

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

**The Effect of HIV on the Orphanhood Method of Estimating Adult
Female Mortality**

Stephen Kramer

Thesis submitted to the Faculty of Commerce in partial fulfillment for
the Degree of Master of Philosophy in Demography

University of Cape Town

August 2009

Abstract

The orphanhood method of estimating adult mortality is widely used in developing countries. The method is subject to a number of assumptions, some of which are violated when a generalized high-prevalence HIV epidemic is present. Non-independence of the mortality of children and mothers, relationships between HIV infection and fertility, and changes in age-specific mortality result in biases that affect the accuracy of the method. An earlier study has examined some of these sources of error, and proposed adjustments to enable continued use of the method. This earlier research, however, uses data from populations with lower HIV prevalence rates than are currently being experienced in much of southern Africa, and is based on specific assumptions about HIV and its effects on mortality and fertility. The effects of HIV on the method are investigated in this research using mathematical modelling of the effects on Black South African females – a population with high HIV prevalence. More is now known about HIV and its effects on mortality and fertility, and these effects are explicitly reflected in the ASSA2002 model which provides much of the data for this research. The research compares the simulated survival of various cohorts of women: those aged 25 in a certain year, women (with an age profile identical to that of mothers), mothers, and mothers as reported by their children. In this way the various sources of error are explicitly identified and the errors quantified. The timing, magnitude, and combined effects of the errors are studied in relation to the emergence and spread of HIV, indicating when the errors might be expected to be large enough to invalidate the method. Errors that bias the outcomes of the orphanhood method take a number of years to develop after HIV starts spreading. Substantial biases in reported survival emerge between 20 and 35 years after the start of an HIV epidemic, in a high prevalence setting. These errors are reduced by the use of antiretroviral and prevention of mother-to-child transmission, but biases remain large enough to invalidate outcomes when the unadjusted method is applied in most southern African countries. An adjusted method has been proposed which substantially reduces error, except when adjusting survival reported by the two youngest age groups. This adjusted method can be applied, but further research to identify revised adjustments would further improve the accuracy of the method.

Table of contents

List of Tables	v
List of Figures	v
Chapter 1: Introduction	1
1.1 Chapter outline	4
Chapter 2: Literature Review	5
2.1 Development of the method.....	6
2.2 Sensitivity to violations of assumptions.....	10
2.3 Impact of a substantial HIV epidemic on the method	12
2.4 Application of the method in countries with generalized high prevalence HIV epidemics, and comparison to other estimates of mortality.....	14
2.5 Conclusion of Literature Review	16
Chapter 3: Methodology	18
3.1 Outline of method and rationale.....	18
3.2 Description of the ASSA2002 model and how it is used	21
3.3 Assumptions within the ASSA2002 model.....	23
3.4 Survival of women aged 25, women, mothers, and mothers as reported by children	24
3.5 Comparison of survival and non-survival: mothers as reported by children vs. mothers; mothers vs. women; and women vs. women aged 25.....	29
Chapter 4: Results	33
4.1 Comparison of survival and non-survival: women, mothers, and mothers as reported by children	33
4.1.1 Individual age-specific cohorts of women and mothers in a particular year	33
4.1.2 Births in specific years – mothers (of various ages) aggregated	41
4.1.3 Women, mothers, and reported mortality of mothers in period format	44

4.1.4 Women, mothers, and reported mortality of mothers in period format grouped in 5-year age bands	51
4.2 Comparison of survival and non-survival: women aged 25 and women (with the age distribution of mothers).....	54
4.2.1 Births in specific years – women (with the age distribution of mothers) vs. 25 year olds	55
4.2.2 Women (with the age distribution of mothers) vs. 25 year old women, in period format	56
4.2.3 Women of child-bearing age vs. 25 year old women, in period format grouped in 5-year age bands	59
4.3 Reported survival of mothers adjusted using coefficients vs. survival of 25 year olds.....	61
4.4 Summary.....	63
Chapter 5: Discussion and conclusions	68
References	73
Appendices	75
Appendix A: Alternative view of Figure 21 - years on x-axis	75
Appendix B: South African Black females – modelled q_x	76
Appendix C: Summary of ratios of survival – “No AIDS” scenario	77
Appendix D: Summary of ratios of survival – “MTCTP and ART” scenario	78

List of Tables

Table 1: Antenatal clinic HIV prevalence - South Africa.....	23
Table 2: Ratio of fertility of HIV-positive women to fertility of HIV-negative women, by year.....	24
Table 3: Summary of ratios of survival of various categories of women - by cause of error, year and age group of respondent.....	64
Table 4: Summary of ratios of non-survival of various categories of women - by cause of error, year and age group of respondent.....	65

List of Figures

Figure 1: Age profile of women vs. mothers in 2005 (from the ASSA2002 model).....	34
Figure 2: Proportion uninfected, or infected for various durations - women and mothers giving birth in 2005 (a) aged 20 and (b) aged 30.....	35
Figure 3: Survival of women, mothers and mothers as reported by children, by duration since birth in 2005: (a) ages 15, (b) ages 25, (c) ages 35, and (d) all ages combined.....	36
Figure 4: Ratio of proportion not surviving of mothers vs. women, reported vs. mothers and mothers as reported by children vs. women, by duration since birth in 2005: (a) ages 15, (b) ages 25, (c) ages 35, and (d) all ages combined.....	38
Figure 5: Proportions not surviving of women, mothers, and mothers as reported by children, by year of birth of their children: (a) 1995, (b) 2005, and (c) 2020.....	42
Figure 6: Ratio of proportions not surviving of mothers vs. women, mothers as reported by children vs. mothers, and mothers as reported by children vs. women, by year of birth of their children: (a) 1995, (b) 2005, and (c) 2020.....	43
Figure 7: Ratio of proportions not surviving by birth year of child (cohort format) - mothers/women.....	45

Figure 8: Ratio of proportions not surviving by calendar year (period format) - mothers/women	46
Figure 9: Ratio of proportions not surviving by birth year of child (cohort format) - reported/mothers	48
Figure 10: Ratio of proportions not surviving by calendar year (period format) - reported/mothers	49
Figure 11: Ratio of proportions not surviving by calendar year (period format) - reported/women.....	50
Figure 12: Ratio of proportions surviving by calendar year (period format) - reported/women.....	51
Figure 13: Ratio of proportions surviving by calendar year (period format) and 5-year age band - reported/women.....	52
Figure 14: Ratio of proportions not surviving by calendar year (period format) and 5-year age band - reported/women.....	53
Figure 15: Ratio of proportions surviving by calendar year (period format) and 5-year age band - reported (adjusted as proposed by Timaeus and Nunn (1997))/women.....	54
Figure 16: Proportions not surviving of women and 25 year old women: (a) 1995, (b) 2005, and (c) 2020	56
Figure 17: Ratio of proportions not surviving by birth year of child (cohort format) - women/(women aged 25)	57
Figure 18: Ratio of proportions not surviving by calendar year (period format) – women/(women aged 25) ...	58
Figure 19: Ratio of proportions surviving, by calendar year (period format) and 5-year age band: women/(women aged 25 at birth of child)	60
Figure 20: Ratio of proportions surviving, by calendar year (period format) and 5-year age band - women(adjusted using coefficients)/(women aged 25 at birth of child).....	61
Figure 21: Ratio of proportions surviving, by calendar year (period format) and 5-year age band - reported(adjusted using coefficients)/(women aged 25 at birth of child).....	62
Figure 22: Ratio of proportions surviving, by calendar year (period format) and 5-year age band – reported (using adjustments for selection bias and coefficients proposed by Timaeus and Nunn (1997))/(women aged 25 at birth of child).....	63

Chapter 1: Introduction

The orphanhood method is a widely used indirect method for estimating adult mortality. The technique involves participants in a survey or census being asked their age and sex, and whether their mother or father is still alive. The proportion alive, as reported by children, constitutes a conditional probability of survival, and can be used to estimate adult mortality, subject to the validity of certain assumptions about the population.

Clearly, the father was alive at the time of conception of his child, and the mother alive at the time of birth. The ensuing period (the age of the child when assessing mother's survival) is the period during which there has been exposure to the risk of dying, and survival during this period can be used to estimate mortality. An advantage of making use of these data is that they can be collected through fairly simple census or survey questions (e.g. "Is the person's (biological) mother/father alive?", "How old is the person?"). The method is however subject to a few key assumptions and limitations, namely:

- The mortality of the parent (whether mother or father) and that of the child should be independent. The answer to the question about parental survival refers to biological parents, and circumstances or cultural factors that cause widespread deviation from this may cause bias (i.e. adoption may cause over-reporting of survival of parents).
- The method estimates survival for women and men that have children – a difference in mortality between women or men who have children and those who do not will mean the mortality estimates are not strictly applicable to the general population. Similarly, the mortality of parents who have large families will be more heavily weighted in the method, and the mortality experience of those whose children do not survive will be underweighted.
- The coefficients used to convert survival ratios of mothers into the life table measure $l_{(25+x)}/l_{(25)}$ are dependent on assumptions about how the survival of women aged 25 relates to the survival of mothers, based on standard life tables.
- The coefficients for estimating the timing of deaths are based on mortality schedules that follow a gradual pattern of increase by age, and a slow and constant decrease over time. Should mortality patterns deviate

from the pattern of a gradual increase by age, or gradual constant decrease over time, the estimates of the timing of deaths are likely to be incorrect.

In many developing countries where this method is applied, a number of the above assumptions are violated to a greater or lesser extent. The prevalence of adoption, non-independence of the mortality of parents and children, and the small differences between the mortality of those who bear children and those that do not may, before the start of the HIV epidemics in various countries, have had relatively small effects.

The presence of a generalized HIV epidemic causes substantial further violations of the assumptions: the independence of mortality between parent and child is impaired; fertility is reduced at longer durations of HIV infection, and is also related to HIV in that more sexually active women are both more fertile and are more likely to become infected; adoption patterns may be distorted in that many adoptive parents may be grand-parents; and mortality levels and patterns show sudden changes over time, causing a “hump” of mortality at certain ages. Some work has been done on understanding the effects of these violations of assumptions, but much of it used data from populations with lower prevalences of HIV than is currently being observed in southern Africa. Most notably, Timaeus and Nunn (1997) proposed an adjustment to reported survival and the use of revised coefficients for calculating $l_{(25+x)}/l_{(25)}$, to accommodate the errors and biases caused by the presence of HIV. The revised method was based on various assumptions and on data from Masaka in Uganda. The epidemic in Uganda had not yet peaked or reached a plateau though, and prevalence was substantially below what we currently see in southern Africa.

Violations of the assumptions of the orphanhood method during a high prevalence HIV epidemic may render the life table measure that the method provides, namely $l_{(25+x)}/l_{(25)}$, inaccurate. Not enough is known about the magnitude of the errors that will result from each of three identified areas of potential bias, or whether the method, as it is currently documented, can be used in countries where high prevalence HIV epidemics exist.

The purpose of this research is to examine violations of assumptions that underlie the use of reported survival and the coefficients to calculate $l_{(25+x)}/l_{(25)}$, and to determine the magnitude of the errors caused by these violations. The size of these errors, at various ages and times, will be a key determinant of when, in terms of the epidemic, the method can be used.

The research investigates three main areas of error or bias. First, it investigates the error resulting from the correlations in mortality between mothers and their children due to mother-to-child transmission. Second, it investigates the error due to the relationship between HIV and fertility (whether one causes the other or is simply correlated with the other). Third, it investigates errors resulting from the use of a cohort of mothers across the age range of child-bearing to represent the mortality of women who were 25 years old in a particular year ($l_{(25+x)}/l_{(25)}$). The combined effect of these sources of bias and the individual contribution of each to the total error are determined.

The research considers only the relationship between mothers and female offspring – it therefore only examines biases in estimating female mortality, using reporting by female children on the survival of their mothers. The inclusion of reporting only by female offspring is done both to reduce the amount data needed to do the modelling, and because women have in many cases been favoured as respondents as they appear less likely to mis-state their age (Blacker 1977) – the outcomes would, however, be very similar if considering male offspring (Blacker 1977).

The research concerns itself with the direct effects of HIV – namely, its effect on fertility, its effect on the relationship between the mortality of mothers and their children as a consequence of mother-to-child transmission of HIV, and its effect on mortality of different cohorts of mothers in the years after they give birth. Indirect effects such as correlations in the mortality of mothers and daughters in later years due to related socio-economic conditions, or effects due to changes in adoption patterns, are not examined in this research. This research also does not investigate the appropriateness of the use of model life tables which have an age structure different to that of a population heavily affected by HIV to translate the initial measure of mortality, $l_{(25+x)}/l_{(25)}$, into other measures such as ${}_35q_{30}$.

An assessment of the extent to which a high prevalence generalized HIV epidemic renders the method inaccurate is needed to determine whether the method can or cannot be used, or whether it can be used during certain periods during an HIV epidemic, or at certain ages. It is possible that revised coefficients or additional adjustments could be identified to convert survival measures more accurately into $l_{(25+x)}/l_{(25)}$ during a high prevalence generalized HIV epidemic, but this is not the purpose of this investigation. The question also arises as to the size of the errors during the transition to a point where the epidemic is stable, as opposed to all

years when the epidemic is present. This research seeks to provide sufficient evidence to allow researchers to make informed decisions about whether to use the method in countries with similar high prevalence generalized HIV epidemics.

1.1 Chapter outline

After the introduction there are four chapters in the following order: a literature review, a description of the method used to do the research, presentation of results, and finally a discussion and conclusion.

The literature review (Chapter 2) has a short introduction, and then deals with the development of the orphanhood method, followed by a summary of research previously conducted that has concerned itself with the sensitivity of the method to violations of the assumptions. This, in turn, is followed by a summary of prior published research on the impact of a substantial HIV epidemic on the accuracy of the method, and a section comparing mortality estimates generated using the method and estimates generated using other methods in countries with generalized high prevalence HIV epidemics.

The chapter describing the method (Chapter 3) first gives an outline of the method and approach, then describes how the future survival of various categories of women (women, mothers, mothers as reported by children, and women aged 25) is calculated. The last part of this chapter describes how comparison of the survival of these groups produces an estimate of the error due to each of the sources of bias, and an estimate of their combined effect.

Chapter 4 presents results of the calculations described in Chapter 3. There are four main sections. The first concerns itself with the errors in the survival measures used by the method (i.e. data on reported survival from a census). The second deals with errors inherent within the method in assuming that a group of women over a wide age range can be used as a proxy for survival of women in one age group near the mean age (i.e. 25 years old in a particular year). The third looks at the combined effect of the sources of bias dealt with in the first two sections. The chapter concludes with a summary of the contributions of each source of bias to the total error.

Chapter 5 discusses the results and the implications of these results for use of the method, and outlines suggested areas where further research would be useful.

Chapter 2: Literature Review

For the purposes of investigating the value of the orphanhood method of estimating mortality in a context heavily impacted by HIV, the literature dealing with the method is categorized into four main areas of research.

The first deals with the development of the method itself. This involves examining the initial descriptions of the ideas behind the method, development of the concept of the method, descriptions of the assumptions and limitations of the method, and examines subsequent improvements made since the methods were initially proposed. Specific areas of research include the development of the coefficients used to convert reported survivorship into life table measures and the development of adjustments that provide time location of the mortality measures. Uses and applications of the method which have led to a better understanding of biases, and the remedies for these biases, are also included in this section. Most of the literature covered predates the start of the global HIV epidemic.

The second area of research delineated is that which covers investigations into the sensitivity of the technique to violations in the assumptions (including limitations in the use of the method). This research explicitly focuses on the accuracy, or degree of inaccuracy, of the method when certain basic assumptions are violated. Some violation of assumptions always exists – but what this section concerns itself with is the extent to which these violations lead to bias in the estimates and the extent to which various errors and biases cancel each other out.

The third area of research looks specifically at the issue of HIV and how, through its noticeable impact on the level of mortality, the age distribution of mortality, and on correlations between adult mortality and that of offspring, it may render the orphanhood method less accurate. The technique is not only tested, but suggestions exist in the literature as to how the method could be adapted to be more accurate in the context of a high-prevalence generalized HIV epidemic.

The fourth area is concerned with the extent of the error in the estimates produced by the orphanhood method. Where possible comparisons with mortality estimates derived using direct techniques or other indirect techniques are made, with a view to identifying any noticeable bias or inconsistency.

2.1 Development of the method

The idea of using reported orphanhood to estimate adult mortality was first developed by Brass and Hill (1973). Using a model based on standard fertility and mortality schedules, the mathematical relationship between proportions orphaned and life table survivorship of the parents of each 5-year cohort was developed and defined. Coefficients allowing proportions orphaned (by 5-year age group of whom the question was being asked) to be converted into life table survivorship of their parents were published (Brass 1973), ensuring a practical tool for use in estimating mortality when the orphanhood question was asked in censuses and surveys. The model was shown to be fairly robust in situations where any changes in either mortality or fertility are gradual. The coefficients were based on model life tables – the assumption being that most populations where the orphanhood method might be applied had mortality schedules of a similar shape to each other, even if the level of mortality was higher or lower. Allowance was made for higher or lower average ages of child-bearing, but not for substantial changes in the age patterns of fertility. Blacker (1977) applied the method in various African countries including Chad, Kenya and Malawi.

The data Blacker used in assessing mortality in Chad were from the 1964 Demographic Survey – a single round retrospective survey. In analyzing these data, Blacker found that age misreporting among men appeared to be greater than amongst women, and suggested that when using data from the orphanhood question asked of females (as opposed to males) the method was more likely to give reliable estimates. Also apparent in the data from Chad was bias resulting from what has since become known as the "adoption effect" - this refers to the deliberate or inadvertent reporting of an adopted parent as a biological parent when answering the question as to whether a parent is alive. Field workers conducting surveys (or respondents) may simply assume that parents are biological parents, and different cultural practices may mean different understandings of parenthood within and between countries. The "adoption effect" is greater when considering the survival of parents of children at younger ages. The reason for this is twofold - younger children are less likely to be aware that they are adopted, and at older ages it is more probable that this bias is eliminated as both adoptive and biological parents are dead.

The data used in assessing mortality in Kenya were from the 1969 census – the orphanhood question being asked of all urban dwellers and a sample of rural dwellers. Once again, Blacker prefers the mortality measures derived from the orphanhood question asked of female children (Blacker 1977).

Assessing the method, having analyzed the three countries, Blacker concluded that all the outcomes appeared to be plausible, and in reasonable agreement with the estimates derived using other methods, but that this did not mean there was no bias. The results found in these applications of the method also led him to conclude that the modelled relationships between orphanhood and mortality hold even where patterns of fertility are quite different from those used to derive the factors. In conclusion Blacker (1977) identified five areas of potential bias. The first concerns what has become known as the “adoption effect”, leading to an underestimation of mortality. The second results from the fertility and mortality patterns being different in the population under study from those assumed in the model population used to develop the method. The third is age misreporting, with implications both for proportions with parent alive, and the mean age of child-bearing. The fourth was the possible effect of declining mortality. The final concern identified was the possibility of a bias flowing from a relationship between the mortality of a parent and the number of children that the parent had – parents of larger families being reported on more often by their many offspring. No mention was made of the potential correlation between the mortality of parents and children, or of a relationship between individual fertility and mortality risk.

Hill and Trussell (1977) investigated various developments in indirect estimation of mortality. Estimates of mortality derived from reported child survival, spouse survival and sibling survival are all given some consideration. However, Hill and Trussell (1977) largely concerned themselves with describing the method developed by Brass and Hill (1973), explaining in detail the derivation of the coefficients. Three areas of bias are mentioned: non-parents and parents of children who have all died and who are thus not being represented; any relationship between the number of surviving children and survival of the parents (implying both fertility and mortality correlations); and the “adoption effect”. Other indirect methods which are examined by Hill and Trussell (1977) (those based on survival of siblings and survival of spouse) are also subject to biases similar to those of the orphanhood method (although the biases may be less significant). Biases resulting from relationships between survival of parents and survival of children, survival of siblings, and survival of spouses are all of relevance when considering the HIV epidemic. In addition the fact that model life tables were used in

generating the coefficients for the method, and in converting estimate of survival from age 25 (for women) ratios into other life table measures means all indirect techniques which rely on model life tables to produce age-specific rates are to a greater or lesser extent subject to distortion, due the fact that the age profile of a life table in a high-prevalence generalized HIV epidemic is so different from the underlying or standard model life tables.

Timaeus (1986), using techniques proposed by Zlotnik and Hill (1981), proposed an adapted version of the orphanhood method based on the idea of hypothetical cohorts. The hypothetical cohort is constructed from data either 5 or 10 years apart, with each 5-year cohort being compared to itself at a time 5 or 10 years earlier. The comparison, in this way, of data from two surveys gives an estimate of mortality for the period between the surveys. The assumption is made that the mortality and likelihood of migration of respondents are not related to the mortality of parents. By "chaining" the results for each cohort, a form of period survival function can be established. In essence the effects of declining (or increasing) mortality over time are stripped out and a more accurate period measure is produced. In addition, certain inconsistencies may come to light that may otherwise have remained unknown were only one set of data used. These include age exaggeration, sampling errors and possibly some evidence as to the extent of the "adoption effect".

