Estimating adult mortality based on maternal orphanhood in populations with HIV/AIDS

Bruno Masquelier[1] & Ian M. Timæus[2,3]

[1] Louvain University, Belgium [2]London School of Hygiene and Tropical Medicine, UK, [3]University of Cape Town, South Africa.

This is a preprint of an article whose final and definitive form will be/is published in Population Studies [copyright Population Investigation Committee]. Population Studies is available online at: https://www.tandfonline.com/journals/rpst20

Abstract

In countries without adequate death registration systems, adult mortality is often estimated using orphanhood-based methods. The HIV pandemic breaches several assumptions of these methods, for example, by increasing the correlation between maternal and child survival. We generated 1152 populations facing HIV epidemics with microsimulations and evaluated orphanhood-based estimates against the underlying mortality rates. We regressed survivorship probabilities on proportions of respondents with surviving mothers, adjusting for trends in seroprevalence and the coverage of antiretroviral therapy. We tested the coefficients on survey and census data from 16 African countries with high HIV prevalence. The original orphanhood method underestimates mortality during an AIDS epidemic: better estimates can be obtained using revised coefficients applied to synthetic measures of maternal survival. The resulting estimates agree well with those of the UN Population Division. Orphanhood-based estimates can fill data gaps on adult mortality, including in countries with high HIV prevalence.

Keywords: Adult mortality, indirect estimation, orphanhood method, HIV/AIDS, microsimulation

Short title: HIV-related biases in the orphanhood method

1 Introduction

The highest levels of adult mortality worldwide are found in Sub-Saharan Africa (SSA). According to the United Nations (2022), the risk of dying between the ages 15 and 50 was greater than 10% in 61 countries in 2022; 48 of these high-mortality countries were located in SSA. However, because of the under-development of systems of Civil Registration and Vital Statistics (CRVS) in the region, the magnitude of this mortality burden is hard to quantify. Only a handful of countries can generate reliable mortality estimates from their death registration system, for example, South Africa and Zimbabwe (Feeney 2001, Joubert et al. 2013), while other countries have high-quality data in their capital city only (Masquelier et al. 2019).

Available adult mortality estimates, such as those developed by the United Nations Population Division and the Institute for Health Metrics and Evaluation, are therefore based on statistical models that synthesize a fairly limited set of primary estimates from censuses and surveys in SSA (United Nations 2022, Wang et al. 2020). These primary estimates typically stem from three main approaches: evaluating intercensal population change by age and sex, eliciting reports on recent household deaths, and assessing survival among close relatives (Hill et al. 2005). In particular, sibling survival histories have proved useful in reconstructing trends and age patterns of mortality (Timæus and Jasseh 2004). These are, however, relatively time-consuming to collect and inappropriate for use in censuses or rapid turn-around surveys. In contrast, orphanhood-based methods generally require only that two questions are asked: "Is your mother alive?" and "Is your father alive?". No information is required on the ages of surviving parents, ages at death, or the timing of the deaths. Mortality is estimated instead from the proportions of surviving parents. Since parents were alive at the time of the birth of their children (or at the time of conception for fathers), the duration of the exposure to the risk of dying corresponds to the age of the respondents. The average age of the parents at the start of the exposure period is simply the mean age at childbearing. Proportions of parents alive reported by individuals aged a are thus closely related to the

probability of surviving from M to $M+a$ ($_ap_M$), where M stands for the mean age at childbearing.

Henry (1960) and Brass and Hill (1973) developed the original orphanhood method, based on the theory of stable populations, which enables the expression of the frequency and survival of close relatives as a function of mortality and fertility rates. Several revisions have been proposed since (Hill and Trussell 1977, Palloni and Heligman 1985, Timæus 1991a; 1992). About a hundred censuses conducted in SSA have included orphanhood questions (Table S1). Nationally-representative surveys, such as the Demographic and Health Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS) also regularly collect data on parental survival among children under 18 years of age. Numerous studies have estimated adult mortality by means of orphanhood-based methods, including studies in populations with high HIV prevalence (Tollman et al. 1999, Feeney 2001, Dorrington et al. 2004, Hosegood et al. 2004, Lesotho NSO 2009, Nhacolo et al. 2006, Chisumpa and Dorrington 2011, Menashe Oren and Stecklov 2018, Odimegwu et al. 2018). Yet, orphanhood-based estimates remain less frequently used to monitor trends in mortality than sibling survival histories. For example, they are not included in the mortality database of the Global Burden of Disease Study (Wang et al. 2020).

