2,108,123.30	1,882,809.70	90,201.36	4,081,134.36
35-39	1,350,959.30	1,829,325.90	49,972.24	3,230,257.44
40-44	859,826.40	1,555,921.40	15,417.75	2,431,165.55
45-49	472,726.00	1,609,095.50	13,280.10	2,095,101.60
50-54	189,483.80	1,581,818.00	1,447.61	1,772,749.41
55-59	69,264.77	1,418,096.10	5,434.70	1,492,795.57
60-64	28,088.56	1,156,105.00	4,505.91	1,188,699.47
65-69	2,947.45	918,847.20	363.06	922,157.71
70-74	0	632,551.40	438.66	632,990.06
75-79	0	396,549.20	171.78	396,720.98
80-84	0	192,856.20	0	192,856.20
85+	0	150,305.90	0	150,305.90
Total	27,274,536.08	18,087,172.17	1,067,408.37	46,429,116.62

	Number of respondents with a surviving mother			
Age Group	*Yes*	*No*	*Don't Know*	*Total*
0-4	5,649,908.50	32,080.10	4,546.11	5,686,534.71
5-9	5,316,205.60	134,281.10	17,516.19	5,468,002.89
10-14	4,292,743.50	369,102.00	8,397.70	4,670,243.20
15-19	3,335,514.40	472,048.60	8,894.54	3,816,457.54
20-24	3,383,683.00	675,095.70	14,804.32	4,073,583.02
25-29	3,687,080.30	880,671.50	12,638.67	4,580,390.47
30-34	3,316,108.80	986,533.40	5,045.95	4,307,688.15
35-39	2,358,099.80	1,017,894.10	282.25	3,376,276.15
40-44	1,681,228.00	867,448.90	15,221.92	2,563,898.82
45-49	1,122,845.80	982,820.57	604.25	2,106,270.62
50-54	706,164.40	1,072,130.40	548.96	1,778,843.76
55-59	428,241.40	1,033,820.30	74.58	1,462,136.28
60-64	231,120.60	936,519.90	0	1,167,640.50
65-69	80,322.29	802,270.10	363.06	882,955.45
70-74	18,645.98	581,613.50	0	600,259.48
75-79	4,704.36	379,268.70	0	383,973.06
80-84	443.94	192,095.14	0	192,539.08
85+	0	150,305.90	0	150,305.90
Total	35,613,060.67	11,565,999.91	88,938.49	47,267,999.07