Timaeus (1986) showed the results of application of variations of the orphanhood method by himself and others to data from Peru. Data used include the 1972 Census, the 1976 demographic survey, the 1977 Peru Fertility Survey and the 1981 Census. Various methods include intersurvey changes in orphanhood using 5-year cohorts in one instance and open-ended age groups in another, as well as synthetic cohorts, and the original method combined with time location calculations as proposed by Brass and Bamgboye (1981). He concluded, in assessing the results from Peru, that using open-ended age groups in the calculations produces implausible results. The other techniques appear to be in broad agreement for respondent ages 20 to 39, with some concern about under-estimation of mortality amongst the very young and the old using the original technique.

The various techniques for estimating mortality from orphanhood data were also tested on data from Kenya (using 1969 and 1979 Census' data). Timaeus (1986) concludes that results from the use of open-ended age groups in reporting orphanhood are again implausible and that estimates based on two sets of data

appeared to be the most reliable. Data from a third country to which the techniques were applied, Malawi (using the 1971 population change survey and 1977 Census), led to the same conclusions.

Timaeus (1986) argued that the removal of trends allows a more reliable analysis of the extent to which older respondents over-report parents alive. The use of two sets of data meant bias introduced by the “adoption effect” was limited, but that bias resulting from age exaggeration and under-reporting of orphanhood at older ages renders data for these age groups less useful. Nevertheless he concluded that combining two sets of data gave reliable estimates of mortality.

Timaeus (1991) proposed an adaptation of the method which uses only orphanhood in adulthood in order to avoid bias resulting from the “adoption effect”. The method sought to estimate mortality by using orphanhood data from those who had a living parent at age 20. The median age of mothers of those who have survived to age 20 will be between 45 and 55, while the median age of fathers will be between 50 and 60. This method may be useful in dealing with some of the underlying problems of the orphanhood method such as the “adoption effect”, but does not deal with many other problems in the application of the method, especially those relevant as a consequence of the HIV epidemic. This revised method also required additional information, as it was necessary to establish whether or not respondents had a living parent at age 20, and was thus likely to be impractical given the data sets available. It may be useful, however, in providing information on underlying non-HIV trends in mortality where sufficient data are available.

Timaeus (1991) proposed another method which estimates mortality from orphanhood before and since marriage. He argued that the conventional orphanhood method has a distinct limitation. It is only able to measure the overall levels and trends in mortality and not able to detect unusual age patterns or short-term fluctuations in mortality. He argued further that the declines in mortality estimated by the orphanhood method are implausibly rapid and that applications of the method to successive surveys show up inconsistencies. Both of these problems he ascribed to the distorting effect of adoption. By using data on orphanhood since marriage, the data are measuring more recent mortality than mortality based on survival since birth, as is the case in the conventional orphanhood method. This means more recent trends are likely to become clear when applying the orphanhood method only to orphanhood since marriage. In addition, Timaeus argued, the restriction of recollection to events occurring at ages when respondents would be old enough to remember them results in fewer reporting errors. This variation requires many of the same assumptions as the

conventional method. Timaeus tested this variation on two sets of data (from a survey conducted in 1987 as part of the DHS program in Morocco; and on data from the DHS program in Burundi at a date not given in the paper), and found the results were consistent with estimates derived using other methods (and in some cases more believable).

The revised method is similar in its applicability to the method using orphanhood in adulthood, and thus suffers from similar shortcomings in general and when applied specifically to an HIV impacted population. Timaeus (1991) concluded that indirect techniques as a whole are unable to give age-specific or time-specific estimates of mortality comparable to those given by direct estimates, but that the estimates from indirect methods should instead be taken to be broad indicators of the levels and trends of mortality in populations studied.

2.2 Sensitivity to violations of assumptions

Palloni *et al.* (1984), used mathematical modelling techniques to analyze systematically the sensitivity of the orphanhood method to various assumptions and biases. Quantitative evaluations were carried out to test sensitivity to three basic assumptions. These, as listed by Palloni *et al.* (1984), were: constancy of mortality prior to the survey; absence of adoption or the “adoption effect”; and third, the absence of correlations between fertility of the mother and mortality of the mother, fertility of the mother and mortality of the offspring, and mortality of the mother and mortality of the offspring. Expanding on the third area of concern, they argued that children of highly fertile mothers have higher mortality, resulting in less reporting on the survival status (be it higher or lower) of these women. Palloni *et al.* (1984), recognized other sources of bias, such as age mis-statement or under-enumeration (although they do not state whether by age, or by survival status of parent), and deviation of experience from the underlying mortality models, but consider the three areas on which they focus to be the ones likely to result in sizeable errors. Palloni *et al.* (1984) stressed the size of the potential bias resulting from declining mortality, reinforcing the conclusions of Blacker (1977). Subsequent to the conclusions of Blacker (1977), Brass and Bamgboye (1981) developed a technique to give estimates of mortality time locations. Palloni *et al.* (1984) conclude that by making the correct adjustments to ensure accurate time locations one can derive an accurate measure of mortality at various times. The method used to generate these time locations is either dependent on coefficients or on a formula, which are derived

from model life tables. The dependency of the techniques on the underlying model life tables, and the problem of significant deviation from the age patterns of mortality are not dealt with, presumably because they were assumed to be of little consequence at a time before the impact of the HIV/AIDS epidemic was apparent.

In assessing the bias resulting from the “adoption effect”, Palloni *et al.* (1984) showed that mortality is likely to be underestimated, when applying the orphanhood method, by at least 7 percent and possibly by as much as 10 percent for certain age groups. Important variables considered applicable in determining the size of the bias include the extent to which adoption happens, the median age of child-bearing, and the age of both adopted mother and of the child at the time at which adoption occurs.

In considering the relationships between fertility of mothers and the mortality of both mothers and offspring, and how this might affect the accuracy of the technique, a number of assertions were made. First, that the mortality of children is correlated to that of mothers due to similar living conditions in mother-child pairings, causing a bias which is likely to understate mortality. Second, that pregnancy and childbirth themselves are associated with morbid conditions resulting in heightened risk of mortality of both mother and child, causing a bias likely to underestimate mortality. Third, that women who have more children to report on their survival are also likely to experience higher mortality, the effect being an overstatement of mortality. Palloni *et al.* also hypothesize that mothers who survive longer are “more prolific”, but acknowledge that children of higher fertility women have higher mortality. In considering the fertility of mothers, the mortality of their offspring, and the mortality of mothers - and the relationships between these three variables, there are a number of biases influencing the estimates of mortality in conflicting directions resulting in a small net effect (Palloni *et al.* 1984). These conclusions were reached at a time when HIV was still relatively new and before clear evidence had emerged as to the scale and effects of the high-prevalence generalized epidemics on populations. It is not clear whether the conclusions reached by Palloni *et al.* (1984) remain valid in the context of HIV, but it seems unlikely that the effects described will remain negligible. The manner in which various biases cancel each other out, or reinforce each other, may be quite different in the presence of a high-prevalence generalized HIV epidemic.

2.3 Impact of a substantial HIV epidemic on the method

Timaeus and Nunn (1997), specifically considered the effect of a large HIV epidemic on the accuracy of the orphanhood method. The approach of their paper involved five main areas of investigation.

In the first a model was developed that articulated mathematically some of the biases that may arise due to AIDS. Three main sources of bias were identified, namely: bias as a consequence of the correlation between the mortality of mothers and children resulting from mother-to-child transmission of HIV, bias resulting from reductions in fertility due to HIV infections, and bias resulting from the coefficients used to convert reports on survival of parents into life table survivorship derived using inappropriate life tables and age-specific fertility schedules (with an age pattern substantially different from that existing in an AIDS epidemic).

In the second section of the paper the model developed in the first section was applied using HIV specific data, such as the effect of HIV infection on fertility, expected vertical transmission of HIV and expected survival of both infected mother and infected child. This model sought to quantify theoretically the extent of bias in a population with an HIV epidemic of a certain level. Some simplifying assumptions were made where limited data existed, or to limit the complexity of the modelling required. For example they assumed that all infected children die in the first 5 years of life; that 25 per cent of children of infected mothers are infected at birth; that mean adult survival time from infection with HIV is 6.7 years; and that all women sero-positive at the time of giving birth die within 5 years.

In the third section various parameters were adjusted upwards or downwards to test sensitivity and results were produced reflecting the biases estimated by the model. Having done this, the main conclusion reached is that biases resulting from vertical transmission and reduced fertility of HIV-positive women lead to an overestimation of adult survivorship of only a few percentage points. They derived a simple formula for correction of the estimate that can be used where sero-prevalence at the average time of birth of the children in each group is known. The adjusted survival $S(a) = \frac{1-0.25P}{1+0.25P} \cdot S^*(a) \approx (1-0.5P) \cdot S^*(a)$ is a function of only the prevalence of mothers P (at the time, on average, at which the births occur), and the unadjusted survival $S^*(a)$. They concluded, on the basis of comparisons produced using the model and having done

sensitivity analyses, that the existing regression equations produce reasonable outcomes where children are aged 15 years and over.

In the fourth section coefficients for conversion of survivorship into life tables measures are produced, using the model, which take into account various aspects of the HIV epidemic identified as being relevant to the orphanhood method. Data from the Masaka district in Uganda are used to generate the new coefficients. The same variables as are used in the sensitivity testing – prevalence of HIV amongst women aged 15-49, effect of HIV on fertility of those sero-positive, extent of vertical transmission of HIV and background life expectancy – are used to generate both expected reported survival under model conditions as well as life table survivorship, using the model. The relationship between expected reported survivorship and life table survivorship is shown to be linear. They conclude that this indicates that regression coefficients can be generated for an estimation of survivorship from orphanhood data in circumstances where there is a substantial HIV epidemic. They caution that the coefficients are provisional, as they have been generated using only data from the Masaka district, and also caution that these coefficients would not be ideal for a population in which the peak of prevalence by age is amongst older women (as opposed to peaking amongst those in their twenties). The assumption of the peak of prevalence amongst women in their late twenties likely results from the assumption of such a short survival period for infected women. Longer survival, given the same rate of incidence by age (peaking amongst those in their early twenties), would result in the peak of HIV prevalence being amongst women in their early thirties. They suggest that the coefficients be revised as more data become available from populations heavily impacted by HIV.

Lastly, the results are tested against data from Masaka in Uganda. In comparing the results of the model described in the paper with the evidence from Masaka in Uganda, it must be noted that this population had a prevalence of only 8 percent amongst adults and about 12 percent amongst pregnant women, which is far lower than many populations heavily affected by HIV. They conclude that the revised orphanhood method works quite well for this population, but caution that this is to some extent a consequence of the fact that the revised coefficients used to generate results were themselves generated using the Masaka data.

Many of the underlying assumptions were considered in the paper, and useful techniques are developed to ensure reduction or elimination of the biases generated by the violations of the assumptions that are likely in a high-prevalence HIV epidemic. The inappropriateness of using coefficients generated from unsuitable life

tables and fertility schedules was recognised and to some extent resolved by Timaeus and Nunn (1997). However, the problem conversion of $l_{(25+x)}/l_{(25)}$ into other life table measures remains. Using an inappropriate life table is likely to distort significantly the estimates of mortality which are produced. The very high rate of mortality likely affecting mothers shortly after the median age of child-bearing (given an HIV epidemic on the scale seen in parts of sub-Saharan Africa) would result in inappropriate conversions using standard life tables (which do not reflect the impact of HIV on mortality). Allowing for the peak of HIV-related mortality amongst women in their thirties seems central to an accurate conversion of $l_{(25+x)}/l_{(25)}$ into other life table measures.

2.4 Application of the method in countries with generalized high prevalence HIV epidemics, and comparison to other estimates of mortality

In order to compare the performance of the orphanhood method with other direct or indirect methods of mortality estimation in circumstances where HIV may be having an effect on the method, certain conditions are necessary. It is clear that estimates based on the orphanhood method are unlikely to be biased substantially in the first two decades after an epidemic has become generalized (generally taken to mean that the epidemic has spread beyond defined "high risk" populations – i.e. when prevalence amongst adults is more than 1% (UNAIDS 2007)). The main biases concern age-specific mortality rates and correlations between the mortality of mothers and their off-spring. Neither of these will have substantially affected estimates within the first 20 years of even a high prevalence generalized HIV epidemic – prevalence was too low at the time births of the reporting children occurred, and those infected in adulthood have not yet started experiencing increased mortality (due to the long incubation period of HIV). The epidemic must have reached a level of prevalence and maturity such that the impact on the reporting of the mortality of parents is noticeable. This will depend on the scale of the epidemic and the pace at which prevalence increases, but in most cases the impact will probably become significant between 20 and 35 years after the epidemic exceeds a prevalence of one percent. The orphanhood method, using data from a single survey, estimates mortality between 7 and 25 years before the date of the survey. When using time location, the most recent estimates (based on responses from the very young) are often more questionable than estimates relating to the more distant past (based on responses from

older children). This is a consequence of the “adoption effect”, where adopted parents of younger children are either reported as biological parents in the survey, or are assumed by survey officials to be biological parents. Thus the most reliable mortality estimates flowing from use of the orphanhood method probably relate to time locations more than seven years before the date of the survey.

A study by Hosegood *et al* (2004) examined data gathered in the rural district of Umkhanyakude in northern KwaZulu Natal. Two sets of data were collected: demographic and health data from 11 314 households (89 132 individuals) collected every four months from January 2000 to December 2000; and information from each of the participants in the survey on the survival of biological parents. The data on orphanhood are influenced only to a very limited extent by HIV as few young respondents interviewed will have been born at a time when infection levels were above a few percentage points. The results from this study are plausible – the data showing lower mortality rates in earlier years and indicating a rapid rise in mortality in the late 1990’s.

Timaeus and Jasseh (2004) used data on sibling survival from Demographic and Health Surveys from a wide variety of countries to assess mortality trends in a large number of African countries. The DHS’s used were all conducted between 1992 and 2000. Two sets of data were used to calculate mortality in the investigation – data on survival of siblings and data on orphanhood. Mortality estimates were derived from orphanhood data using a slightly altered version of that the revised method proposed by Timaeus and Nunn (1997). Mortality estimates derived using sibling survival data appear to have been done without any adjustment of the method to deal with the effects of HIV. The authors concluded that there are no consistent differences between the mortality estimates derived using the two methods, and that there is no tendency for estimates to diverge at higher levels of mortality (where HIV could be having a greater impact). This was taken to indicate that neither method produces biased estimates in the presence of an HIV epidemic. Hill and Trussel (1977) indicated (at a time when HIV was not yet known about) that the sibling survival method might be expected to suffer from similar sources of potential bias as the orphanhood method, in that correlations between the mortality of siblings are likely to exist. The suggested correlation between the mortality of siblings (i.e. sibling survival method) is likely to be greater in a context of high HIV prevalence, even though increased correlation is probably substantially smaller than that of mother and offspring (i.e. affecting the orphanhood method). The impact of this increased correlation would be a smaller relative rise in the estimate of mortality

using the sibling survival method as compared to a similar estimate using the orphanhood method, given an underlying rise in mortality due to increased HIV prevalence or a maturing HIV epidemic.

The full set of data is not given in the paper, but identification of graphically presented points is possible for key countries. The highest levels of mortality, as estimated by the orphanhood method and excluding Rwanda are, in order from highest to lowest: Uganda in 1991, Uganda in 1996; Zimbabwe in 1995 and Zambia in 1992. Of these the only estimate which might be expected to be subject to substantial HIV-related bias when using the orphanhood method, given the scale and duration of the epidemic, is Uganda in 1996. The estimate of mortality for Uganda in 1996, based on the orphanhood method as revised by Timaeus and Nunn (1997) shows a decrease in mortality, while the estimate derived using the sibling survival method indicates a substantial increase.

2.5 Conclusion of Literature Review

Many sources of bias that have an impact on the orphanhood method have been studied and analysed. A generalized HIV epidemic, however, introduces violations of the assumptions of the method at a level which was not foreseen or dealt with in any of these studies. The manner in which various biases described by Palloni *et al.* (1984) cancel each other out, or reinforce each other, may be quite different in the presence of a high-prevalence generalized HIV epidemic.

Timaeus and Nunn (1997) have produced plausible and practical adjustment techniques to allow for the areas of bias that have the greatest impact, but have concentrated on a population with relatively low prevalence of HIV. The prevalence amongst pregnant women in the populations studied by Timaeus and Nunn (1997) was about 12 percent. In many sub-Saharan countries today prevalence amongst pregnant women is higher than 30 percent. A degree of error is also possible in light of the data they used to generate revised coefficients. The very different mortality patterns of cohorts of women of child-bearing age and women aged 25 at the time of birth of their children, would cause errors.

The effect of a generalized high prevalence HIV epidemic on the accuracy of the orphanhood method requires further investigation. Biases resulting from correlated mortality between mother and child, the relationship between HIV and fertility, and the effects of HIV on age-specific mortality rates require

investigation. The significant problem of the conversion of $l_{(25+x)}/l_{(25)}$ into other life table measures needs to be examined, but is not be dealt with in this research.

University of Cape Town

Chapter 3: Methodology

3.1 Outline of method and rationale

The orphanhood method essentially concerns itself with two measures of survival. The first measure is the survival, from the time of birth of their children, of women of a wide range of ages who become mothers. The question asked in a census or survey about survival of mothers seeks to establish the survival of this group. The second measure is the survival from age 25 to subsequent ages ($l_{(25+x)}/l_{(25)}$). The orphanhood method adjusts the proportion of mothers alive to estimate $l_{(25+x)}/l_{(25)}$, using factors derived from models which seek to predict the deviation of the one from the other (using age-specific fertility rates, model life tables, etc). In circumstances where model life tables and age-specific fertility rates reflect actual experience reasonably, and reported survival of mothers reflects survival of women in general, these factors are such that the method could be used to approximate $l_{(25+x)}/l_{(25)}$ for women in general. This may not be true however, if fertility or mortality rates change rapidly, or if age-specific fertility or mortality rates have changed in different directions. Transition to a generalized high prevalence HIV epidemic typically causes these sorts of rapid changes and age-specific variations.

Where the mortality of women increases gradually and consistently with age, survival of mothers can be used to approximate survival of those near the mean age of mothers, with some adjustment. However, when the risk of mortality no longer increases with age at all ages, there is potential for this relationship to break down. In a situation where HIV is widespread the mortality of the cohort of women aged 25 might increase far more rapidly in the years that follow than the mortality of women across a wider age range, and might then decrease after the age of peak HIV mortality has been passed (while the mortality of the women age 15-49 is still increasing). In addition, mortality as reported by children may not be a true reflection of the mortality of women of child-bearing age, due to mother-to-child transmission and the relationship between HIV and fertility. A number of potential areas of error require investigation when a generalized HIV epidemic is present. The various areas of investigation are outlined below.

First, potential errors due to the relationship between fertility by age and HIV infection, are dealt with. Within each age group of women, the proportion HIV positive is different for women in general as compared to women who become mothers (i.e. give birth in the year in question). At younger ages sexual activity tends to be related both to higher fertility as well as higher HIV incidence – HIV status and fertility are thus highly correlated. At older ages, where more women are in the later stages of HIV infection, HIV infection is associated with lower fertility (Gregson *et al.* 2002). The effect of this is investigated by comparing the simulated survival of all women, age weighted to represent mothers, with simulated survival of those women who actually become mothers (i.e. each age group is of identical size, but proportions HIV positive and HIV positive by duration since infection differ).

Second, potential errors arising from the correlations of mortality between mothers and their children are identified. Of the women who give birth, some will transmit HIV to their children and some will not. There is a relationship between the survival of infected children and the survival of their mothers, due to vertical transmission. The children who are less likely to survive have mothers who are less likely to survive. The reporting by children on the survival of their mothers is thus a biased measure of survival due to this relationship. The effect is investigated by comparing a simulation of the reported survival of mothers, assuming accurate reporting (i.e. dependent only on the survival of children who report on their mothers' survival), with a simulation of survival of mothers.

Third, the age profile of mortality is known to change dramatically during a generalized HIV epidemic. Women in their late twenties and early thirties typically experience mortality far higher than women a few years younger or a few years older. The orphanhood method seeks to use the survival of mothers (as reported by children) as an estimator for the survival of 25 year olds, but 25 year olds are likely to experience (immediately or after they have aged a few years) higher mortality than women across the age range 15 to 49. For comparability an age profile weighted identically to that of mothers is used when modelling survival of women. The extent of this error is investigated by comparing the simulated survival (and non-survival) of 25 year old women with the simulated survival of women across the age range of child-bearing, age-weighted identically to women who give birth.

By carrying out the three comparisons above, it is possible, within the limitations of modelling and simulation, to assess the likely extent of each of the three main factors affecting the correctness or otherwise of

reported survival rates. The final analysis of the how accurate reported survival rates estimated by the orphanhood method are likely to be in estimating $l_{(25+x)}/l_{(25)}$ involves combining all three effects. This is the total error likely to be experienced when using the method.

Five steps are used in doing the comparisons outlined:

- Survival for each of the groups (women aged 25, women, mothers, mothers as reported by children) and ages is calculated. This is done by birth-year of child (Y), age of mother at birth of child (a), and age of child (N).
- Survival of each of the groups is calculated, aggregated across age of mother at birth of child (a) (using the same age-weightings as for mothers).
- Survival of non-aggregated and aggregated groups is compared, in cohort and period format, to provide an understanding of the age-specific effects, as well as effects that occur across ages (of mothers).
- Survival aggregated by children's age (N) in 5-year bands for periods is then compared. This simulates the format of the data required by the method.
- The method is applied to the point where the relevant correction factors are used to convert survival into $l_{(25+x)}/l_{(25)}$. This is done using simulated period data relevant to a range of years.