The patchy use of orphanhood data for mortality estimation is probably due to concerns over data quality. The reporting of fostered orphans as non-orphans, an error referred to as "adoption effect", is thought to be common (Blacker and Mukiza-Gapere 1988, Robertson et al. 2008). Methods have been developed to correct for this, however, either by estimating mortality from orphanhood among adults only (Timæus 1991b) or by constructing synthetic cohorts from two sets of data on orphanhood (Timæus 1986; 1991a).

Another source of scepticism about orphanhood-based methods is related to selection biases. These arise if the probabilities of dying of mothers and children are correlated, if the fertility

of mothers is associated with their mortality, or when the mortality of children varies with the number of their siblings. In normal circumstances, these selection biases tend to cancel each other out (Palloni et al. 1984). In recent decades, however, HIV epidemics have amplified these biases. The transmission of HIV from mothers to children ranges from 15-45% in the absence of treatment (De Cock et al. 2000). Because of vertical transmission, fewer orphans will survive among those born to HIV-positive parents. Thus, seronegative parents will be oversampled in reports from censuses and surveys. Lower fertility of seropositive mothers will also bias the estimates downwards. Additional errors are introduced when proportions of parents alive are converted into measures of mortality with coefficients that were calculated based on standard age patterns of mortality. This is because such age patterns do not reflect the "hump" in adult mortality rates that is typical of populations experiencing a generalized HIV epidemic (Masquelier et al. 2017). Finally, biases are also introduced because the rapid changes in mortality during the course of the pandemic violate the assumption of a regular trend in mortality which underlies the calculation of reference periods for the estimates (Brass and Bamgboye 1981).

The only attempt to adapt orphanhood-based methods for use in countries facing HIV/AIDS epidemics dates back to the mid-1990s. Timæus and Nunn (1997) developed approximate expressions for the HIV-related selection biases and proposed an adjustment for the proportions of mothers alive. They also suggested a new set of coefficients to convert the adjusted proportions into survivorship probabilities for females, based on age-specific mortality rates reflecting the burden of AIDS. They warned, however, that these coefficients were provisional because they were based on prospective mortality data collected in a single rural community in Uganda (Asiki et al. 2013). In addition, their method did not account for antiretroviral therapy (ART), which was only introduced in the area in 2004 (Kasamba et al. 2012).

In this study, we assess the sensitivity of orphanhood-based methods to HIV-related bias using

a more diverse set of simulations than in the initial study by Timæus and Nunn (1997). We use microsimulated populations that model vertical transmission of HIV, reduced fertility of HIV-positive mothers and shifts in age patterns of mortality due to AIDS. The impact of antiretroviral therapy (ART), including the Prevention of Mother to Child Transmission (PMTCT), is also explicitly modelled. A new procedure for making the estimates is developed for use when at least two series of maternal orphanhood reports are available from successive surveys or censuses, in addition to estimates of HIV prevalence and treatment coverage. We develop this new approach in our simulated environment and evaluate it using survey and census data for 16 countries where the peak in HIV prevalence exceeded 5% in females (UNAIDS 2022)[1].

2 Data and methods

2.1 The conventional orphanhood method

The conventional orphanhood method is summarized below. We refer readers to Timæus (2013) for a detailed explanation and for Excel templates that facilitate the application of the method. The method is best expounded by starting from a child of age a, taken at random from a population whose fertility rates and survival function are $m(x)$ and $l(x)$ (Keyfitz and Caswell 2005). The probability that the mother is still alive, conditional on her having given birth at age x, is ${}_ap_x$. To eliminate this condition, the survival probabilities should be averaged over all reproductive ages, weighting each age x by the number of births that occurred at this age. In a stable population, the age distribution of the female population is constant and depends on l_x, m_x and r, the intrinsic growth rate. This leads to the equation introduced by Lotka (1931) for the probability that a child

[1]These countries are: Botswana, Cameroon, Central African Republic, Cote d'Ivoire, Eswatini, Kenya, Lesotho, Malawi, Mozambique, Namibia, Rwanda, South Africa, Tanzania, Uganda, Zambia, Zimbabwe. Trends in HIV prevalence and coverage of ART and PMTCT were extracted from Spectrum (https://www.avenirhealth.org/software-spectrum.php). We used version 6.23 based on the 2022 UNAIDS estimates (www.unaids.org). Gabon and Equatorial Guinea also experienced a severe HIV epidemic with a prevalence among women aged 15-49 exceeding 5%, but these two countries are excluded from the analysis due to the paucity of data on orphanhood.