A further potential error may occur in using the method, but is not investigated in this thesis. $l_{(25+x)}/l_{(25)}$ generally requires conversion into other life table measures in order to be of wider use, especially to estimate changes in mortality over time. To do this conversion one needs standard life tables which reasonably approximate the shape and scale of mortality patterns in the population being studied. Life tables which take into account age-specific HIV related mortality at the relevant reference time point are needed to ensure appropriate conversions in populations with high-prevalence generalized HIV epidemics. Currently the standard life tables being used do not reflect HIV-related age-specific mortality patterns, although it is possible that use of the INDEPTH model life tables (INDEPTH network 2004) could resolve this problem. Certain practical difficulties may arise in using the INDEPTH model tables however. Apart from any shortcomings in the INDEPTH tables, such as underestimation of the mortality at older ages, it is not possible with the use of only two parameters, one for level (α) and one for shape (β) to reproduce all the mortality curves likely

to be experienced over the course of an epidemic. The INDEPTH tables were developed using data from some populations with high HIV prevalence, but the prevalence levels were still substantially below that of some of the countries where we may wish to apply the orphanhood methods. Although these practical issues may not be insurmountable, this research does not investigate the use of the INDEPTH tables, focussing as it does on the accuracy of the method as it is currently described in the literature. The problem of how to convert $l_{(25+x)}/l_{(25)}$ into any other common life table index of mortality is not considered.

3.2 Description of the ASSA2002 model and how it is used

The ASSA2002 AIDS and Demographic Model (ASSA 2005) is a mathematical model developed by the Actuarial Society of South Africa. The model is described in more detail by Johnson and Dorrington (2006), and can be downloaded from www.assa.org.za. The model seeks to reproduce as accurately as possible the current and past patterns of HIV spread and HIV-related demographic impact in South Africa, in addition to reproducing non-AIDS related mortality, fertility and migration. The 'full' version of the model is disaggregated into four race groups, and projects the male and female demographic components by individual age and year. Each of these groups (race, age, sex) is further divided, for modelling purposes, into four risk categories. These are, in order of risk of infection with HIV, "PRO" (sex worker or other highly sexually active individuals with many partners and regularly infected with sexually transmitted infections), "STD" (those with multiple partners, but fewer than people in the "PRO" group, and at heightened risk of sexually transmitted infections), "RSK" (one new partner per year, and limited sexually transmitted infections), and "NOT" (those either not sexually active or sexually active in only one monogamous relationship). In the model the HIV epidemic begins in 1985, with an assumed number of 'imported' infections, and spreads most readily in the riskier groups first, gradually moving to the less risky groups and becoming generalized. The amount of sexual activity and distribution of sexual partners by age, for males and females at each age, is mathematically modelled to reflect reality as closely as possible. In addition fertility patterns are captured by making both pregnancy and infection functions of the amount of sexual activity, and by allowing for the reductions that long duration infection cause in fertility levels.

The ASSA2002 model has five HIV-related interventions integrated into the model. The effects of these interventions are subject to various assumptions. Information and education campaigns, voluntary counselling and testing, prevention of mother-to-child transmission, antiretroviral therapy and treatment of sexually transmitted infections are included in the model. These interventions can be switched on or off, or be enhanced or reduced, depending on the purpose of the outcomes. The default assumptions are used to produce a “best-estimate” outcome, while other variations would be used to understand outcomes if interventions were enhanced or reduced (i.e. sensitivity testing).

The outputs are generated by adapting the model to produce additional output, which exists internally within the model for each year of calculation. The data needed for this research include the number of women of each age, in every year, by duration of HIV infection, as well as age-specific fertility rates for both HIV-negative and HIV-positive women (for each duration of infection). Section 3.3 describes how these data are used to calculate the errors and biases that this research seeks to identify and quantify. The processing of data in accordance with the mathematical functions described in section 3.3 was done using Matlab® software (2006b).

The research presented here uses outputs generated using the model with anti-retroviral treatment and mother-to-child transmission prevention (i.e. pharmaceutical interventions) removed, but allowing for non-pharmaceutical interventions such as information and education campaigns, and voluntary counselling and testing.

Only data related to Black/African South African women is used. This particular group is chosen for this study for a number of reasons. Choosing only females (and female off-spring) simplifies considerably the mathematical modelling required to assess the impact of HIV on biases in the orphanhood method. While the detailed results would be slightly different if adult males were the subject of the study, the broad outcomes and conclusions would be no different. Modelling using male off-spring would produce almost identical outcomes. Black South African females were also chosen as this group is experiencing a severe HIV epidemic which has been relatively well documented and recorded. HIV prevalence rates for pregnant women have been readily available for a number of years and the timing of increases in HIV infection and age-specific differences are known. The effect of the HIV epidemic on Black South African women is reasonably well represented by public antenatal clinic data (which is likely not the case for other ethnic groups).

Having chosen Black South African women as the population to be studied in this research, it is worth noting some of the relevant socio-demographic characteristics of this group. The total fertility rate of this group is estimated to have been 3.04 in 2001 and 3.49 in 1996 (Moultrie and Dorrington 2004). The 2007 Community Survey indicates that total fertility continued to fall, to 2.7 children per woman by that year (Community Survey 2007). In 2007, 12.8% of Black South Africans aged over 20 (combined male and female) had had no schooling, and 18.8% had only some primary schooling. Around 47% had completed primary education or had some secondary education, while 15.4% had completed high school and 5.6% had tertiary education (Community Survey 2007). The following table shows HIV prevalence amongst antenatal clinic attendees in South Africa (Dept. Health 2009). Approximately 90% of attendees are Black – the data are thus not an exact representation of prevalence amongst Black South African women, but present a reasonable view of the level and timing of HIV infection in this group.

Antenatal clinic HIV prevalence - South Africa			
1990	0.7%	2000	24.5%
1991	1.7%	2001	24.8%
1992	2.2%	2002	26.5%
1993	4.0%	2003	27.9%
1994	7.6%	2004	29.5%
1995	10.4%	2005	30.2%
1996	14.2%	2006	29.1%
1997	17.0%	2007	29.4%
1998	22.8%	2008	29.3%
1999	22.4%		

Table 1: Antenatal clinic HIV prevalence - South Africa

3.3 Assumptions within the ASSA2002 model

The assumptions of the ASSA model which are relevant to this research are described below.

The transmission of HIV from infected mothers to children (without PMTCT) is assumed to occur in about 33% of cases. Perinatal transmission is assumed to occur in 20% of births and transmission via breastmilk is assumed to occur in 16% of births which are not perinatally infected. Use of antiretroviral drugs

is assumed for our purposes not to occur – neither for the purposes of preventing mother-to-child transmission or for other treatment purposes.

The median term to death for Black Females aged 14 to 24 at infection is assumed to be 11 and a half years. For those aged 25 to 34 this is assumed to be 10 and a half years, and for those aged 35 or older the median term to death is assumed to be 9 years. For children infected perinatally, median term to death is 1.14 years, and for those infected via breast-milk the median survival term is 8.95 years.

The ASSA2002 model contains a number of assumptions about fertility by duration of infection and age. Women infected for longer durations are assumed to be less fertile, and at younger ages the model assumes a strong non-causal positive relationship between HIV infection and fertility (both being the result, in most cases, of unprotected sexual activity). Table 2 shows the ratio of fertility of infected women to uninfected women for selected ages and years. Infected women under 28 are on average assumed in the model to have more children than uninfected women, and vice versa for older ages.

Age	Year				
	1990	1995	2000	2005	2010
15	6.20	6.23	6.29	6.28	6.07
16	3.64	3.68	3.75	3.76	3.74
17	2.48	2.52	2.60	2.61	2.60
18	1.88	1.92	1.99	2.00	2.00
19	1.54	1.57	1.64	1.65	1.65
20	1.33	1.36	1.43	1.44	1.44
25	1.02	1.04	1.07	1.07	1.07
30	0.99	0.99	1.00	0.95	0.93
35	0.97	0.97	0.97	0.91	0.85
40	0.96	0.95	0.94	0.89	0.82
45	0.96	0.94	0.91	0.88	0.81

Table 2: Ratio of fertility of HIV-positive women to fertility of HIV-negative women, by year

3.4 Survival of women aged 25, women, mothers, and mothers as reported by children

In this model we consider only the survival of mothers and of their female off-spring. All survival and transition probabilities are thus for females. All data used as in these simulations are derived from the ASSA2002 version of the ASSA Aids and Demographic models (ASSA 2005).

Let:

$I_{y,a,d}$ = the number of infected women aged a last birthday, in the middle of year y , who have been infected for curvate duration d years

$U_{y,a}$ = the number of uninfected women aged a last birthday, in the middle of year y

$f_{y,a}^-$ = fertility rate of uninfected women aged a last birthday for year y , i.e. from time y to time $y+1$

$f_{y,a,d}^+$ = fertility rate of infected women aged a last birthday, who have been infected for curvate duration d years, for year y .

τ = the rate of vertical transmission of HIV from mother-to-child at birth or during breast-feeding

$S_{y,25,N}^{25}$ = proportion of all women aged 25 in the middle of year y who survive N years

$S_{y,a,N}^W$ = proportion of all women aged a at the middle of year y who survive N years

$S_{y,a,N}^M$ = proportion of mothers aged a at the middle of year y who survive N years

$S_{y,a,N}^R$ = proportion of mothers aged a at the middle of year y reported as having survived for N years by their surviving daughters

Also let:

$S_{y,N}^W$ = proportion of women (with age weighting identical to that of mothers) at the middle of year y who survive N years

$S_{y,N}^M$ = proportion of mothers at the middle of year y who survive N years

$S_{y,N}^R$ = proportion of mothers at the middle of year y reported as having survived for N years by their surviving daughters

Assuming:

- a constant sex ratio at birth
- $I_{y,a,d} = 0$ if $d > a$
- reporting by daughters is accurate and complete

$S_{y,25,N}^{25}$, $S_{y,a,N}^W$, $S_{y,a,N}^M$ and $S_{y,a,N}^R$ are then calculated as follows:

The survival of 25-year olds over a period of N years is equal to those alive aged $(25+N)$ at time N , whether infected or uninfected, regardless of duration.

$$S_{y,25,N}^{25} = \frac{U_{y+N,25+N} + \sum_{d=0} I_{y+N,25+N,d}}{U_{y,25} + \sum_{d=0} I_{y,25,d}}$$

The survival of women is similarly calculated, but with all ages considered.

$$S_{y,a,N}^W = \frac{U_{y+N,a+N} + \sum_{d=0} I_{y+N,a+N,d}}{U_{y,a} + \sum_{d=0} I_{y,a,d}}$$

The survival of mothers is similarly calculated, but requires weighting of those that were HIV-negative or HIV-positive at the start (using fertility by infection status and duration of HIV infection) to match survival by HIV status. All women after N years that are uninfected, or with HIV infection of duration less than N were clearly HIV-negative at the start, and are weighted using the fertility of HIV-negative women. All others were HIV-positive, and are weighted using duration-specific fertility of HIV-positive women.

$$S_{y,a,N}^M = \frac{\left(U_{y+N,a+N} + \sum_{d=0}^{N-1} I_{y+N,a+N,d} \right) f_{y,a}^- + \sum_{d=0} \left(I_{y+N,a+N,d+N} f_{y,a,d}^+ \right)}{U_{y,a} f_{y,a}^- + \sum_{d=0} \left(I_{y,a,d} f_{y,a,d}^+ \right)}$$

The calculation of survival as reported by children is affected by whether children are infected with HIV at birth or not. The rate of vertical transmission of HIV, τ , is thus introduced. Uninfected children include all those born to HIV-negative mothers, as well as $(1-\tau)$ children of HIV-positive women. The weighting is applied to both the number of women after N years (numerator) and the number in the birth year (denominator). In the function below the first term multiplies the proportion of children uninfected with the reported survival of mothers of these children (uninfected mothers and infected mothers who did not transmit HIV to their children). The second term does the same with the proportion of children infected and mothers infected who did transmit HIV to their children. The proportions of children HIV-negative or HIV-positive are, in essence, used to weight the survival probabilities of their respective cohorts of mothers.

$$S_{y,a,N}^R = \left(\frac{A_{y,a,N}}{A_{y,a,N} + B_{y,a,N}} \right) \times \left[\frac{\left(U_{y+N,a+N} + \sum_{d=0}^{N-1} I_{y+N,a+N,d} \right) f_{y,a}^- + (1-\tau) \sum_{d=0}^{N-1} \left(I_{y+N,a+N,d+N} f_{y,a,d}^+ \right)}{U_{y,a} f_{y,a}^- + (1-\tau) \sum_{d=0}^{N-1} \left(I_{y,a,d} f_{y,a,d}^+ \right)} \right] \\ + \left(\frac{B_{y,a,N}}{A_{y,a,N} + B_{y,a,N}} \right) \times \left[\frac{\sum_{d=0}^{N-1} \left(I_{y+N,a+N,d+N} f_{y,a,d}^+ \right)}{\sum_{d=0}^{N-1} \left(I_{y,a,d} f_{y,a,d}^+ \right)} \right]$$

where:

$A_{y,a,N}$ is the number of uninfected children born to women aged a in year y , still alive N years

later, i.e.

$$A_{y,a,N} \approx 0.5 \times \left(U_{y+N,N} + \sum_{d=0}^{N-1} I_{y+N,N,d} + U_{y+N,N-1} + \sum_{d=0}^{N-1} I_{y+N,N-1,d-1} \right) \times \frac{U_{y,a} f_{y,a}^- + (1-\tau) \sum_{d=0}^{N-1} \left(I_{y,a,d} f_{y,a,d}^+ \right)}{\sum_{a=15}^{49} \left(U_{y,a} f_{y,a}^- + (1-\tau) \sum_{d=0}^{N-1} \left(I_{y,a,d} f_{y,a,d}^+ \right) \right)}$$

and

$B_{y,a,N}$ is the number of infected children born to women aged a in year y , still alive N years later,
i.e.

$$B_{y,a,N} \approx 0.5 \times (I_{y+N,N,N} + I_{y+N,N-1,N-1}) \times \frac{\sum_{d=0}^{49} (I_{y,a,d} f_{y,a,d}^{*+})}{\sum_{a=15}^{49} \sum_{d=0}^{49} (I_{y,a,d} f_{y,a,d}^{*+})}$$

The numbers of females infected or uninfected can be derived directly from the ASSA2002 model (ASSA 2005), provided we assume that $I_{y,a,d} = 0$ if $d > 30$. The fertility rates for the year y , $f_{y,a}^{-}$ and $f_{y,a,d}^{*+}$, are approximated by averaging the fertility rates from the model for age a and duration d for the years $y-1$ and y , since the flows (births, deaths, etc) produced by the ASSA2002 model (ASSA 2005) for the year labelled y are for the year starting in the middle of the year to the middle of the next year.

The accurate measure of survival is given by $S_{y,a,N}^{\#}$, which approximates $NP_{a+\frac{1}{2}}$ (for the period $y+\frac{1}{2}$ to $y+N+\frac{1}{2}$).

The orphanhood method uses survival of mothers as reported by surviving children aggregated into 5-year age bands. The survival rates need to be expressed in period format in 5 year age bands. Let G denote each 5-year group, such that $G = \text{Integer}(N/5)$ and let calendar year be denoted c .

Survival for women aged 25 at the birth of their child, women, mothers and mothers as reported by children, aggregated into 5-year bands in calendar year c , is calculated as follows:

$$S_{c,G}^{25} = \frac{\sum_{N=0}^4 (S_{c-5G-N,5G+N}^{\#} \cdot W_{c-5G-N}^{25})}{\sum_{N=0}^4 W_{c-5G-N}^{25}}$$

where

$$W_{c-5G-N}^{25} = U_{c-5G-N,25} f_{c-5G-N,25}^- + \sum_d I_{c-5G-N,25,d} f_{c-5G-N,25,d}^+$$

and similarly

$$S_{c,G}^W = \frac{\sum_{N=0}^4 (S_{c-5G-N,5G+N}^W \cdot W_{c-5G-N})}{\sum_{N=0}^4 W_{c-5G-N}},$$

$$S_{c,G}^M = \frac{\sum_{N=0}^4 (S_{c-5G-N,5G+N}^M \cdot W_{c-5G-N})}{\sum_{N=0}^4 W_{c-5G-N}} \quad \text{and}$$

$$S_{c,G}^R = \frac{\sum_{N=0}^4 (S_{c-5G-N,5G+N}^R \cdot W_{c-5G-N})}{\sum_{N=0}^4 W_{c-5G-N}}$$

where

$$W_{c-5G-N} = \sum_{a=15}^{49} \left(U_{c-5G-N,a} f_{c-5G-N,a}^- + \sum_d I_{c-5G-N,a,d} f_{c-5G-N,a,d}^+ \right)$$

$S_{c,G}^M$ for example, is thus the proportion of mothers who have survived between 5G and 5G+4 years after the birth of their child, to be alive at the middle of year c.

3.5 Comparison of survival and non-survival: mothers as reported by children vs. mothers; mothers vs. women; and women vs. women aged 25

The comparisons are divided into three sections initially: the first dealing with those effects that impact the data that are used in the orphanhood method, namely the effects of HIV on fertility, and the effects of HIV

that result in a correlation in mortality between mother and daughter. The second deals with the changes in the age patterns of mortality and how well future survival of a cohort of women across the child-bearing age range can represent the future survival of women aged 25 - the method has coefficients which relate the one to the other, and this comparison is therefore treated separately. Lastly, a comparison will be made combining all these effects – showing how well reported survival of mothers (adjusted using the coefficients in the method) is likely to estimate $l_{(25+x)}/l_{(25)}$, in various years and at various reporting ages.

In the first section the following comparisons are done:

$\frac{S_{y,a,N}^M}{S_{y,a,N}^W}$ indicates (for each inception (birth) year and age of mother) the extent to which survival is

affected by the differences in HIV prevalence (and durations of HIV infection) between mothers and women in general. A similar ratio is calculated for non-survival (cumulative mortality), which is especially useful in demonstrating differences in the earlier years when both the denominator and numerator are close to 1.

Comparisons of $\frac{S_{y,N}^M}{S_{y,N}^W}$ (i.e. aggregated across ages), and $\frac{S_{y+N,N}^M}{S_{y+N,N}^W}$ (in period format) are also shown. $\frac{S_{c,G}^M}{S_{c,G}^W}$

(i.e. N aggregated into 5-year bands and in period format) shows the magnitude of error using data which replicates that used in practice in the method.

$\frac{S_{y,a,N}^R}{S_{y,a,N}^M}$ indicates (for each inception (birth) year and age of mother) the extent to which reported

survival is affected by the relationship between survival of mothers and survival of children during a generalized HIV epidemic. As for the ratio of survival of mothers to women, comparisons are done using survival rates aggregated across ages (of mothers), arranged in period format, and with N aggregated into 5-year bands.

$\frac{S_{y,a,N}^R}{S_{y,a,N}^W}$ indicates (for each inception (birth) year and age of mother) the extent to which reported

survival is affected by both the relationship between survival of mothers and survival of children during a generalized HIV epidemic, and by the differences in HIV prevalence between mothers and women in general. As for other similar ratios, comparisons are done using survival rates aggregated across ages (of mothers), arranged in period format, and with N aggregated into 5-year bands.

In the second section the following comparison is undertaken, with data aggregated into appropriate 5-year bands, both with and without adjustment using the coefficients of the method that transform reported survival into $l_{(25+x)}/l_{(25)}$:

$\frac{S_{y,N}^W}{S_{y,25,N}^W}$ indicates (for each inception (birth) year and age of women) the extent to which a cohort of

women with an age profile matching that of mothers experiences survival which is different to that of women that this cohort of surviving mothers is taken (with some adjustment) to represent (i.e. 25 year olds). A similar

ratio can be calculated for non-survival, and $\frac{S_{y+N,N}^W}{S_{y+N,25,N}^W}$, and $\frac{S_{c,G}^W}{S_{c,G}^{25}}$ are also shown. In addition,

$\frac{f(S_{c,G}^W, a_0, a_1, a_2)}{S_{c,G}^{25}}$ is shown, where a_0 , a_1 , and a_2 are the coefficients used to convert survival into

$l_{(25+x)}/l_{(25)}$

In the third section $\frac{S_{c,G}^R}{S_{c,G}^{25}}$ and $\frac{f(S_{c,G}^R, a_0, a_1, a_2)}{S_{c,G}^{25}}$ are shown to demonstrate the total likely error in

using reported survival of mothers to estimate $l_{(25+x)}/l_{(25)}$ (both with and without application of the

coefficients). $f(S_{c,G}^R, a_0, a_1, a_2)$ is calculated using both a set of standard coefficients and those proposed by Timaeus and Nunn (1997).

University of Cape Town

Chapter 4: Results

4.1 Comparison of survival and non-survival: women, mothers, and mothers as reported by children

4.1.1 Individual age-specific cohorts of women and mothers in a particular year

Three cohorts of women, and women who became mothers in the year 2005, aged 15, 25 and 35 at that time, have been chosen to illustrate the impact of HIV on their survival.

The age distributions of women who give birth and of all women of childbearing age are very different (Figure 1). The impact of this difference on the proportion surviving of the cohort aged 15-49 (in aggregate) is minimal since the weighted average age of the two groups of women (i.e. mothers vs. all women), weighted by proportion surviving at each age, are fairly similar. In 2005 mothers are aged 26.6 on average, whereas women are on average aged 29.4. In order to facilitate more precise comparisons, however, cohorts of women referred to in this research have been weighted identically to mothers at the start year (i.e. year of birth for mothers). Although the age-weighting is identical for women and mothers, the proportions HIV-positive (and HIV-positive by duration) in each age group will be different. In this section the research examines survival of single-age groups of women and mothers in order to identify and understand how age groups are affected differently by the various sources of error.

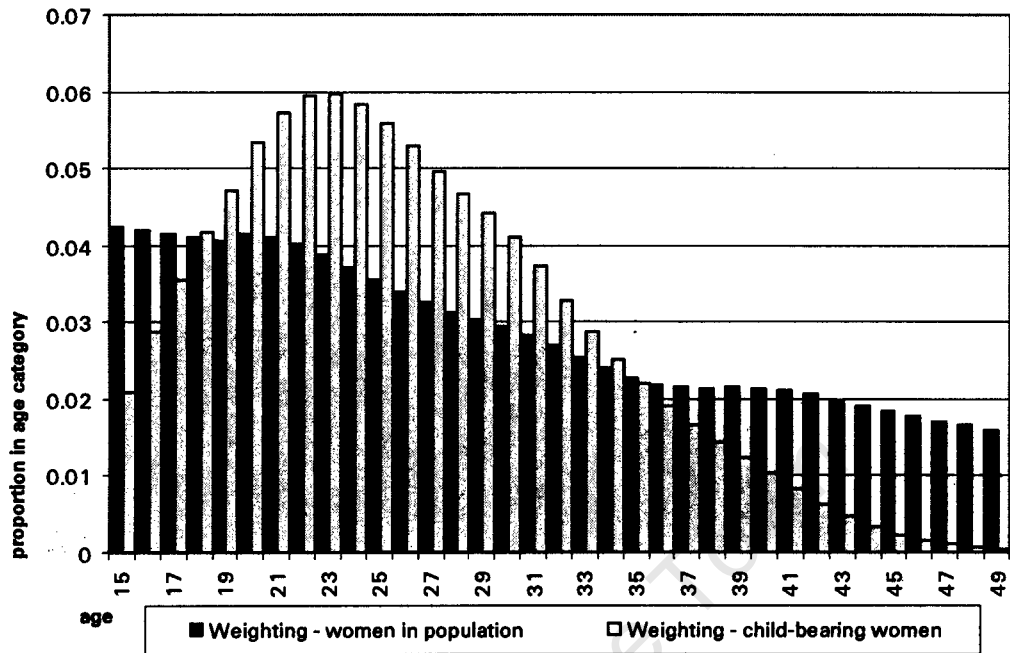


Figure 1: Age profile of women vs. mothers in 2005 (from the ASSA2002 model)

Looking at individual ages, for the most part different survival patterns are exhibited by mothers as compared to all women at the same age, in the presence of a generalised HIV epidemic. At younger ages, newly infected women tend to be those that have higher levels of sexual activity, and hence are more likely to become mothers, than other women (i.e. the uninfected) of the same age. Those who are older and tend to be at later stages of HIV disease progression are likely to experience lower fertility due to the disease. At the youngest child-bearing ages, where very few are infected before giving birth, the overall effect of the difference between the cohort of mothers and women is small, despite the higher fertility of the HIV-positive women. Figure 2 shows the proportion uninfected and infected by duration, for ages 20 and 30 in 2005. A higher proportion infected is found amongst mothers than amongst women in general at age 20, whereas the opposite is true for age 30 (where those who have been infected for longer are assumed to have lower fertility).

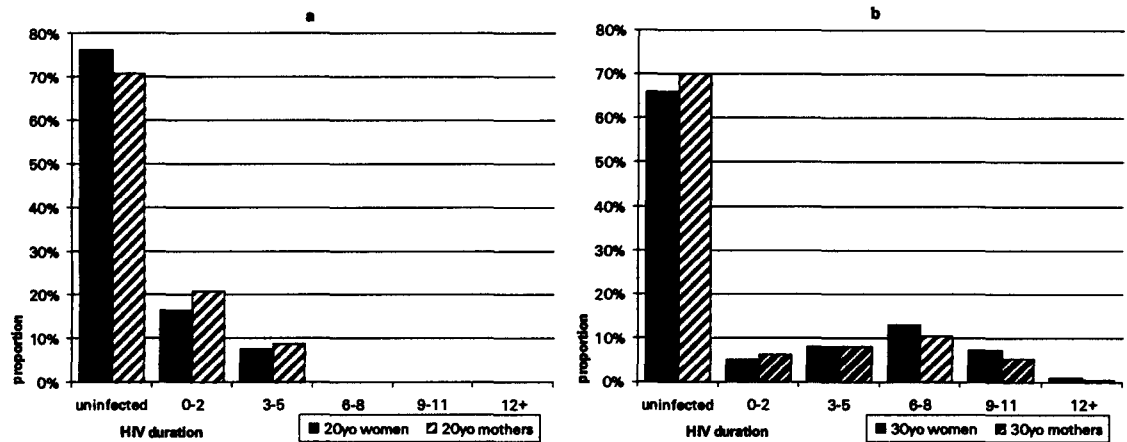


Figure 2: Proportion uninfected, or infected for various durations - women and mothers giving birth in 2005 (a) aged 20 and (b) aged 30

In addition to the effect of HIV on fertility, reporting of survival of mothers is affected by the numbers of children infected through mother-to-child transmission. The children most likely to die soon after birth will have been born to women most likely to die in the next 10 to 20 years. This relationship will result in under-reporting of the mothers' deaths if the reporting is done by their children, as is the case in the orphanhood method.

Figure 3 shows survival of cohorts of those who were 15, 25 and 35 years old in 2005. The survival of women (age-weighted identically to mothers) and mothers, and the reported survival of mothers (reported by the surviving children born in 2005), are shown. From this we note that there is very little difference between the survival of all mothers and all women when looked at in aggregate across all ages, but that for older ages the difference is large and for younger ages the difference is smaller and in the opposite direction. However, at all ages except the youngest the survival of mothers as reported by their surviving children substantially overestimates the survival of women.

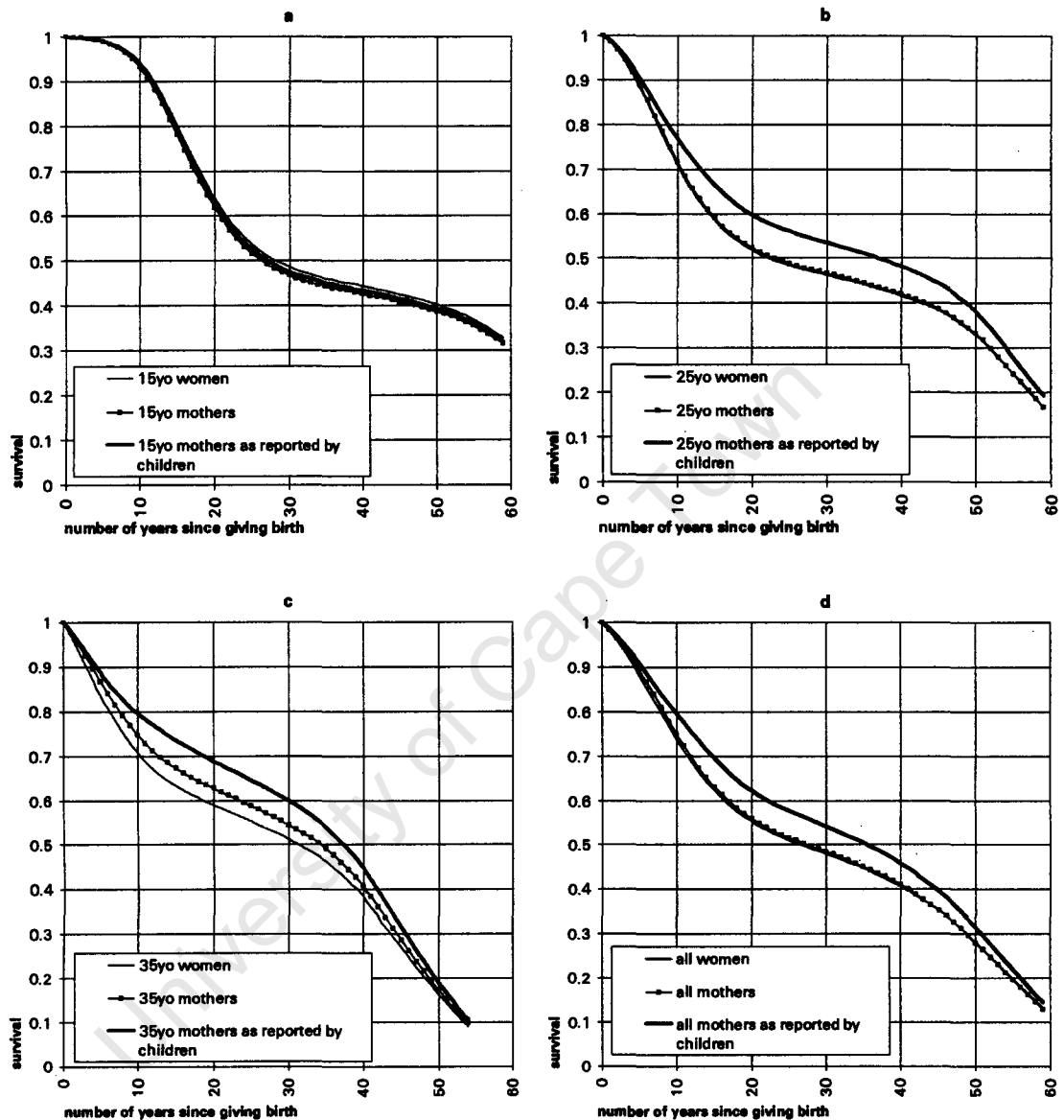


Figure 3: Survival of women, mothers and mothers as reported by children, by duration since birth in 2005: (a) ages 15, (b) ages 25, (c) ages 35, and (d) all ages combined

Figure 4 shows a slightly different representation of the same data. Instead of survival the focus is on cumulative mortality (or non-survival), and all figures given are ratios comparing the three categories studied (women, mothers, and mothers as reported by children). The figure compares the ratios of the proportions not

surviving (e.g. the ratio of the proportion of mothers not surviving to that of all women). The proportion of mothers not surviving compared to that for all women, shows the effect of HIV on fertility (i.e. on the profile of mothers by age and duration since infection). The reported non-survival of mothers compared to actual non-survival of mothers, shows the effect of correlations between the mortality of children and the mortality of their mothers. The solid lines in Figure 4 show the combined bias due to HIV. From this we see that, consistent with the pattern of survival in Figure 3, the ratio of the proportion of all mothers not surviving to the proportion of all women not surviving is close to one after about 15 years. For the older women the difference is larger, and mortality is underestimated. For younger women mortality is overestimated. When we add in the effect of HIV on reporting of mothers deaths, however, significant under-reporting of mortality is shown to occur at all ages, except the youngest.

University of Cape Town

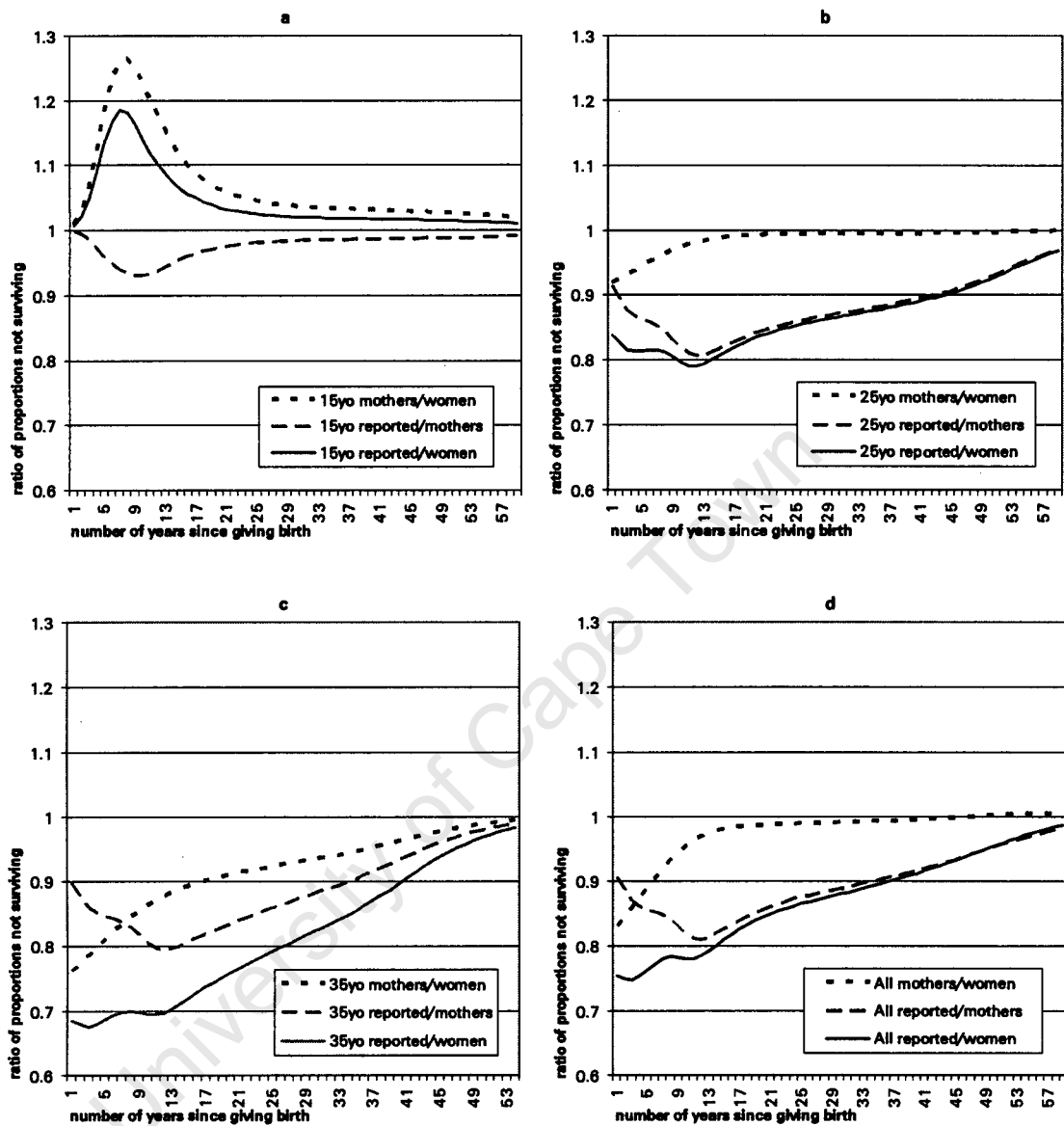


Figure 4: Ratio of proportion not surviving of mothers vs. women, reported vs. mothers and mothers as reported by children vs. women, by duration since birth in 2005: (a) ages 15, (b) ages 25, (c) ages 35, and (d) all ages combined

Women and mothers at age 15

At age 15 HIV-positive women exhibit a higher fertility rate than HIV negative women (the model assumes this to be 50% higher), as they are by definition sexually active, but the proportion of this age group that are

HIV-positive is very small and those that are infected are newly infected. Women who became mothers at age 15 in 2005 thus experience a higher rate of mortality than 15 year old women in general. The difference in cumulative non-survival reaches a peak at about 8 years after which it falls.

Most women in this cohort who will ultimately become infected with HIV have yet to be infected – the steep drop in survival (see Figure 3) after about 14 years is a reflection of the high level of new infections that occur during young adulthood, affecting both the mothers and women in general. The resultant high mortality between ages 25 and 35 causes survival to drop temporarily below that of women who were older when they gave birth in 2005.

The reporting of survival of mothers by their children appears to be only slightly biased when looking at Figure 3, but is more pronounced when examined as a ratio of the proportion not surviving to that of all women of that age (i.e. shown in Figure 4). The effect peaks at about 9 years, after which the bias is reduced by the effect of elevated HIV-related deaths on the cohort in general (i.e. the effect of the correlation between deaths of mothers and children becomes smaller in relative terms due to the diluting effect of high levels of unrelated mortality).

The combined effect of both fertility effects and correlation of mortality is shown by the solid line. The relationship between the mortality of mothers and their children results in an over-statement by reporting children of the proportion not surviving (i.e. survival of women is under-estimated).

Women and mothers at age 25

At age 25, HIV-positive women have been infected for a range of durations. A slightly higher proportion of mothers are HIV-negative than women in general (an estimated 64.3% as compared to 62.9% respectively, from the model). The net effect is a slightly lower proportion not surviving amongst mothers than amongst women in general, especially in the first few years – this age cohort is already experiencing elevated mortality due to HIV, as shown by the steepness of the drop in survival in the first few years (Figure 3).

The relatively high number infected at the start means that many children of these women are infected. This results in the proportion of children alive to report on the survival status of their mothers rapidly becoming biased in favour of those born to HIV-negative mothers. This in turn results in substantial under-

statement of the cumulative non-survival of the mothers after a few years, when reported by children. On reaching age 20 the children of these mothers would report that 62% of their mothers are alive, when survival of mothers was 54% and survival of women in general was 53% (Figure 3b). The “double-dip” in the ratio of the proportions not surviving (reported/women or reported/mothers in Figure 4b) evident in the 12 years following birth reflects the bimodal distribution of mortality amongst those infected by their mothers. Those infected perinatally are assumed by the model to have a median term of survival of a little over 1 year while those infected through breastfeeding are assumed to have a median term of survival of nearly 9 years. Once again, a peak differential occurs (in this case after 12 years), after which the difference becomes smaller. This decline in difference is the result of other deaths, where there is no direct relationship (i.e. due to HIV affecting fertility or resulting in mother-to-child transmission) between the mortality of the mother and child, diluting the earlier related mortality patterns.

Women and mothers at age 35

At age 35, many HIV positive women are in the later stages of HIV infection, and less fertile than uninfected women. Women who become mothers at 35 are more likely to be HIV-negative than women in general, and those that are HIV-positive and do become mothers are more likely to be HIV-positive women that are relatively newly infected. Mothers thus experience lower mortality than women in general. Mother-to-child transmission of HIV does occur, however, resulting in child-reported survival of mothers being far higher than actual survival.

Note that the effect of HIV on fertility is larger at age 35 than at age 25 (see the “mother/women” ratios in Figure 4). The model assumes declining fertility as duration since infection increases – 35 year olds have been infected for longer than 25 year olds on average. The “reported/mother” ratio, however, indicates that mother-to-child transmission and its effect on bias remains fairly constant at these two ages. The model does not assume any change in risk of vertical transmission by duration since infection.

Women and mothers – all ages aggregated

When the survival rates for women in general and for mothers, and the survival rates as reported by children, are compared, we notice that the differences are predominantly in the first 20 years after birth (i.e. the period during which most currently infected mothers die and the vast majority of infected children die). Mortality resulting from new infections of mothers and their children after this time does not involve a direct mother-child relationship, and biases tend to either remain the same or diminish as these unrelated deaths begin to make up a higher proportion of all deaths. It should be noted that an indirect relationship may well occur, and the bias may not diminish as rapidly as is modelled here because of this. Investigating this is, however, beyond the scope of this research.

4.1.2 Births in specific years – mothers (of various ages) aggregated

The aggregated impact for cohorts whose survival is examined from 2005 is shown in Figure 3 as “all women”, “all mothers”, etc. The over-representation of HIV-positive women amongst the younger mothers is counter-balanced by under-representation of those that are HIV-positive amongst older women. The net effect is that survival of mothers is only slightly higher than that of women in general. The bias due to vertical transmission is however more significant. Survival of the cohort of women between ages 15 and 49 is about 56% after 20 years, while 57% of those women who became mothers in 2005 are still alive after 20 years, and 62% of children who are alive would report their mothers as being alive. Survival of mothers as reported by the cohort of children is thus about 9% higher than actual survival (implying that the proportion of mothers reported as not surviving this period is about 14% lower than the actual proportion not surviving).

Figure 5 and Figure 6 show similar data, aggregated across ages, for different years of birth of the child. Figure 5 shows non-survival and Figure 6 shows ratios of non-survival.