aged a has a surviving mother under the prevailing conditions of mortality and fertility:

$$S(a) = \int_{\alpha}^{\beta} {}_{a}p_{x}\ {}_{x}p_{0}\ e^{-rx} m(x)dx \tag{1}$$

From this, the proportion of mothers surviving among those who gave birth to a child who is now aged y to $y + 5$ years can be expressed as:

$${}_{5}S_{y} = \frac{\int_{y}^{y+5} e^{-r(a)}\ {}_{a}p_{0} \int_{\alpha}^{\beta} e^{-r(x)}\ {}_{x+a}p_{0}\ m(x)dx\ da}{\int_{y}^{y+5} e^{-r(a)}\ {}_{a}p_{0} \int_{\alpha}^{\beta} e^{-r(x)}\ {}_{x}p_{0}\ m(x)dx\ da} \tag{2}$$

By specifying a series of fertility and mortality rates through standard age patterns, Eq. 2 can be used to approximate numerically the proportions ${}_{5}S_{y}$ (Brass and Hill 1973). The proportions can be connected to the probabilities ${}_{n}p_{M}$, calculated in the life tables from which they were generated, for example through linear regression, which yields a set of coefficients for each age group y. For convenience, M is often replaced by 25, a round number close to the mean age at childbearing among women and an estimate of the mean age at childbearing is included in the regression as a covariate to control for the actual timing of fertility. It can be obtained as the average age of women giving birth in a 12-month period about y years ago. Timæus (1992) used the following equation:

$${}_{n}p_{25} = \beta_{0}(n) + \beta_{1}(n)\ M + \beta_{2}(n)\ {}_{5}S_{n-5} \tag{3}$$

where ${}_{n}p_{25}$ is the chance of a woman surviving between age 25 and $25 + n$, M is the mean age at childbearing and ${}_{5}S_{n-5}$ is the proportion of respondents in the age group $n - 5$ to n whose mother is still alive. The β coefficients are presented in Table S2. Other regression equations have been tested and provide different sets of coefficients (Hill and Trussell 1977, United Nations 1983, Palloni and Heligman 1985). In the absence of HIV, the estimated probabilities ${}_{n}p_{25}$ are not very sensitive to the choice of the coefficients, especially when based on reports from young respondents ($n < 35$) (Masquelier 2010). Timæus (1992) also developed coefficients to estimate

men's mortality from paternal orphanhood.

The proportions of parents reported to be alive by adult respondents refer to mortality and fertility rates over a longer period than those derived from the reports of young children. A time to which the estimate refers should therefore be calculated. Existing time location procedures assume a linear trend in mortality levels, captured through the α parameter of the Brass logit system (Brass and Bamgboye 1981) or trends in life expectancy (Palloni and Heligman 1985). These procedures also assume a steady increase in mortality by age. Once time-located, the probabilities, ${}_np_{25}$, obtained from the reports of respondents in different five-year age groups $n-5$ to n need to be converted into a common index of mortality, such as the probability ${}_{35}p_{15}$, to be comparable and depict the general trend in mortality. This can be achieved either through relational models (such as the Brass logit model) or by interpolating within other families of model life tables (Coale et al. 1983, INDEPTH 2004).

2.2 Distortions due to HIV/AIDS

As mentioned earlier, the HIV epidemic undermines the validity of the conventional orphanhood method. Three important sources of bias exist. First, selection biases are magnified by the vertical transmission of the virus and the reduced fertility of seropositive women, which both inflate the proportions of mothers reported to be alive. Second, HIV epidemics generate atypical age patterns of mortality: the risk of dying rises rapidly with age across the early adult ages (between 15 and 30 years) and then increases more slowly with age than in standard model schedules. This leads the standard coefficients to overestimate survivorship. Third, increases in mortality due to AIDS, and the declines that have followed due to the uptake of ART and behavioural changes, violate the assumption that the trend in all-cause mortality is linear and unidirectional. Thus, the series of estimates made from respondents of different ages can no longer be interpreted as indicative of the period trend in mortality.