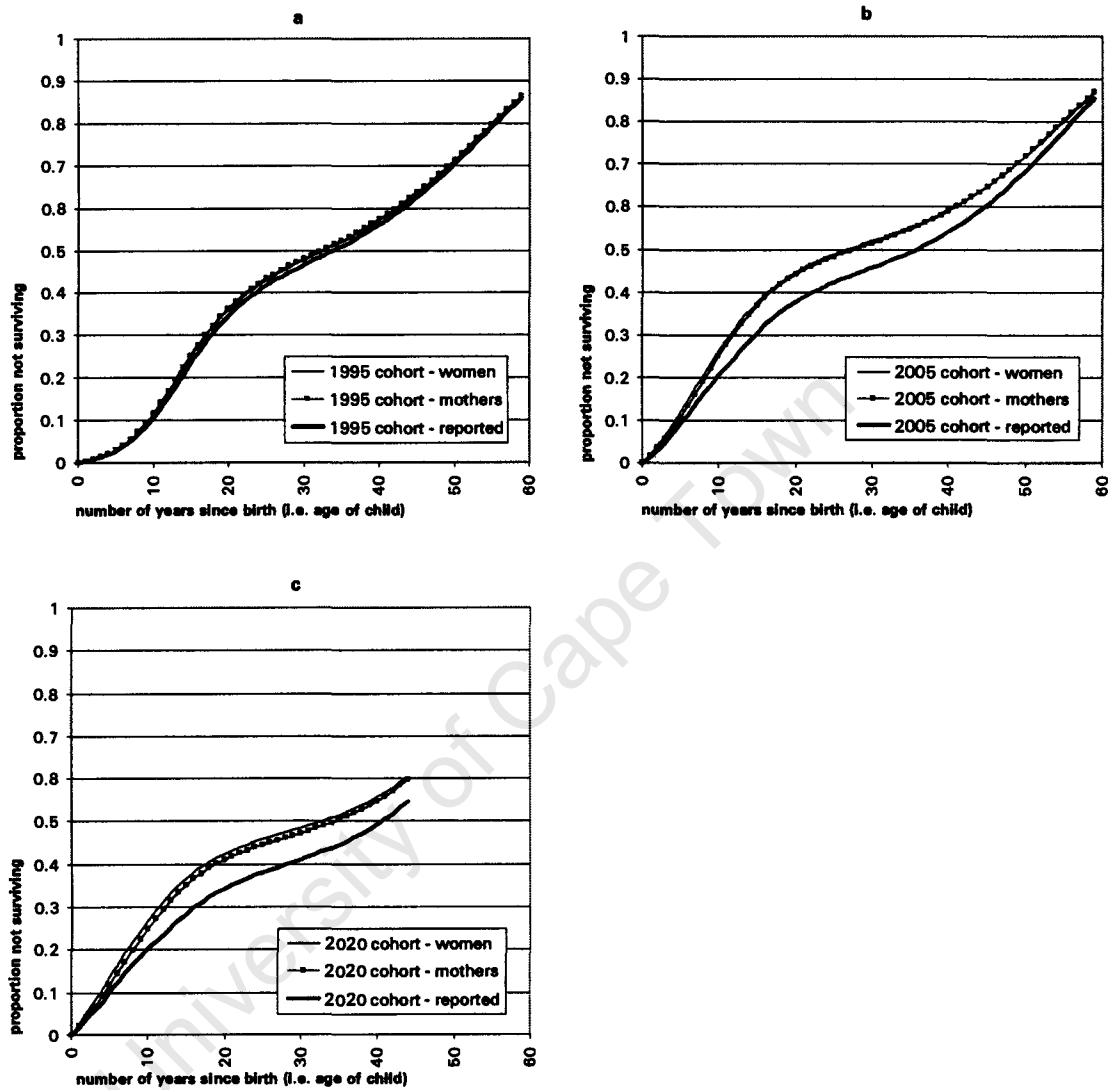


Figure 5: Proportions not surviving of women, mothers, and mothers as reported by children, by year of birth of their children: (a) 1995, (b) 2005, and (c) 2020

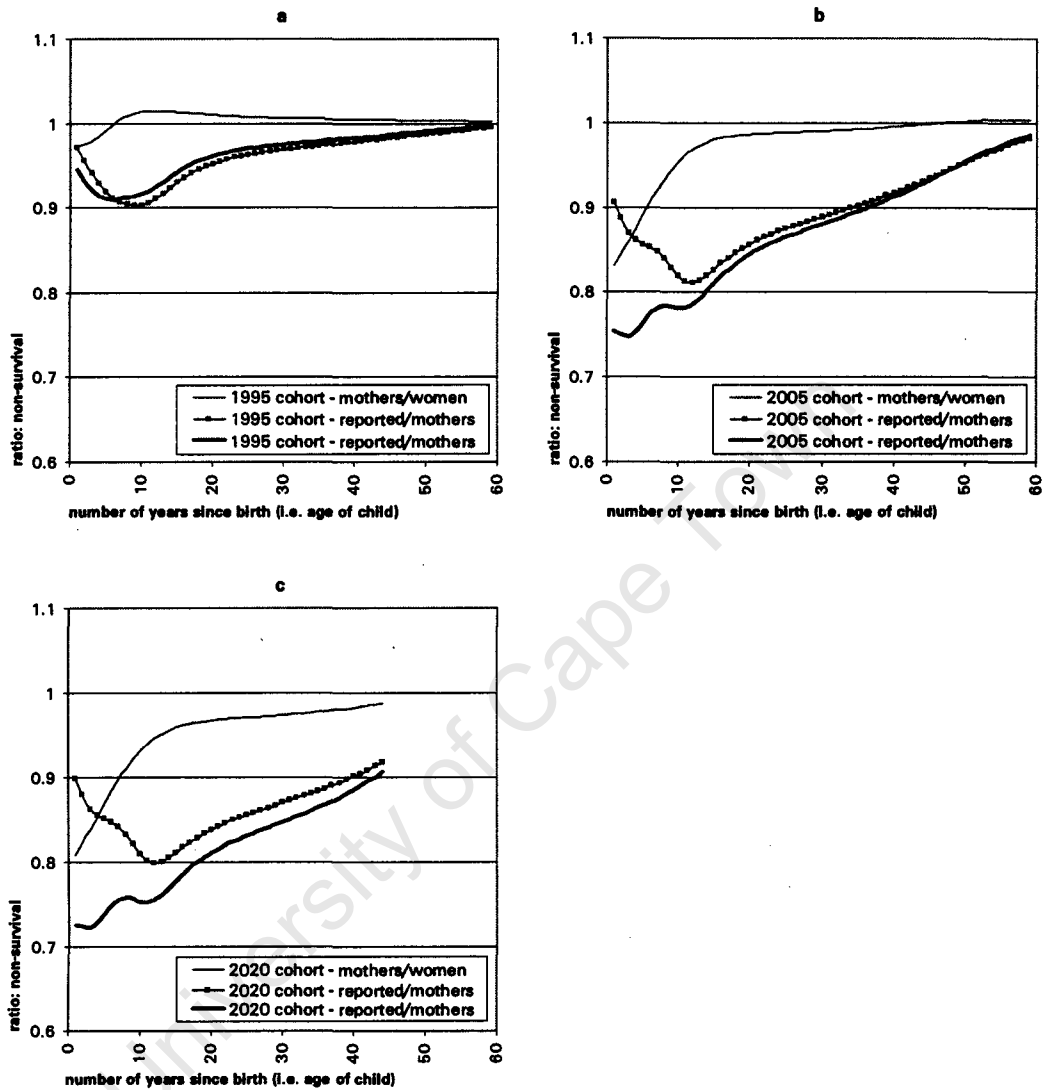


Figure 6: Ratio of proportions not surviving of mothers vs. women, mothers as reported by children vs. mothers, and mothers as reported by children vs. women, by year of birth of their children: (a) 1995, (b) 2005, and (c) 2020

Birth year 1995

For birth years soon after the epidemic has begun, HIV has a net effect on fertility that is opposite to that which is the case in later years. This is due to the fact that, at this early stage in the epidemic, there are

relatively few women in stages of HIV infection that inhibit fertility. There are, however, a relatively large number in the early stages of HIV, where higher fertility and HIV infection are related. The effect is similar to that shown for age 15 in Figure 4, but instead of only those at age 15 having infected women mainly at early stages, all infected women are relatively newly infected. As time progresses, and more women have been infected for longer, the effect of longer duration HIV infection on fertility begins to play a larger role. The mother/women line in Figure 6 shows that in 1995 this transition is in progress. The combined effect of the bias caused by the relationship between fertility and infection rates, together with the bias caused by transmission from mother to child, is that mothers reported by children are less likely to be infected and to die (see solid line for 1995).

Birth year 2005 and later

The ratio comparing non-survival of mothers to women in general (Figure 6) shows that the effect of HIV on fertility on this component of bias continues to grow after 2005. The change between 2005 and 2020 is smaller when examining the ratio of non-survival as reported by children as compared to actual non-survival of mothers (i.e. bias in children reporting deaths of mothers). This is because the prevalence of HIV amongst mothers does not change much (i.e. mother-to-child transmission is not that much higher). In essence, the change in average duration since infection after 2005 drives an increase in fertility related bias (ratio of mothers to women), and a much smaller increase in bias related to transmission of infection from mothers to children (ratio of reported to mothers).

4.1.3 Women, mothers, and reported mortality of mothers in period format

The orphanhood method depends on reported survival of mothers, as reported by children of different ages at the time of a the survey. Regrouping the data calculated and shown in cohort form into a period format will show what biases are likely to be shown by different ages or age groups in specific calendar years. Single age groups in various years are examined first (Figure 7), and the data are then grouped as they would be when used in the orphanhood method (Figure 8).

Effect of HIV on fertility – mothers vs. women

Figure 7 and Figure 8 show two representations of the same data (ratios of non-survival of mothers to that of women), first in cohort format, and then regrouped in period format. Non-survival of mothers as compared to women in general reflects the effect of HIV on survival through its impact on fertility, and demonstrates the extent to which non-survival of mothers does not reflect non-survival of women (of the same age profile).

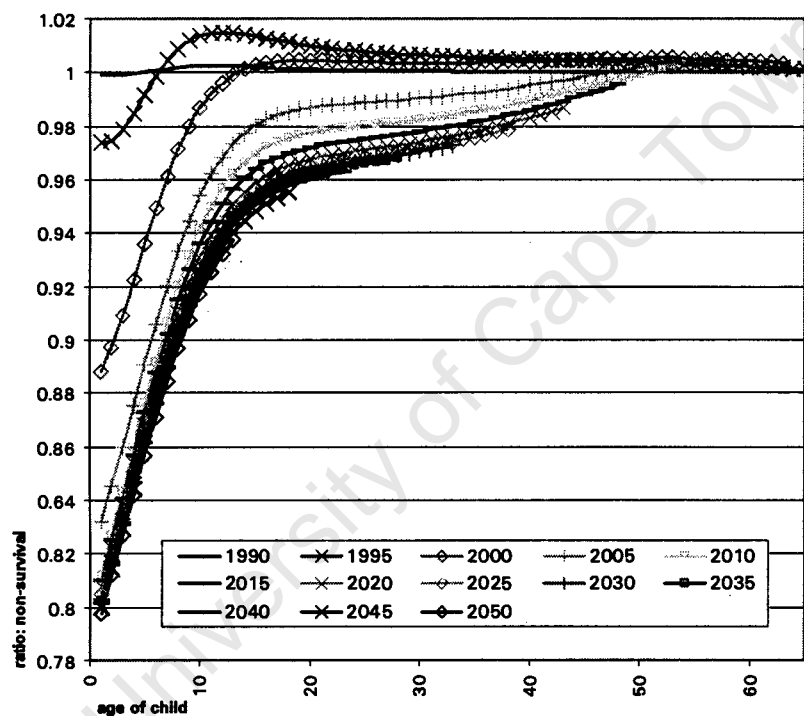


Figure 7: Ratio of proportions not surviving by birth year of child (cohort format) - mothers/women

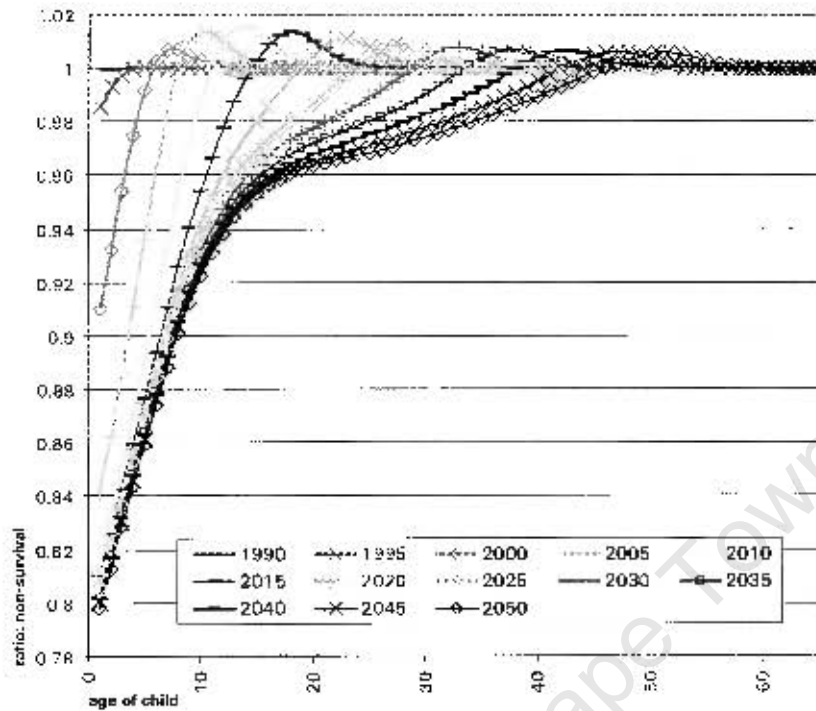


Figure 8: Ratio of proportions not surviving by calendar year (period format) - mothers/women

Cohort of women giving birth in 1995

Looking at the cohorts of women, and women who gave birth, in 1995 (Figure 7) – mothers initially experience lower cumulative mortality, but in later years this is reversed. Mothers initially include fewer women at longer durations of infection, and more women at short durations since infection. This results in lower mortality in early years, but higher mortality later as HIV infection in the younger mothers progresses to cause AIDS deaths.

Later cohorts of women giving birth

Later cohorts, however, are characterised at the start by women who have been infected for longer. In these cohorts infected women are less likely to become mothers. A smaller proportion of mothers are thus infected with HIV than in early cohorts and thus experience a lower surge in mortality than the earlier cohorts did. The

ratio of non-survival of mothers vs. women is thus lower than for the 1995 and other earlier cohorts (Figure 7). As the cohort becomes much older however (i.e. after 50 years), the older mothers will all have died, resulting in only the younger mothers in that cohort of mothers being left. This results in the bias relevant to younger mothers becoming dominant. Younger mothers are at early durations of infection, where higher infection levels correlate with higher fertility. When only the young mothers are represented, the bias is such that a higher proportion of mothers are subject to non-survival than women in general (see Figure 4 for age-specific non-survival ratios).

Regrouped in period format

Figure 8 contains the same data as in Figure 7, but regrouped in period format. It is clear how the early cohorts, which at the start were subject only to early stage HIV infection, experience a bias opposite to that of later cohorts. The early and opposite bias appears as a “moving” hump, found at the oldest ages where HIV has a noticeable effect (i.e. following the emergence of HIV). This hump widens slowly as ageing cohorts also experience selection for younger mothers (who were at low HIV durations when giving birth). The biases will only have an effect up to ages where fertility was affected by HIV – in early years predominantly only the youngest ages, but for later periods an increasing range of ages.

Effect of HIV on survival of mothers as reported by children – reported vs. mothers

Figure 9 and Figure 10 demonstrate, in cohort and period format, the effect of mother-to-child transmission of HIV (i.e. causing correlations in non-survival amongst those infected). The largest increase in bias happens in birth years up to 2010, but after this the pattern and scale of bias appears to change only slightly. This reflects the growth in prevalence and expected stabilisation of the epidemic in the South African population. Figure 10 shows the same data regrouped in period format. Once again, only the survival of women with children old enough to have been subject to HIV (i.e. born after 1985) is affected. The increase in bias that occurs up to birth year 2010 is, however, strongly evident in the periods that follow. Children born before 2010 will, in later

years in their lives, report on the survival of their mothers in a way that creates a significantly biased view of mother's and hence women's mortality.

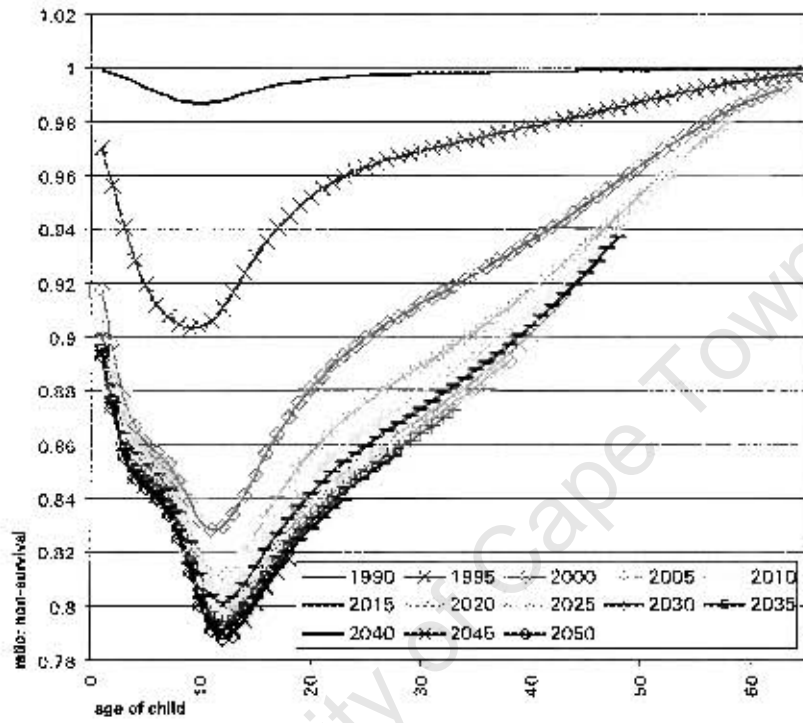


Figure 9: Ratio of proportions not surviving by birth year of child (cohort format) - reported/mothers

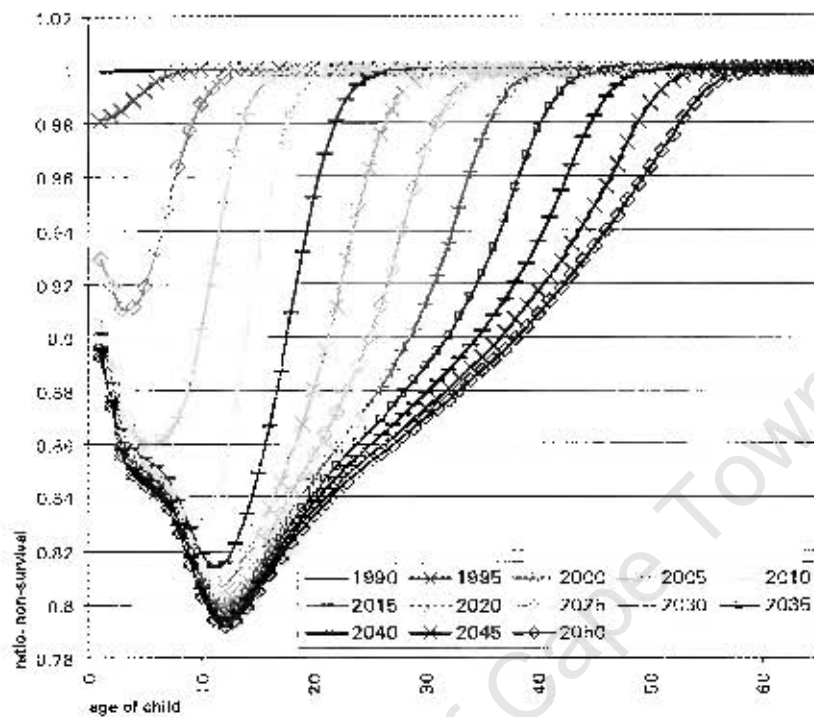


Figure 10: Ratio of proportions not surviving by calendar year (period format) - reported/mothers

Combined effect of HIV on fertility and survival

Figure 11 and Figure 12 show the combined effect of the relationship between HIV and fertility, and of mother-to-child transmission. Figure 11 shows non-survival, and Figure 12 shows survival, both grouped in period format. In 2035, reported survival of mothers by their children would be 14% higher than the survival of all women, but when examining non-survival the difference would be as much as 28% in the same year. The relative difference is at its largest for non-survival in the earliest years, where the numbers of deaths are small and numbers surviving large. In these years the relative difference remains small when looking only at survival. It is clear from Figure 11 and Figure 12 that, when looked at in period format, the biases caused by HIV take many years to become significant, and continue to grow well after the prevalence of HIV has reached its peak or has levelled off.

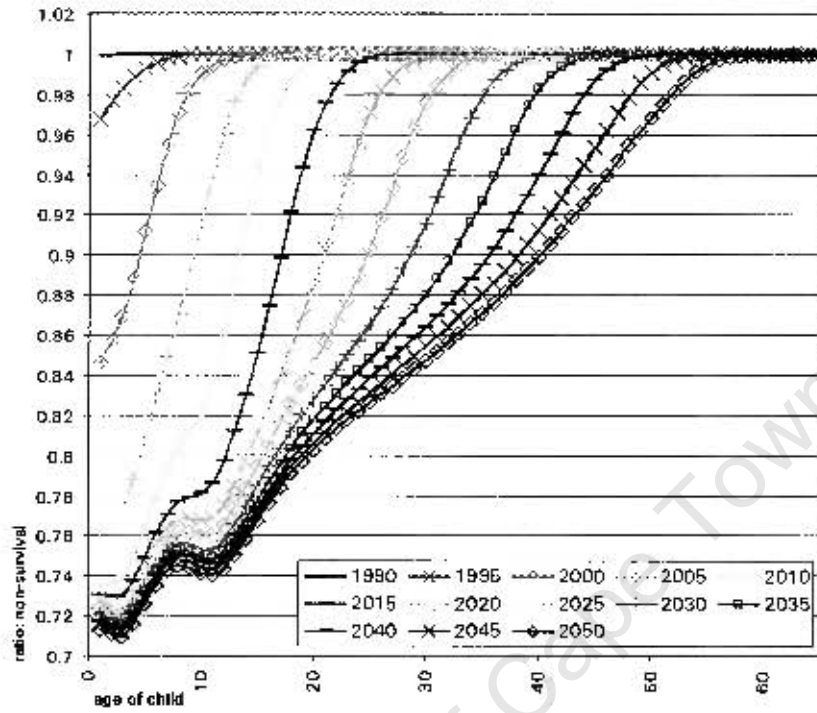


Figure 11: Ratio of proportions not surviving by calendar year (period format) - reported/women

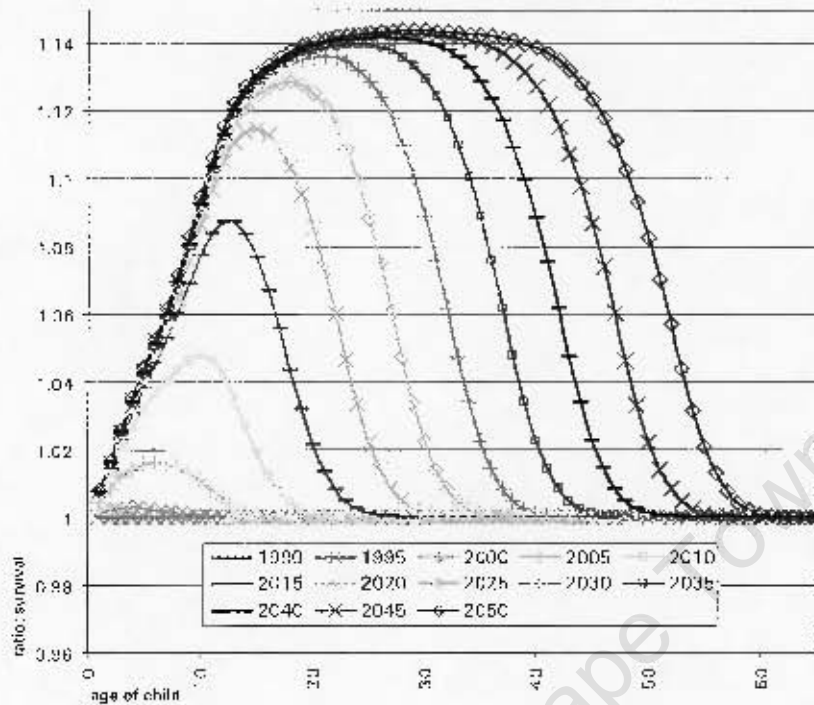


Figure 12: Ratio of proportions surviving by calendar year (period format) - reported/women

4.1.4 Women, mothers, and reported mortality of mothers in period format grouped in 5-year age bands

In practice data are likely to be collected and analysed in 5-year age bands. Figure 13 shows the same data as is shown in Figure 11, but grouped as they would be in practice. Figure 14 shows the corresponding non-survival ratios, and Figure 15 shows the survival ratios if reported survival is adjusted for selection bias as proposed by Timaeus and Nunn (1997). The method involves reducing the reported survival, and the use of revised coefficients (Timaeus and Nunn 1997). The reported survival is reduced by 25% of the prevalence at the time of birth of the reporting child for age group 5 to 9, and by 50% of prevalence at the time of birth for all other ages. The data shown in Figure 15 are the outcomes of applying these adjustments and the coefficients to the simulated reported survival rates produced by the calculations described in Chapter 3.

As those birth cohorts that were subjected to more substantial vertical transmission get older, more age groups in the various calendar years shown become affected by the biases related to HIV. Significant biases

emerge amongst younger ages by 2005, but these biases only become evident in a significant way amongst older groups well into the 21st century. The adjustments proposed by Timaeus and Nunn (1997) substantially reduce the magnitude of the error, but appear to overcompensate for the effects of HIV, especially in age group 10-

14.

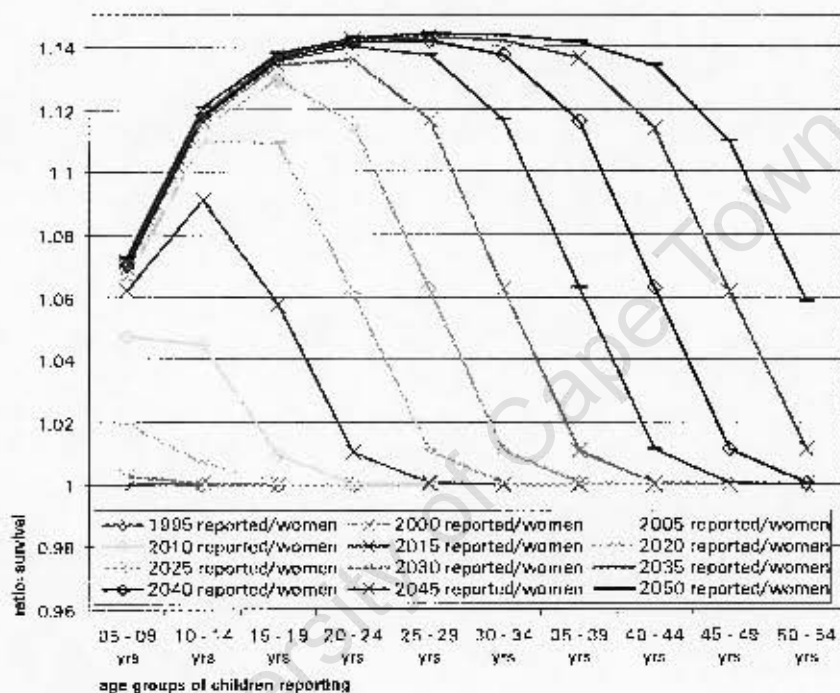


Figure 13: Ratio of proportions surviving by calendar year (period format) and 5-year age band - reported/women

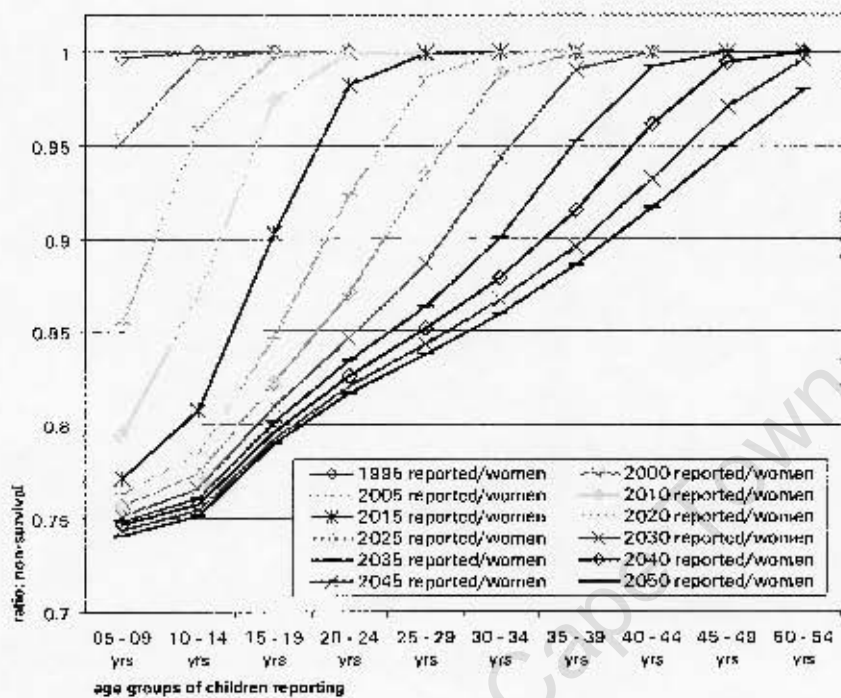


Figure 14: Ratio of proportions not surviving by calendar year (period format) and 5-year age band - reported/women

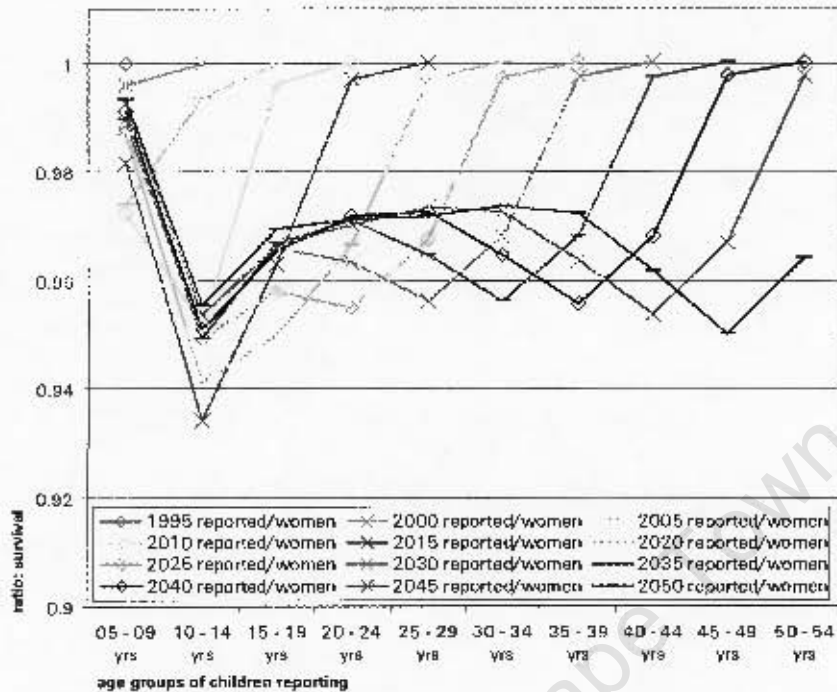


Figure 15: Ratio of proportions surviving by calendar year (period format) and 5-year age band - reported (adjusted as proposed by Timaeus and Nunn (1997))/women

4.2 Comparison of survival and non-survival: women aged 25 and women (with the age distribution of mothers)

The results shown in section 4.1 demonstrate that there is substantial error present when using reported survival of mothers as an estimator for the survival of women of the same age profile as mothers. In this section the unbiased future survival (i.e. not affected by the fertility and vertical transmission effects of HIV) of women will be compared to that of women aged 25, to establish the error created by the change in age-specific mortality caused by HIV. The orphanhood method makes use of coefficients to convert survival of women (with an age range identical to that of mothers) into $l_{(25+x)}/l_{(25)}$. The coefficients are applicable only to 5-year age bands in period format – earlier aspects of the calculations are shown without the conversion factors, but this section also shows results in 5-year bands after correction.

4.2.1 Births in specific years – women (with the age distribution of mothers) vs. 25 year olds

When the HIV epidemic has been infecting women for 20 years or more, 25 year olds are just before the peak age for HIV-related mortality, and this individual age group has far higher mortality for a number of years than the cohort of women of child-bearing age. Survival of the cohort of women of child-bearing age would overstate survival if used as an indicator of survival of 25 year olds. The relationship reverses itself after 8 to 10 years (i.e. the cohort of women of child-bearing ages has higher mortality rates than those that were 25 at inception), but takes between 20 and 30 years to completely reverse the survival difference that is established in the first 8 to 10 years. Note that in the first few years of the epidemic, 25 year olds do not initially have higher mortality than the cohort of women of child-bearing age. The effect appears only to take hold after those in their late teens and early 20's become the dominant group subject to HIV incidence (i.e. when the levels of new infections at higher ages decline). Survival of the cohort of women who were between ages 15 and 49 in 2005 is about 56% after 20 years, while 52% of those who were 25 years old in 2005 are still alive after 20 years. Survival of women of child-bearing age is thus about 6.5% higher than survival of 25 year olds.

Figure 16 shows the same ratio (average for women 15-49 at birth) for different years of birth of the child.

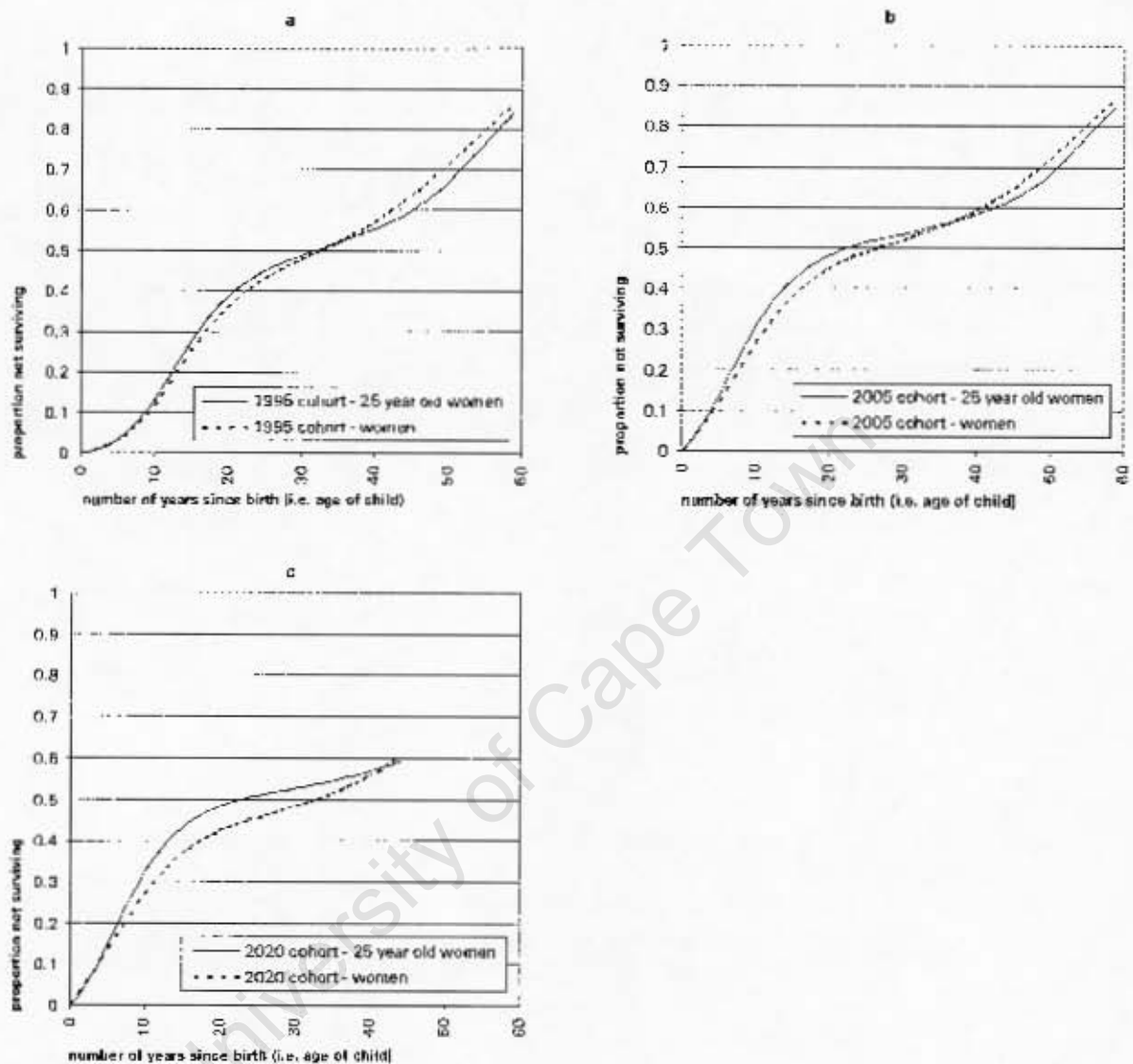


Figure 16: Proportions not surviving of women and 25 year old women: (a) 1995, (b) 2005, and (c) 2020

4.2.2 Women (with the age distribution of mothers) vs. 25 year old women, in period format

Figure 17 and Figure 18 shows the ratio of future survival of women (without adjustment to estimate survival from age 25), to survival of 25 year olds. Regrouping the data shown in cohort form into a period format shows what errors are likely to be present when examining various ages in certain calendar years. Single age groups in various years are examined first (Figure 18), and the data are then grouped.

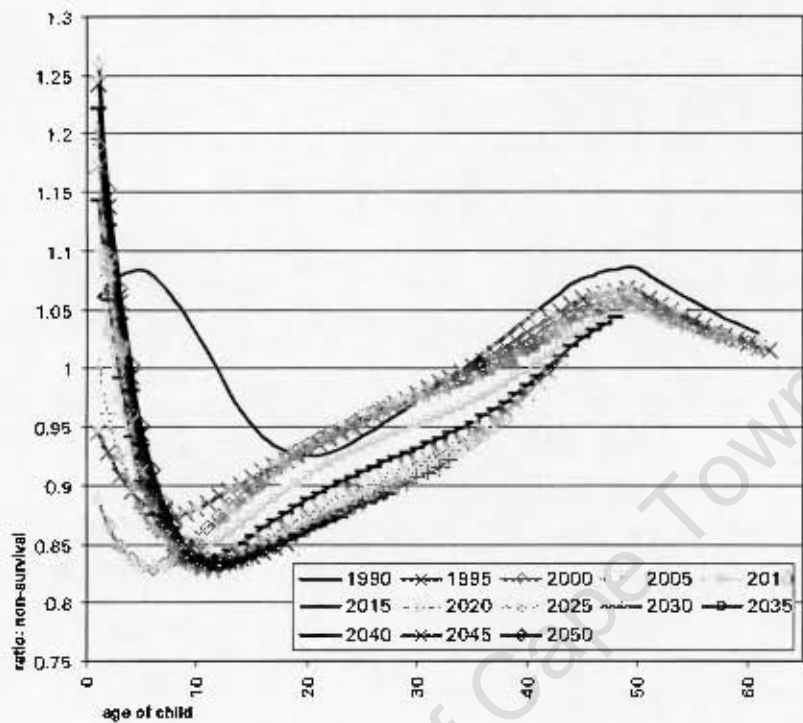


Figure 17: Ratio of proportions not surviving by birth year of child (cohort format) - women/(women aged 25)

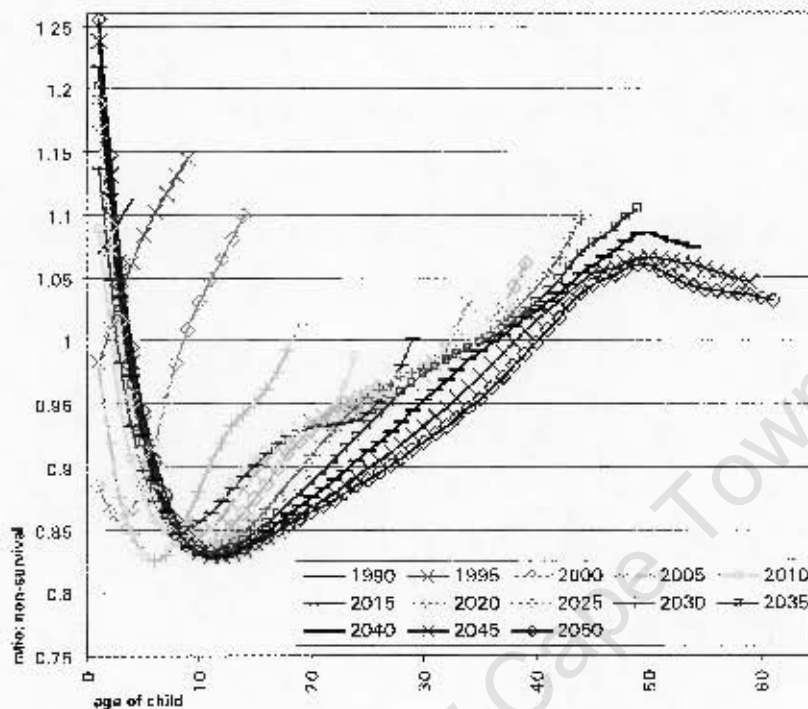


Figure 18: Ratio of proportions not surviving by calendar year (period format) – women/(women aged 25)

Various cohorts of women giving birth

Looking at the outcomes for birth year 1995, women across the age range of mothers have slightly lower mortality than women aged 25 in the first six years, but in the years that follow the opposite is the case. For later years of birth, women across the age range have lower mortality than 25 year olds for an increasing number of years, but this relationship is always reversed at some point in later years. The period during which women across the age range of mothers have lower mortality ultimately reaches a plateau at about 10 years, when the HIV epidemic has reached something resembling a stable state. In the early years of the epidemic new infections occur across a wide age range, and resulting HIV-related deaths occur across a wide age range. When the epidemic matures, the age group where most new infections occur is the late teens and early

twenties, resulting in ages from 25 to 35 being the core ages where deaths occur. The resulting significant difference in non-survival can be seen in Figure 18.