To address the first problem, Timæus and Nunn (1997) re-arranged Equation 2 to distinguish between seronegative mothers who remain uninfected, seropositive mothers at the time of birth, and mothers who become infected after the birth of their child. They proposed an adjustment to the observed proportions of mothers remaining alive. This adjustment is based on two parameters; F, the ratio of the fertility of seropositive to seronegative women (assumed to be age-invariant), and h, the proportion of children born to seropositive mothers who become infected in the perinatal period. In addition to h and F, an estimate of the prevalence of HIV infection among women attending prenatal clinics (P) is needed. The corrected proportions of mothers alive (${}_5S_n^{'}$) are obtained from the observed proportions (${}_5S_n^{*}$), such that:

$$ {}_5S_n^{'} = \frac{1 - hP}{1 + \frac{1-F}{F} \times P} \times_5 S_n^{*} \tag{4} $$

If an estimate of the HIV prevalence in the population is used (P^*), the equation becomes:

$$ {}_5S_n^{'} = [1 - (1 - (1 - h) \times F) \times P^{*}] \times_5 S_n^{*} \tag{5} $$

One can assume that the risk of mother-to-child transmission is about a third, and the fertility of seropositive women is about 75% of that of seronegative women (De Cock et al. 2000, Chen and Walker 2010). Thus, a suitable adjustment might be:

$$ {}_5S_n^{'} = [1 - 0.5 \times P^{*}] \times_5 S_n^{*} \tag{6} $$

This correction is easy to implement but assumes that all mothers who were already infected when the respondents were born died before the survey. Because this is unrealistic for young respondents, Timæus (2013) later recommended that the adjustment applied to reports from respondents aged 5-9 be halved ($1 - 0.25 \times P^*$) and reduced by a quarter for 10- to 14-year-olds ($1 - 0.375 \times P^*$).

To address the second problem, related to distortions introduced in age patterns of mortality, Timæus and Nunn (1997) developed simulations based on prospective mortality data from the Masaka Health and Demographic Surveillance System (HDSS) in Uganda in 1990-95. Their simulations were based on stable population theory, assuming that the characteristics of the HIV epidemic at that time are kept constant. These simulations allowed them to compute a set of coefficients that can be used to convert the proportions of mothers alive into life table survivorship estimates when HIV prevalence is 5% or greater. These coefficients are reproduced in Table S3. The third problem, associated with the time trend in mortality, has not been explicitly addressed in the literature. Most attempts to estimate mortality from orphanhood in settings with high HIV have ignored this problem and used time-location procedures that assume a smooth and unidirectional trend in mortality (e.g. Feeney (2001), Hosegood et al. (2004), Nhacolo et al. (2006), Menashe Oren and Stecklov (2018), Odimegwu et al. (2018)). Estimates produced in this way will inevitably smooth out the sudden reversals and accelerations in mortality trends to be expected in populations experiencing a generalized HIV epidemic.

2.3 The microsimulation set

To produce a more confident assessment of the magnitude of the HIV-related biases in orphanhood-based estimates than Timæus and Nunn (1997) obtained by an analytic approach, we resorted to demographic microsimulations. These are models in which individuals experience vital events as a result of stochastic experiments with pre-defined probabilistic rules (Zagheni 2015). We used SOCSIM, a discrete-time microsimulation model that keeps track of kinship links between individuals (Wachter et al. 1997, Verdery et al. 2020). The simulations were run by periods, during which all parameters are kept constant; the first period corresponds to conditions of a stable population and lasts from year 0 to 200. Populations reach about 100 000 individuals at that point. Ten periods of 5 years follow, during which the population face an HIV epidemic. The growth rate during these 50 years evolves according to the severity of the epidemic; some

populations reach 150 000 individuals in the year 250, while others decline to 85 000 survivors.

For the period preceding the onset of the HIV epidemic, the mortality and fertility rates are similar to those used by Timæus (1992), although fewer parameters are retained to allow for the introduction of additional parameters related to HIV while limiting the number of simulations and the computational burden of producing them. The populations were exposed to various levels of non-AIDS mortality, modeled with the Brass relational model, by specifying three values for the α_m parameter (capturing variations in the level of mortality) and two values for the β_m parameter (capturing differences in age patterns), using Brass' general standard (Brass 1971). Fertility was also modeled with a relational model, with two values of α_f (capturing the age location of the fertility schedule) and two values of β_f (capturing the spread of the fertility schedule) (Brass 1974). The standard used for the fertility schedule was created by Booth (1984) for populations with high fertility. The waiting time to each event was generated randomly from a piecewise exponential distribution. The stable-equivalent population obtained analytically from the survival curve, the shape of the fertility schedule and the growth rate was used to specify the age structure of the starting population. The initial growth rate was set at 2% and kept constant until the onset of the epidemic, to be consistent with the method's original calculations (Hill and Trussell 1977, Timæus 1992). Using a single value of the growth rate is adequate as variations in age structure have little effect on mortality estimates derived from orphanhood (Timæus 1992). The corresponding non-AIDS life expectancies at birth range from 43.1 to 68.3 years, while the mean ages at childbearing range from 25.1 to 29.8 years.