In Figure 17, the 1990 cohort stands out as unusual. The underlying output from the model (which may or may not accurately reflect reality) was examined in an attempt to understand the pattern shown. In 1990 the mortality rate for 25 year olds was slightly lower than for women aged 15-49 (age weighted to match the distribution of mothers). AIDS-related mortality is relatively insignificant this early on in the epidemic. By 1995 AIDS-related mortality has grown in both groups, but the mortality of those that were 25 in 1990 remains below that of the weighted 15-49 year old cohort from 1990. Shortly after 1995 this changes as AIDS-related mortality impacts the single-age cohort more heavily. The cumulative lower non-survival of the single-aged cohort starts to be reversed by HIV-related deaths. Higher mortality in the single-age cohort continues approximately until this cohort is 45, and the age range of the broader group spans 35-69. At this stage HIV-related mortality is rapidly declining in both cohorts, and eventually slightly lower non-AIDS (and all-cause) mortality in the single-aged cohort causes a second inflection. The pattern of lower mortality in the single-age cohort continues for another 30 years after 2010. By 2040 the elder members of the broad aged cohort have been depleted, and the cohort starts to be represented by an increasing proportion of younger members of the original cohort. This causes the third inflection. Thus the pattern for the 1990 cohorts can be explained, but whether it reflects reality depends on the reasonableness of the model.

4.2.3 Women of child-bearing age vs. 25 year old women, in period format grouped in 5-year age bands

In practice data are likely to be collected and analysed in 5-year age bands. Figure 19 shows similar data to that shown in Figure 18, but they are grouped as they would be in practice and are showing survival rather than non-survival. For the effect of HIV on the orphanhood method to become significant, there needs to be an effect in the birth year in the first instance (i.e. higher infection rates in women aged 25 at the birth of their children), followed by the consequences of this being enacted over many years. In Figure 19 it is clear that the most substantial effect, when considering the effect of HIV on mortality age patterns, occurs only from about 2020 onwards. If we consider that the HIV epidemic in South Africa grew most rapidly from 1995 to about

2005, and only matured after this, it makes sense that only birth cohorts that occurred after 2000 would be significantly affected, and it would take time for the effects to be noticeable.

Figure 20 shows similar data to Figure 19, but in this case the simulated survival of women has been used, together with the coefficients and mean age of mothers, to calculate $l_{(25+t)}/l_{(25)}$. The comparison of the calculated $l_{(25+t)}/l_{(25)}$ directly from the model, and the simulated $l_{(25+t)}/l_{(25)}$ using calculations shown in section 3.3, shows the extent of error that results from using future survival of women of child bearing age to estimate future survival of 25 year old women in the context of a high prevalence generalized HIV epidemic.

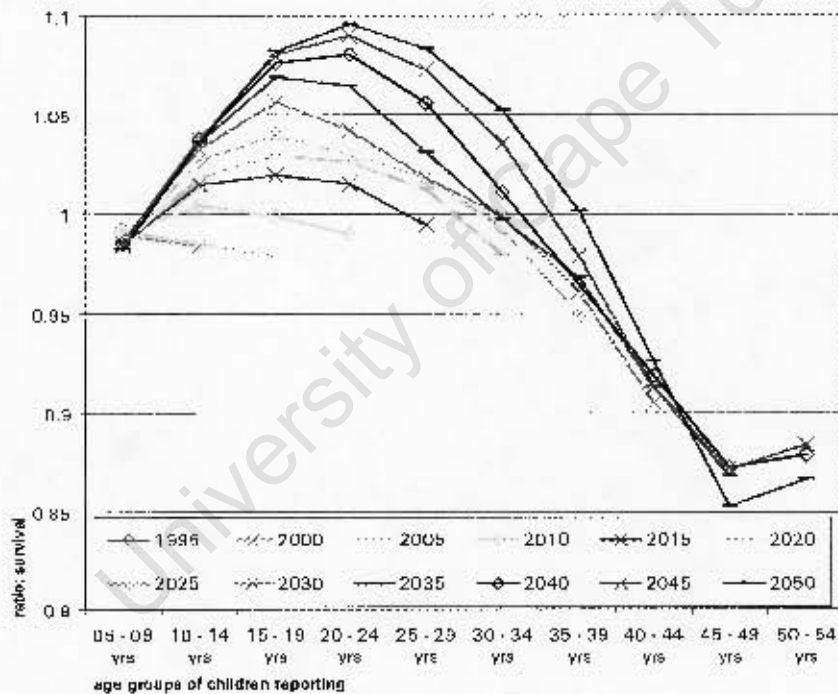


Figure 19: Ratio of proportions surviving, by calendar year (period format) and 5-year age band: women/(women aged 25 at birth of child)

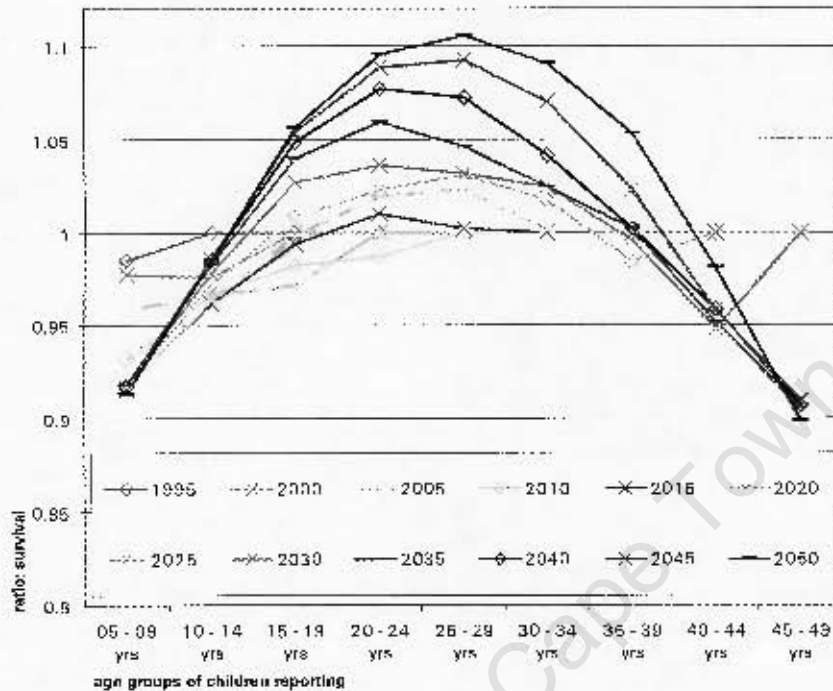


Figure 20: Ratio of proportions surviving, by calendar year (period format) and 5-year age band - women(adjusted using coefficients)/(women aged 25 at birth of child)

4.3 Reported survival of mothers adjusted using coefficients vs. survival of 25 year olds

The reported survival of mothers, as reported by surviving children, is subject to two main factors causing error, namely the effect of HIV on fertility and vertical transmission of HIV. An additional source of error exists in the use of a group of women of child-bearing age to estimate survival of 25 year old women, as shown in section 4.2.3. The combined effect of all these sources of error, as they would manifest when using the orphanhood method (United Nations 2002) to estimate $l_{(25+3)}/l_{(25)}$, are shown in Figure 21. For example, survival of mothers as reported by the 20-24 year old group in 2025, converted into $l_{(25+3)}/l_{(25)}$ using the

orphanhood method (United Nations 2002), would overstate survival by about 14.5%. An alternative view of this, with years on the X-axis, can be seen in Appendix A.

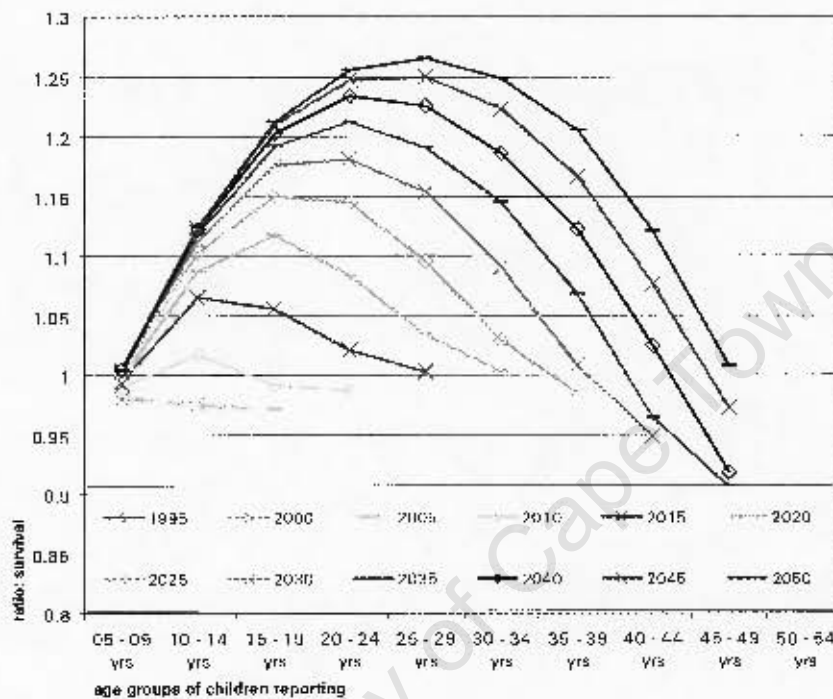


Figure 21: Ratio of proportions surviving, by calendar year (period format) and 5-year age band - reported(adjusted using coefficients)/(women aged 25 at birth of child)

An alternative set of coefficients is given by Timaeus and Nunn (1997), for ages below 30, that have been derived from data in a population (Masaka) with a lower prevalence of HIV than is being seen in South Africa. Timaeus and Nunn do warn that the coefficients they give are provisional and based entirely on data from Masaka district. The purpose of examining outcomes using these revised coefficients is to determine whether they reduce error. The adjustments for selection bias and revised coefficients appear to increase error in the two youngest age groups, and substantially decrease error in all other age groups. Timaeus and Nunn (1997) state that their adjustments for selection bias in the two youngest age groups will achieve less accurate outcomes if the median survival time of infected women is longer than that assumed in their research. Infected

women in the modelling shown here are assumed to survive between 9 and 11.5 years after infection, depending on age, as compared to 6.7 years in Masaka. The outcomes in Figure 22 reflect this.

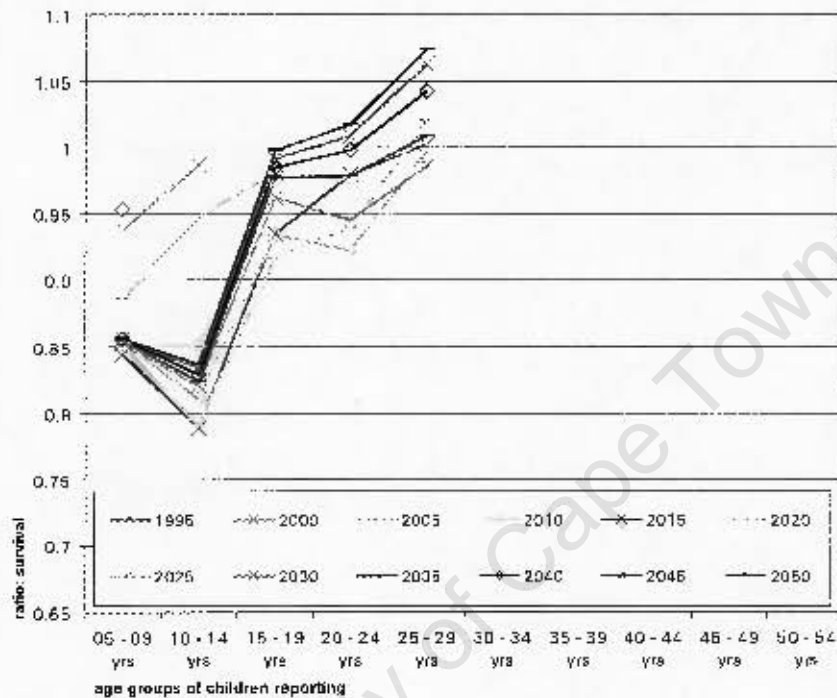


Figure 22: Ratio of proportions surviving, by calendar year (period format) and 5-year age band – reported (using adjustments for selection bias and coefficients proposed by Timaeus and Nunn (1997)) / (women aged 25 at birth of child)

4.4 Summary

Table 3 shows, in summary, the errors that are likely to occur as a consequence of some assumptions in the orphanhood method, when using the reported survival of Black African mothers in South Africa. ‘Reported vs. Mothers’ refers to the error that results from the related risk of mortality of mothers and children during a high prevalence generalized HIV epidemic. ‘Mother vs. Women’ refers to the error that results from changes in fertility patterns and the relationship between fertility and HIV infection. ‘Women adj. vs. 25 year old’ refers to the survival of women, adjusted using the coefficients in the method to be equivalent to $l_{(25-c)}/l_{(25)}$, in

comparison to $l_{(25+x)}/l_{(25)}$ simulated as outlined in section 3.3. This shows the error that results from the change in age-specific mortality rates (especially at the ages where peak HIV mortality occurs) and the effect this has due to the assumption that a group of women of child-bearing age will have similar survival to a cohort of women aged 25 at the birth of their child (an age close to the mean age of child-bearing).

'Combined error' shows the error that results if all these are taken into account. In the earliest years of the HIV epidemic, some of the errors appear to cancel each other out. In later years the error caused by the correlation between the mortality of mothers and children dominates, but the error due to changes in the age profile of mortality also plays a significant role. The combined error, when examining survival, only appears to become significant between 2010 and 2020. Table 4 has similar ratios to Table 3 but is instead showing non-survival. At younger ages (of reporting children) relatively few deaths will have occurred and differences in non-survival would appear trivial in a table showing survival. It is clear from Table 4 that significant errors in non-survival emerge as early as 2000 in younger reporting age groups.

		05 - 09 yrs	10 - 14 yrs	15 - 19 yrs	20 - 24 yrs	25 - 29 yrs	30 - 34 yrs	35 - 39 yrs	40 - 44 yrs	45 - 49 yrs
Reported vs. Mothers	2000	1.003	1.000							
	2010	1.036	1.047	1.012	1.001					
	2020	1.042	1.091	1.103	1.088	1.014	1.001			
	2030	1.043	1.091	1.112	1.118	1.111	1.070	1.014	1.001	
	2040	1.043	1.090	1.109	1.115	1.119	1.121	1.112	1.073	1.016
	2050	1.044	1.091	1.109	1.113	1.116	1.118	1.121	1.123	1.115
Mothers vs. Women	2000	1.000	1.000							
	2010	1.012	0.998	0.997	1.000					
	2020	1.023	1.018	1.008	0.994	0.997	1.000			
	2030	1.025	1.024	1.020	1.018	1.005	0.993	0.997	1.000	
	2040	1.028	1.028	1.025	1.023	1.020	1.015	1.003	0.991	0.998
	2050	1.027	1.027	1.028	1.026	1.025	1.023	1.018	1.010	0.995
Women adj. vs. 25 year old	2000	0.977	0.978							
	2010	0.932	0.967	0.982	0.987					
	2020	0.917	0.962	0.996	1.019	1.023	1.002			
	2030	0.917	0.978	1.027	1.036	1.031	1.025	0.998	0.948	
	2040	0.917	0.985	1.048	1.077	1.072	1.041	1.002	0.959	0.907
	2050	0.914	0.985	1.056	1.095	1.105	1.091	1.053	0.981	0.898
Combined error (Reported adj. vs. 25 year old)	2000	0.980	0.978							
	2010	0.991	1.017	0.992	0.987					
	2020	0.998	1.067	1.117	1.083	1.034	1.002			
	2030	1.003	1.112	1.175	1.181	1.153	1.090	1.007	0.948	
	2040	1.004	1.122	1.203	1.233	1.225	1.185	1.122	1.024	0.918
	2050	1.003	1.124	1.213	1.256	1.265	1.246	1.205	1.121	1.007

Table 3: Summary of ratios of survival of various categories of women - by cause of error, year and age group of respondent

		05 - 09 yrs	10 - 14 yrs	15 - 19 yrs	20 - 24 yrs	25 - 29 yrs	30 - 34 yrs	35 - 39 yrs	40 - 44 yrs	45 - 49 yrs
Reported vs. Mothers	2000	0.947	0.996							
	2010	0.839	0.863	0.968	0.999					
	2020	0.831	0.813	0.863	0.915	0.982	0.999			
	2030	0.827	0.805	0.834	0.862	0.891	0.937	0.988	1.000	
	2040	0.824	0.800	0.827	0.851	0.870	0.891	0.917	0.956	0.992
	2050	0.821	0.797	0.823	0.845	0.863	0.880	0.899	0.922	0.947
Mothers vs. Women	2000	1.005	1.001							
	2010	0.947	1.007	1.007	1.000					
	2020	0.919	0.965	0.992	1.008	1.004	1.000			
	2030	0.910	0.953	0.972	0.982	0.995	1.006	1.003	1.000	
	2040	0.906	0.947	0.963	0.971	0.979	0.987	0.998	1.006	1.002
	2050	0.902	0.944	0.960	0.968	0.971	0.978	0.986	0.994	1.002
Women adj. vs. 25 year old	2000	1.544	1.405							
	2010	1.312	1.094	1.048	1.037					
	2020	1.321	1.070	1.002	0.978	0.971	0.998			
	2030	1.318	1.040	0.967	0.963	0.970	0.978	1.004	1.051	
	2040	1.319	1.027	0.939	0.920	0.932	0.966	0.999	1.029	1.059
	2050	1.336	1.027	0.929	0.899	0.899	0.920	0.957	1.013	1.060
Combined error	2000	1.469	1.400							
(Reported adj. vs. 25 year old)	2010	1.042	0.951	1.021	1.036					
	2020	1.007	0.839	0.847	0.901	0.957	0.997			
	2030	0.990	0.795	0.781	0.813	0.867	0.918	0.984	1.050	
	2040	0.983	0.776	0.746	0.757	0.788	0.840	0.905	0.983	1.052
	2050	0.987	0.770	0.731	0.730	0.747	0.781	0.835	0.914	0.996

Table 4: Summary of ratios of non-survival of various categories of women - by cause of error, year and age group of respondent

Table 3 and Table 4 show that biases emerge quickly as a consequence of the adjusted survival of women not accurately representing the survival of 25-year olds. This bias is soon reversed to some extent, except in the youngest group (for an explanation see Figure 18 and section 4.2.2). Subsequent biases take far longer to emerge and various ages are impacted to varying degrees and at different times. The start year for the epidemic in South Africa is assumed, in the modelling shown here, to be 1985. The combined error in survival in the most affected age-group is a little over 1% by 2010. The errors in the same year, when examining non-survival, are less than 5%, although they were temporarily as large as 40% (10 years before) in the two youngest age groups. In 2010 the combined error (visible in both tables) is mainly due to both an overstatement of survival due to correlations in mortality between mother and child, and an understatement due to lower survival of 25 year olds. The net effect hides larger underlying errors.

By 2020 the combined bias has become highly significant – at this time all cohorts that gave birth after 1995 are substantially affected, except the youngest. The combined error in both survival and non-survival in age group 5-9 is relatively small – the net effect of substantial biases in different directions..

When observing the various sources of error, the errors follow varying patterns in how and when they emerge. Errors due to correlations of mortality between mother and child (reported vs. mothers) all occur in

the same direction, but are of different magnitudes by age. HIV incidence peaked in South Africa in the late 1990's – all cohorts that gave birth after this period of rapid spread of HIV are heavily impacted by correlations of mortality.

Errors due to the relationship between HIV and fertility (mothers vs. women) take a bit longer to develop. At some ages and durations HIV infection is non-causally related to higher fertility, and at other ages and at longer durations of HIV infection, HIV causes lower fertility. Early and younger cohorts experience the first effect, and later and older cohorts the latter. It takes time for the latter effect to become dominant – only cohorts related to birth years after 2000 appear to show any substantial error, and these errors never become especially large as age-specific errors balance each other out to some extent. Older reporting ages are also substantially less affected than the younger reporting ages.

Errors due to the cohort of mothers not accurately reflecting the mortality of 25 year olds due to changes in age-specific mortality patterns (“women adj. vs. 25 year old”), emerge quickly for younger ages, but take longer to emerge for older ages. Examination of Figure 19 and Figure 20 makes it clear that the coefficients used to convert survival rates into $l_{(25+x)}/l_{(25)}$ contribute to the errors that emerge at younger ages. Cohorts related to births after 2005 show some error at older ages (i.e. 15-19 and older), but the cohorts born 10 years later show errors twice as large at age 25-29, when examining both survival and non-survival. Consider that most infections occurred in the late 1990's - during this period a relatively wide age range of women were being newly infected. The pattern that eventually emerges, of HIV incidence being mostly in ages 15 to 22, is one that occurs only later in the epidemic. The reason for this is that ages following sexual debut continue to provide susceptible women who are subject to risk of infection long after the pool of susceptible women in other age groups has been exhausted. If a wide age range of women were infected in the late 1990's, younger ages would begin to dominate incidence shortly after this - ages 15 to 22 become the age group experiencing heaviest incidence after about the year 2000. Given that HIV-related deaths follow 8-12 years later, this would mean that 25 to 35 year-olds experience higher mortality relative to other ages in the years following 2010 (see Appendix B for illustration of age-specific mortality by year). The rapid growth in the error shown as “women adj. vs. 25 year old” occurs most rapidly in cohorts that relate to births after 2010.

Chapter 5: Discussion and conclusions

This research clearly shows that the orphanhood method is subject to significant bias due to violations of the assumptions of the method in all but the earliest years of a high prevalence generalized HIV epidemic. The version of the method published by the United Nations Population Division (United Nations 2002) over-estimates survival and produces estimates of mortality that are significantly below what is being experienced in many countries with large HIV epidemics. The method of adjustment proposed to Timaeus and Nunn (1997) results in under-estimating survival for the youngest reporting age groups but is within 5% of the correct rate of survival for other groups.