The parameters used to model the HIV epidemic are inspired by those underpinning the UN-AIDS Spectrum package (Stover et al. 2012; 2017). Following Zagheni (2011), HIV/AIDS was modelled in SOCSIM by splitting the population into sub-groups: (i) seronegative individuals, (ii) those that have been infected with HIV through sexual transmission but are in the asymp-

tomatic phase, (iii) those who have developed acquired immunodeficiency syndrome (AIDS), (iv) children infected with HIV through vertical transmission, and (v) individuals who have initiated antiretroviral treatment (Figure S1). HIV infection is governed by age-specific rates, depending on trends in incidence. A gamma distribution was used to impose a plausible shape on the HIV incidence curve (Heuveline 2003). It depends on two parameters α_{hiv} and β_{hiv}, as follows (Clark et al. 2012):

$$\Gamma_{t_2-t_1} = \int_{t_1}^{t_2} \frac{x^{\alpha_{\text{hiv}}-1} e^{-x/\beta_{\text{hiv}}}}{(\alpha_{\text{hiv}}-1)!\beta_{\text{hiv}}^{\alpha_{\text{hiv}}}} dx \tag{7}$$

This curve defines the trend in HIV incidence between times t_1 and t_2; the scale of the epidemic was determined by an additional parameter H, such that the proportion of individuals that are uninfected at time t_1 and alive and HIV-positive at time t_2 is:

$$i_{t_1} = 1 - exp\left\{-\Gamma_{t_2-t_1} H\right\} \tag{8}$$

The age distribution of HIV infections was obtained as an average of patterns derived from cross-sectional measurements of HIV prevalence in DHS (Stover et al. 2010). Once infected, individuals remain exposed to background mortality and progress to AIDS according to a Weibull distribution (with a median time from infection to AIDS of 8.55 years for females) (Fig. S2). These progression rates were used in the Spectrum program before the approach was revised to accommodate changes in the criteria for eligibility for ART (Stover 2009). In our simulations, the transition to AIDS is governed solely by the time since infection. The transition from the AIDS stage to death was also modelled with a Weibull distribution with a median time from AIDS to death of 1.95 years. The fertility of infected women relative to uninfected women was fixed at 1.26 for women aged 15 to 19, 0.76 for women aged 20 to 24, and 0.67 for those aged 25 or more, based on ratios between age-specific fertility in HIV-positive women and HIV-negative women observed in 19 community-based studies (Lewis et al. 2004).

The probability of vertical transmission *in utero* or during delivery for a child born to an HIV-positive mother was assumed to be 20% in the absence of treatment (Stover et al. 2012). The probability of infection through breastfeeding varies by age of the child and depends on an average proportion of children who received exclusive or mixed breastfeeding. The survival of HIV-infected children is defined by a double Weibull curve and varies with the age at infection, based on Stover et al. (2012). Based on data from 16 countries where the prevalence of HIV reached at least 5% according to UNAIDS, we developed three scenarios for the expansion of the coverage of treatment with ART and PMTCT: rapid treatment scale-up, slow treatment scale-up, and no treatment (see appendix). In the simulations with treatment, the proportion of children infected vertically was revised downwards based on PMTCT treatment trends. We assumed that there is no dropout from ART treatment, that patients on ART also benefit from PMTCT and that the fertility of women receiving ART is similar to that of those who have not yet initiated treatment. The probability of survival on ART is fixed at 85% in the first year, and 95% for each additional year on treatment (Stover et al. 2008). Adults on ART are also exposed to mortality from other causes. All the simulations were run twice, once with and once without vertical transmission and reduced fertility due to HIV.

The values of the main parameters used to set up these simulations are shown in Table 1. Their combination results in 1152 different simulations. Figure 1 presents the following model inputs: (1) the survival curves of the non-AIDS life tables, (2) the fertility schedules, and (3) the incidence curves. It also displays the resulting trends in HIV prevalence among women aged 15 to 49.

3 Results

3.1 Effect of the HIV epidemic on orphanhood prevalence and adult mortality

As a calibration exercise, Figure 2a presents trends in orphanhood prevalence among children aged 5-9 years and 10-14 years in one simulation set, selected to broadly reflect the HIV epidemic

in Zimbabwe. For this illustration, the years of the simulation have been recoded such that the onset of the epidemic is around 1977. Proportions of orphaned children aged 5-9 and 10-14 observed in surveys and censuses in Zimbabwe are represented by blue circles[2]. In Zimbabwe, the prevalence of maternal orphanhood among 5-9 year olds increased from 2% in the 1982 census to 9% in the 2009 MICS, then declined to 3% in the 2019 MICS. The orphan prevalence recalculated from the simulation follows a similar trend, hovering around the expected prevalence in the stable equivalent population until the mid 1980s (2%), then increasing rapidly to peak at 10% in 2006, before declining to reach 3% in 2019.

In this simulation, AIDS-free life expectancy at birth is 68 years for women and the risk of dying between the ages of 15 and 50 is 89 per thousand before the HIV epidemic unfolds (Figure 2b). This probability then increases to reach 513 per thousand in 2004. Sibling histories collected in the DHS surveys in Zimbabwe depict a similar mortality increase, from 142 per thousand in the 1994 DHS to 443 per thousand in the 2005 DHS (Central Statistical Office Zimbabwe and Macro International 1995; 2007). The prevalence of HIV infection among women of reproductive age peaks at 23% in this simulation in 2001, which is close to the prevalence measured in the 2005-2006 DHS (21.1%). According to UNAIDS, the peak in prevalence in women aged 15-49 was earlier and higher, in 1996, at 28.3%. The simulation, therefore, does not reproduce the evolution of the HIV epidemic in Zimbabwe, but follows a similar enough pathway to be used here for illustrative purposes. Figure S4 compares the proportions of maternal orphans in all DHS conducted in high-HIV countries to those observed in the microsimulation set, as well as adult and child mortality rates in these two series. These comparisons suggest that simulations are well calibrated.

[2]These estimates were extracted from the 1982, 1992, 2002 and 2012 censuses, the 1997 Intercensal Survey, the 1994, 1999, 2005-2006, 2010-2011, and 2015 DHS, as well as the 2009, 2014, and 2019 MICS surveys.

3.2 Estimates obtained from one survey/census and standard coefficients

We first examine HIV-related biases in the orphanhood estimates obtained from the conventional method. Proportions of maternal orphans classified by five-year age groups are computed from simulations as if a survey or census had been conducted every five years. Indirect estimates are obtained using the coefficients developed by Timæus (1992) and time-located following Brass and Bamgboye (1981). The "true" mortality rates are obtained by dividing the number of deaths by age, sex and year by the corresponding exposure, using the exact dates of birth and death of all females who ever lived in the simulation. Fig. 2(c) is based on the simulation that approximates the trends observed in Zimbabwe. It confronts the indirect mortality estimates obtained from maternal orphanhood in respondents aged 5-9 and 15-19 (dashed lines) with the "true" mortality rates (solid lines). The blue circles correspond to estimates derived from surveys and censuses in Zimbabwe. In the simulations, the indirect estimates agree well with the direct measures in the pre-HIV period. After the onset of the epidemic, the probability of dying between exact ages 25 and 35 ($_{10}q_{25}$) inferred from orphanhood starts to deviate from the underlying mortality rates. It is underestimated by as much as 58% in 2004. The probability $_{20}q_{25}$, based on respondents aged 15-19, is also substantially biased, with a 44% underestimate in 2004. To quantify the magnitude of the errors across all simulations with different mortality rates and compare age groups, we compute the ratio of the odds of surviving according to the indirect estimates to the "true" odds of surviving, as follows:

$$\frac{{}_{n}p_{25}^{\text{indirect}}}{1 -_{n} p_{25}^{\text{indirect}}} \times \frac{1 -_{n} p_{25}^{\text{true}}}{{}_{n}p_{25}^{\text{true}}} \tag{9}$$

Fig. 2(d) displays the median ratios for the 576 simulations with reduced fertility and vertical transmission (see also Table S4). As in the Zimbabwe example, the indirect estimates are close to the underlying mortality rates before HIV is introduced; the median ratios range between 1.00 and 1.05. (Small deviations are to be expected as we did not use exactly the same parameters as in Timæus (1992) to build the simulations). Once HIV is introduced, the median ratios first decline below 1, indicating that mortality is overestimated in the 10 to 15 years following the onset

of the epidemic. This is because trends are overly smoothed by the conventional time location procedure, estimates for the beginning of the epidemic are capturing some of the subsequent mortality increase. This effect is most pronounced when estimating ${}_{20}p_{25}$ (median ratios = 0.93, 7 years after the onset of HIV). As the epidemic unfolds, the median ratios rise substantially for all survivorship probabilities, before shrinking as the epidemic recedes. The bias is larger when estimates are inferred from reports from younger respondents, but the errors are still substantial for the older age groups. The odds of surviving between ages 25 and 35 are overestimated by as much as 85% about 30 years after the onset of the epidemic, against 46% for the probability ${}_{45}p_{25}$.