The spread of HIV began within the general population in South Africa in approximately 1985. Prevalence amongst pregnant women has exceeded 20% since 1998, and has been close to 30% for the last few years. The biggest errors in the method, when applied to Black African Females in South Africa, thus emerge between 2010 and 2020. The adjustments proposed by Timaeus and Nunn (1997) appear least able to correct for the errors in these years. Many other countries in sub-Saharan Africa have HIV epidemics which began earlier and which in some cases have been more severe than the epidemic in South Africa. Application of the original method (United Nations 2002) in these countries would now result in over-estimating survival, leading to under-estimates of mortality. Application of the revised method as proposed by Timaeus and Nunn (1997) would probably now slightly under-estimate survival, leading to over-estimates of mortality. The question arises as to how this research compares with other research on mortality. Timaeus and Jasseh (2004) examine two methods of estimating mortality (the orphanhood method and a method using sibling histories) in a number of African countries. The results are compared with each other, and with other data where they exist for some countries. Orphanhood data is collected only for ages 5-9 and 10-14, and the adjustments for selection bias applied (using rates of adjustment that are slightly lower than those proposed by Timaeus and Nunn (1997)). Outcomes for the orphanhood method and the sibling-history method are largely similar, but in one country where comparisons can be made with estimates derived using the registered death data (Zimbabwe), it appears mortality is under-estimated at time durations more distant from the present. The time location of this work is important. The investigation uses data from Zimbabwe in the 1990's. If the epidemic

in Zimbabwe started a couple of years before it did in South Africa, this would mean that by the time the data used by Timaeus and Jasseh was collected Zimbabwe was around 15 years into its epidemic. The research presented here shows that at that stage in the epidemic, after adjustment using the revised coefficients proposed by Timaeus and Nunn (1997), there should be very little bias. The outcomes of the study by Timaeus and Jasseh (2004) differ slightly from those presented here, but are probably the result of data errors rather than bias due to HIV/AIDS.

The question arises as to what extent the results of the research undertaken here are a consequence of the model used. A few assumptions in the ASSA2002 model may affect the outcomes. Both the assumed age of sexual debut and the median age of child-bearing (and the age-specific fertility rates) may have an impact on the outcomes. If sexual debut occurs later, or median age of child-bearing is younger (i.e. the gap between the two is smaller), the relationship between age and fertility may be affected. This could mean fertility is less affected by HIV, as fewer women of child-bearing age would be infected, and those that are infected would be at lower durations since infection. Should the gap between the two be larger, the opposite may be true. The age at which HIV related deaths peak would also change in relation to the HIV related mortality of mothers. Another aspect of the modelling that may affect the outcomes is the use of risk groups (see section 3.2 for a description of the risk groups). After allocation of the population to particular risk groups at a young age, there is no movement in the model between risk groups. The effect of this is that new infections virtually cease above certain ages, as the pool of susceptible people in the riskier groups is depleted. This might distort the age profile of mortality. This aspect of the model, as well as the peak of HIV related deaths in relation to the age profile of child-bearing, could affect the relationship between reported survival and the life table measure $l_{(25+x)}/l_{(25)}$. At this stage, it is thus essential to keep in mind that the outcomes shown here are modelled and will necessarily reflect reality imperfectly. The only way to verify the outcomes shown would be to apply the method to data from a population with a high prevalence generalized HIV epidemic, for which we know the mortality rates. South Africa is clearly a country where death data is now of sufficient quality to allow such a comparison (Dorrington *et al.* 2004).

Further research using data from South Africa's death registration system would give insight into the validity of the findings presented here. Further research could also be undertaken to identify and use data from

other surveys (including the Demographic and Health Survey, and national census data from various countries) to test the validity of this research and its applicability to other countries.

This research relates to a specific population, namely Black South African women and their female offspring. Further investigation could extend this research to populations with similar levels of HIV and patterns of HIV spread. There may be some differences where other populations have lower or higher median ages at which children are born – especially in respect of the error caused by using survival of women of child-bearing age to represent survival of women aged 25.

The research undertaken here can also be extended to identify the extent of bias when examining male mortality (i.e. children reporting on the survival of their fathers). Extending the scope of this work to fathers introduces a further complication in that the vertical transmission of HIV occurs via the woman who bears the child. If the father is infected with HIV, the mother may or may not be infected, and vice versa. This research establishes a foundation that could be extended to include fathers, using certain assumptions, but this is beyond the scope of this research.

This research has shown the timing and the magnitude of some errors likely to emerge when using the orphanhood method to estimate mortality in countries with high prevalence generalized HIV epidemics. The errors covered in this research represent only some of the potential problems that may occur. This research only assesses the method to the point where $l_{(25+x)}/l_{(25)}$ values are generated. Two additional problems are easily identified, and require further research. First, the use of standard life tables to convert the $l_{(25+x)}/l_{(25)}$ measures into other useful measures of mortality is not appropriate. The age-specific variations in mortality caused by HIV should be reflected in any life tables used to generate other measures of mortality or survival. Second, the calculations related to timing also need revision due to changes in age-specific mortality patterns. The timing correction calculations depend on the assumption that mortality continues to increase with age at all ages, which is not the case when HIV is highly prevalent. In light of the other uncertainties in estimating mortality however, the problem of correctly estimating timing is likely to be relatively minor.

Two pharmaceutical interventions require consideration in that they may affect the biases in the orphanhood method due to HIV. These are prevention of mother-to-child transmission of HIV using

antiretroviral drugs (PMTCT), and anti-retroviral treatment (ART). The first intervention is now applied fairly widely in many southern African countries, and the second is subject to more limited use. PMTCT will significantly reduce the correlations in mortality between mothers and children. ART will lengthen the median time to death, be correlated to PMTCT to some extent (as infected mothers who know they are infected may receive both PMTCT and ART, and ART itself has some effect in preventing mother-to-child transmission), and result in higher fertility at longer durations of infection. ART will thus reduce bias due to all three causes dealt with in this research, but will have its greatest effect in shifting the peak of HIV mortality to later ages (see Appendix B for the age profile of mortality), and will therefore reduce the error in the relationship between the mortality of women in general and women aged 25 at the time when they give birth. The extent to which error is reduced by PMTCT and ART is a function of how prevalent these interventions are, and how successfully they are used. The effect of these two interventions, based on assumptions on usage and success which reflect the assumptions of the ASSA2002 model, are shown in Appendix D. With PMTCT and ART interventions, the ratio of survival of mothers as reported by children vs. mothers, reflecting related mortality between mothers and children, shows significant reductions in bias in those cohorts in which a sizeable proportion of women would have benefited from ART, with the effect fading as cohorts age. The ratio of survival of mothers vs. survival of women, reflecting the effects of HIV on fertility, shows increased bias. The ratio of survival of women vs. 25 year old women shows reduced bias at the start but is subject to a slower reversal of that bias – the bias is thus larger for age group 10-14, but smaller for all others. The interventions result in a substantial reduction in the combined error in reported survival, for cohorts that benefited from these interventions at years close to the birth year. The reduction in error is especially significant at the younger reporting ages, where mortality estimates would be more sensitive to errors in reported survival. Any revised adjustments for selection bias, and revisions to coefficients for conversion of survivorship into life table measures should ideally take interventions into account.

Having outlined the errors leading up to the calculation in the method of $l_{(25+x)}/l_{(25)}$ the question arises as to whether the calculation of revised correction factors is plausible. The correction factors proposed by Timaeus and Nunn (1997) can be used to estimate mortality in a population that has a high prevalence of HIV, but will to some extent overcompensate for errors, especially at younger reporting ages. Calculating

revised correction factors is beyond the scope of this research but should be possible. Revised coefficients for conversion of survivorship into life table measures would also be useful. Any coefficients that convert survival rates into $l_{(25+x)}/l_{(25)}$ would probably have to include terms that deal with a number of factors related to HIV. Prevalence in previous years, possibly even historical HIV prevalence specific to each age group would have to be considered. This would probably be important in populations that have experienced rapid increases in mortality due to HIV (as well as changes in age-specific patterns), and are now experiencing declines in prevalence. Some countries are experiencing declines in mortality due to antiretroviral therapy, and are simultaneously experiencing increased prevalence as infected people live longer. Any set of revised coefficients for conversion of survival into $l_{(25+x)}/l_{(25)}$ would have to consider all these factors to be widely applicable and useful in a range of developing countries with high prevalence generalized epidemics.

The research has demonstrated that using the responses of children about the survival of their parents in the current versions of the orphanhood method would ultimately result in a substantial underestimation of the mortality of women in countries that have well established high prevalence HIV epidemics. The adjustments to the method suggested by Timaeus and Nunn (1997) substantially improve the accuracy of the method, but themselves require correction, especially at younger reporting ages. The research also shows that these errors take a number of years to develop, and that errors are likely to be relatively small in the first 20 years of an epidemic. Increased use of antiretrovirals and more effective prevention of mother-to child transmission may in the future reduce or eliminate the need for adjustments to the method. Currently, however, the original method as published by the United Nations Population Division (United Nations 2002) should not be used, and the method with adjustments as proposed by Timaeus and Nunn (1997) can be used for older reporting ages.

References

(2006b). Matlab. Natick, Massachusetts, The MathWorks Inc.

ASSA (2005). ASSA2002 AIDS and Demographic Model, Actuarial Society of South Africa AIDS sub-committee.

Blacker, J. G. C. (1977). The Estimation of Adult Mortality in Africa from Data on Orphanhood. Population Studies 31(1): 107-128.

Brass, W. and B. Bamgboye (1981). The time-location of reports of survivorship: estimates for maternal and paternal orphanhood and the ever-widowed. London School of Hygiene and Tropical Medicine Centre for Population Studies Working Paper 81(1).

Brass, W. and K. H. Hill (1973). Estimating adult mortality from orphanhood. Paper presented at proceedings of the International Union for the Scientific Study of Population.

Community Survey (2007). (Revised version). Statistical release P0301. Pretoria, South Africa, Statistics South Africa.

Dept. Health, S. A. (2009). 2008 National Antenatal Sentinel HIV & Syphilis Prevalence Survey, Department of Health, South Africa.

Dorrington, R. E., T. A. Moultrie, et al. (2004). Estimation of mortality using the South African Census 2001 data. Centre for Actuarial Research (11), Cape Town, South Africa, University of Cape Town.

Gregson, S., B. Zaba, et al. (2002). The impact of HIV-1 on fertility in sub-Saharan Africa: causes and consequences. Expert Group Meeting on Completing the Fertility Transition, New York, 11-14 March 2002, New York, New York, United Nations, Department of Economic and Social Affairs, Population Division.

Hill, K. H. and J. Trussel (1977). Further developments in indirect mortality estimation. Population Studies 31(2): 313-333.

Hosegood, V., A.-M. Vanneste, et al. (2004). Level and causes of adult mortality in rural South Africa: the impact of AIDS. AIDS 18(4): 663-671.

INDEPTH network (2004). INDEPTH Model Life Tables for Sub-Saharan Africa, Ashgate, United Kingdom.

Johnson, L. and R. E. Dorrington (2006). Modelling the demographic impact of HIV/AIDS in South Africa and the likely impact of interventions. Demographic Research 14: 541-74.

Moultrie, T. A. and P. R. Dorrington (2004). Estimation of fertility from the 2001 South Africa Census data. Centre for Actuarial Research Monograph(12).

Palloni, A., M. Massagli, et al. (1984). Estimating Adult Mortality with Maternal Orphanhood Data: Analysis of Sensitivity of the Techniques. Population Studies 38: 255-279.

Timaeus, I. M. (1986). An assessment of Methods for Estimating Adult Mortality from Two Sets of Data on Maternal Orphanhood. Demography 23(3): 435-450

Timaeus, I. M. (1991). Estimation of Adult Mortality from Orphanhood Before and Since Marriage. Population Studies 45(3): 455-472.

Timaeus, I. M. (1991). Estimation of Mortality from Orphanhood in Adulthood. Demography 28: 213-227.

Timaeus, I. M. and M. Jasseh (2004). Adult mortality in sub-Saharan Africa: evidence from demographic and health surveys. Demography 41(4): 757-772.

Timaeus, I. M. and A. J. Nunn (1997). Measurement of adult mortality in populations affected by AIDS: an assessment of the orphanhood method. Health Transition Review 7(Suppl 2): 23-43.

UNAIDS (2007). "Q&A on HIV/AIDS estimates." 3.

United Nations (2002). Methods for Estimating Adult Mortality, Population Division, Department of Economic and Social Affairs, United Nations Secretariat: 55-73.

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

University of Cape Town

Appendices

Appendix A: Alternative view of Figure 21 - years on x-axis

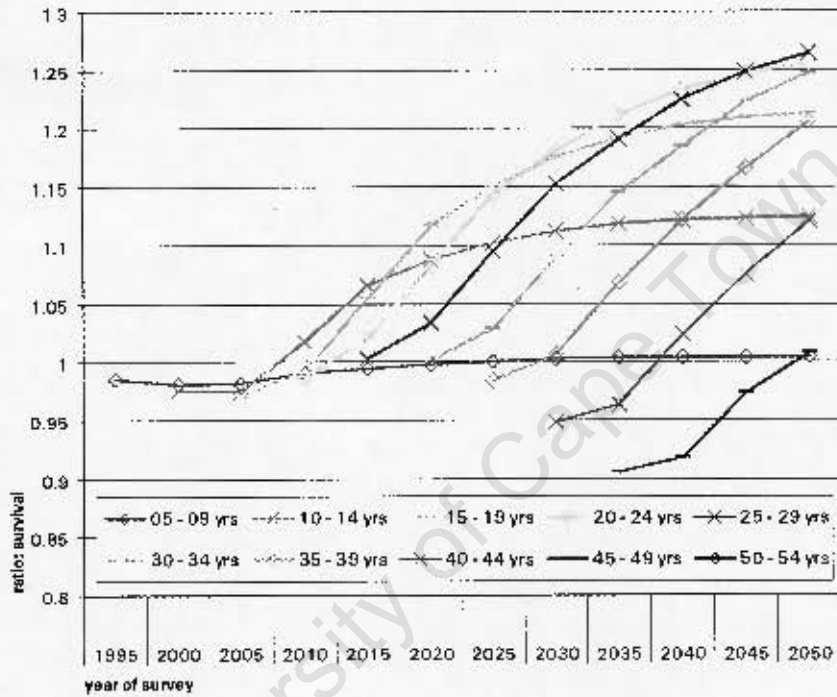


Figure A: Ratio of proportions surviving, by 5-year age band and calendar year: reported(adjusted using coefficients)/ (women aged 25 at birth of child)

Appendix B: South African Black females – modelled q_x

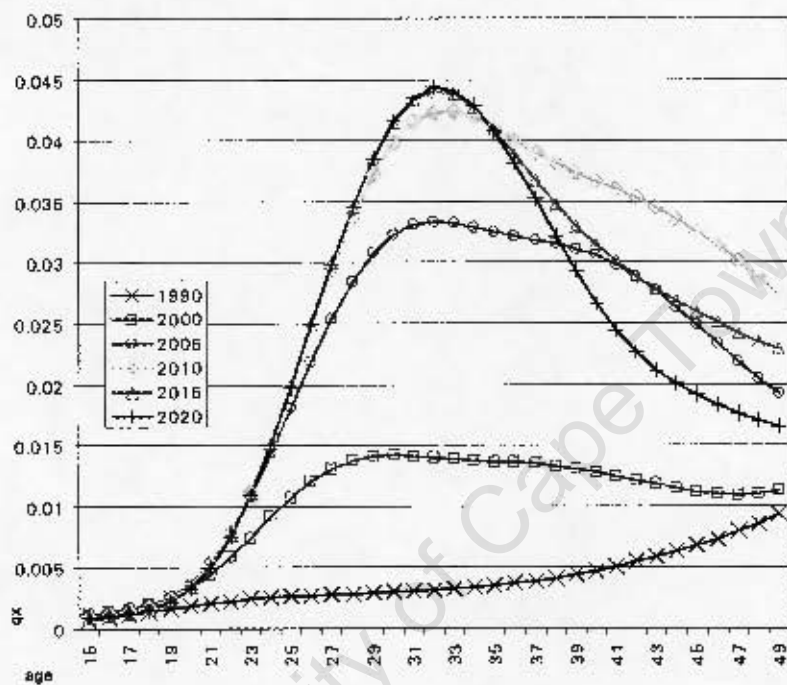


Figure B: Modelled q_x using ASSA2002 – Black South African females

Appendix C: Summary of ratios of survival – “No AIDS” scenario

The following table was produced using data from the ASSA2002 model, using a projection that does not include the spread of HIV. All subsequent calculations identical to those used in the main body of this thesis.

The ratios “Reported vs. Mothers” and “Mothers vs. Women” are not shown – values are all one.

	05 - 09 yrs	10 - 14 yrs	15 - 19 yrs	20 - 24 yrs	25 - 29 yrs	30 - 34 yrs	35 - 39 yrs	40 - 44 yrs	45 - 49 yrs
Women adj. vs. 25 year old women									
2000	0.987	0.983							
2010	0.989	0.984	0.979	0.971					
2020	0.991	0.989	0.985	0.978	0.971	0.959			
2030	0.993	0.990	0.988	0.986	0.980	0.971	0.959	0.930	
2040	0.994	0.992	0.990	0.988	0.986	0.985	0.974	0.943	0.908
2050	0.995	0.994	0.993	0.991	0.990	0.989	0.986	0.968	0.932

Table A: Summary of ratios of survival of various categories of women in “No AIDS” scenario - by cause of error, year and age group of respondent

University of Cape Town

Appendix D: Summary of ratios of survival – “MTCTP and ART” scenario

The following table was produced using data from a projection of the ASSA2002 model that includes the use of prevention of mother-to-child transmission and anti-retroviral therapy. The assumptions are based on the default assumptions of the ASSA2002 model. Assumptions relevant to survival, ART usage and effects, and PMTCT usage and effects, are shown in Table . All subsequent calculations are identical to those described in Chapter 3: .

		05 - 09 yrs	10 - 14 yrs	15 - 19 yrs	20 - 24 yrs	25 - 29 yrs	30 - 34 yrs	35 - 39 yrs	40 - 44 yrs	45 - 49 yrs
Reported vs Mothers	2000	1.003	1.000							
	2010	1.020	1.039	1.012	1.001					
	2020	1.017	1.040	1.063	1.061	1.013	1.001			
	2030	1.017	1.041	1.058	1.067	1.081	1.070	1.014	1.001	
	2040	1.017	1.040	1.067	1.067	1.074	1.078	1.085	1.073	1.016
	2050	1.018	1.041	1.067	1.068	1.072	1.076	1.077	1.077	1.086
Mothers vs Woman	2000	1.000	1.000							
	2010	1.011	0.999	0.998	1.000					
	2020	1.020	1.017	1.007	0.996	0.997	1.000			
	2030	1.024	1.028	1.026	1.020	1.006	0.993	0.997	1.000	
	2040	1.025	1.030	1.034	1.034	1.027	1.017	1.003	0.991	0.996
	2050	1.026	1.031	1.036	1.036	1.036	1.032	1.024	1.012	0.995
Women Adj vs Is	2000	0.977	0.978							
	2010	0.944	0.976	0.986	0.987					
	2020	0.926	0.962	0.991	1.013	1.016	0.997			
	2030	0.921	0.953	0.988	1.016	1.028	1.022	0.995	0.947	
	2040	0.921	0.956	0.991	1.020	1.046	1.038	1.001	0.958	0.907
	2050	0.918	0.954	0.994	1.031	1.062	1.059	1.034	0.980	0.899
Combined Error	2000	0.980	0.978							
	2010	0.983	1.018	0.995	0.987					
	2020	0.971	1.026	1.067	1.071	1.026	0.997			
	2030	0.972	1.030	1.079	1.108	1.119	1.086	1.006	0.947	
	2040	0.973	1.035	1.090	1.128	1.156	1.136	1.092	1.022	0.918
	2050	0.971	1.036	1.095	1.142	1.179	1.176	1.143	1.074	0.979

Table B: Summary of ratios of survival of various categories of women in “PMTCT and ART” scenario - by cause of error, year and age group of respondent

Usage rates	2000	2001	2002	2003	2004	2005	2006	2007	2008	Later
Mother-to-child transmission prevention (PMTCT)	0.0%	10.0%	40.0%	80.0%	80.0%	90.0%	90.0%	90.0%	90.0%	90.0%
Antiretroviral treatment (ART)	1.8%	3.7%	5.6%	7.5%	16.2%	25.0%	33.7%	41.7%	49.7%	49.7%
Mortality assumptions (non-ART)										
	Years									
Median term to death of HIV+ (14-24)	11.50									
Median term to death of HIV+ (25-34)	10.50									
Median term to death of HIV+ (35+)	9.00									
Vertical transmission rates (HIV positive mother to child)										
Perinatal transmission rate	20%									
Breastmilk transmission rate	16%									
MTCTP assumptions										
HAART take-up rate (for 80% counselled and tested)	100%									
Formula milk take-up rate	50%									
Reduction in perinatal transmission	47%									
Reduction in transmission through breast milk (for mother on ART & formula feeding)	100%									
ART assumptions										
% reduction in AIDS morbidity on ART - adults	75%									
% reduction in AIDS morbidity on ART - children	75%									

Table C: Assumptions in ASSA2002 (ASSA 2005) relevant to antiretroviral therapy (ART) and prevention of mother-to-child transmission (PMTCT)