3.3 Existing adjustments for HIV-related bias

We detailed above the different adjustments developed by Timæus and Nunn (1997). Figure 2(e) shows the risks of dying obtained from the simulation resembling Zimbabwe and from surveys or censuses conducted in the country, after applying these adjustments when HIV prevalence at the time of birth is higher than 5%. The indirect estimates of mortality are now much closer to the underlying mortality for the youngest respondents (5-9 years), but seem too high in the most recent periods. In addition, the indirect estimates for ${}_{20}q_{25}$ remain considerably lower than the underlying mortality rates. To generalize over all simulations with reduced fertility and vertical transmission, Figure 2(f) displays the median ratios of the estimated to the true life table odds of surviving. For the youngest age group (n = 10), the errors are substantially reduced compared to estimates obtained without any adjustment for HIV. However, the odds of survival are overestimated about 20 years after the onset of the epidemic, before being underestimated, with median ratios approaching 0.6. Biases are much reduced for the probabilities ${}_{15}q_{25}$ to ${}_{20}q_{25}$, but the latest ratios are also lower than 1. For the older age groups, biases are similar to those observed with the original method. This is because at the time Timæus and Nunn (1997) conducted their study, Uganda was only about 15 years into its HIV epidemic and no evidence existed as to how large

an impact it would have in future on the mortality of older women.

The remaining errors might have three sources, again related to bias in proportions of mothers surviving, the conversion of proportions into life-table survivorship and the procedure for estimating mortality trends. First, the adjustment for proportions of surviving mothers does not incorporate the effects of PMTCT on the risk of vertical transmission. The effect of ART on fertility and survival in HIV-positive women is not accounted for either. The adjusted proportions of surviving mothers will therefore be too low once treatment has been scaled-up. Figure 3 illustrates this. The left-hand plot shows proportions of respondents with surviving mother in a simulation that has the same parameters as the one resembling Zimbabwe, except that nobody receives ART. The right-hand plot shows the same proportions, with treatment coverage reaching 90%. The dashed lines refer to the proportions that, according to the simulations, one would observe in surveys, while the solid lines represent the unbiased proportions. These unbiased proportions are obtained without any HIV-related bias and are computed in simulations with exactly the same parameters, except that we disabled vertical transmission and reduced fertility of HIV-infected mothers. The proportions adjusted as suggested by Timæus and Nunn (1997) are displayed with triangles. In the absence of treatment (Figure 3(a)), the adjusted proportions are consistent with the unbiased proportions, suggesting that vertical transmission and reduced fertility could indeed be accounted for in this way before ART and PMTCT were introduced. By contrast, in the simulation that includes ART and PMTCT (Figure 3(b)), the adjusted proportions are found to be too low. Timæus and Nunn (1997) had to assume that all HIV-positive mothers die when their child reaches 5 years of age to develop their adjustment, and then reduced the adjustment by a fixed age factor to account for the fact that many of the mothers of young children would still be alive at the time of data collection. However, this procedure is no longer adequate as the bias cannot be approximated based solely on h and F once a substantial proportion of infected women are receiving ART.[3].

[3]This is because the numerators of equations (1) and (2) in Timæus and Nunn (1997) are no longer identical.

A second source of bias is that the coefficients developed in the 1990s were based on the limited evidence available at that time and also do not account for the introduction of treatment, which became available later. Timæus and Nunn (1997) used prospective mortality data from a HDSS where the HIV prevalence among the adult population was about 8%, a level well below the peak of HIV prevalence observed since in several countries. The age pattern of HIV prevalence was also assumed to be fixed and seroprevalence peaked in women in their mid-twenties. In our simulations, the age pattern of HIV varies over time. In the first 10 years after HIV is introduced, prevalence peaks among women aged 20-24, but it gradually shifts to older ages and by 20 years into the epidemic peaks among women aged 25-29 in about half of the simulations. In addition, the scale-up of ART changes age patterns of mortality. Since the mid-2000s, the coverage of ART has dramatically increased in Sub-Saharan Africa. By 2021, it had reached 82% in West and Central Africa and 79% in East and Southern Africa (UNAIDS 2022). As noted earlier, a third source of bias is that deriving a series of dated estimates from a single set of proportions is inappropriate when mortality trends have been highly disrupted.

3.4 A revised method based on two sets of proportions of mothers alive

To address these three sources of errors, we first re-examined the relationship between the unbiased proportions and the proportions affected by fertility reduction and vertical transmission. In Figure S5 in appendix, the three graphs on the left present the ratios of the unbiased to the observed proportions, according to HIV prevalence at birth, without any adjustment to the proportions, based only on simulations without treatment (S_n/S_n^*). Bias is linearly associated with prevalence, as established previously by Timæus and Nunn (1997). The coefficient of the slope of the linear regression is -0.248 for the 5-9 age group, -0.378 for the 10-14 age group, and -0.477 for the 15-19 age group. These coefficients are remarkably consistent with the adjustments that Timæus and Nunn (1997) developed analytically (i.e. $1 - 0.25 \times P$, $1 - 0.375 \times P$, and $1 - 0.5 \times P$, see section

Using their formulation, $N^+(a)$ (corresponding to the number of living HIV-positive women who have given birth a years ago if their fertility was the same as other women) is no longer zero in the presence of the treatment.

2.2). In the middle panel, we present the ratios based on proportions after using such adjustments ($S_n/S_n^{'}$), but considering now all simulations, including those with treatment. These ratios are too high. Moreover, for the first age group, the error is directly related to PMTCT coverage at the time of birth. For subsequent age groups, the magnitude of the bias is associated with ART coverage at the time of the survey.

We tested seven regression models to predict the bias in the proportions based on covariates that are made available by UNAIDS for all countries. We included as covariates, various combinations of HIV prevalence, PMTCT and ART coverage measured at time of the survey or at birth (Table S5). The dependent variable was the ratio of the unbiased to the observed proportions. All the models were run separately for each age group. To evaluate model performance, we randomly selected 80% of the simulations to fit the models and calculated the out-of-sample RMSE based on predictions in the remaining simulations. The prediction errors are displayed in Table S5. Across age groups, the best-performing model was as follows:

$$\frac{{}_5S_n}{{}_5S_n^*} = \beta_0(n) + \beta_1(n)\ [HIV_{t-n+2.5} \times (1 - PMTCT_{t-n+2.5})] + \beta_2(n)\ ART_t \tag{10}$$

where $HIV_{t-n+2.5}$ and $PMTCT_{t-n+2.5}$ are HIV prevalence and PMTCT coverage respectively, at the time when the respondents were born, while ART_t refers to ART coverage at the time of the survey. The β coefficients are presented in Table 2. The ratios obtained from proportions corrected by means of these coefficients are displayed in the right-hand panel ($S_n/S_n^{''}$) of Figure S5.

Once the proportions have been corrected for HIV-related bias in the reports, they can be related to the underlying mortality levels, but this raises the question of the reference period to consider. It does not seem possible to develop a new procedure for dating the estimates that captures the diversity of temporal changes in mortality in HIV/AIDS-affected settings or that adequately addresses the built-in tendency of the lifetime proportions to smooth out abrupt dis-

continuities in mortality trends. To circumvent this problem, one can use two sets of proportions from successive surveys or censuses to construct synthetic cohorts. The construction of synthetic cohorts was first proposed by Zlotnik and Hill (1981), who suggested chaining together changes experienced by a given age cohort (in terms of the survival of their parents) during the intercensal or intersurvey period. Alternatively, Preston (1987) proposed working with the changes experienced by a given age group between two censuses or surveys. A correction factor is computed based on the growth rate of the proportion of parents alive. This provides the proportion of parents alive that pertains to the intercensal period and would be observed in a stationary population. This approach was originally developed to generate estimates that refer to more recent periods than the cohort measures and that are less biased by the under-reporting of parental deaths (Timæus 1986). In contexts affected by HIV, synthetic cohorts offer the additional advantage that they do not require any assumption made about the trend in mortality prior to the collection of the first set of data.

Proportions for a hypothetical cohort can be obtained by adjusting each set of proportions for HIV-related bias based on Equation 10, and chaining successive sets as suggested by Preston (1987). The remaining task is then to convert these proportions into life table survivorship estimates. Using microsimulations, Figure 4a presents the relationship between the probabilities ${}_{10}p_{25}$ and proportions ${}_5S(h)_5$, in simulations without treatment. The simulation highlighted with large dots has the same parameters as the one that resembles Zimbabwe, but without any treatment. When HIV prevalence is less than 5%, the relationship between the two series is well represented by the coefficients from Timæus (1992) (straight green line, for a value of M of 25). However, as HIV prevalence increases, the slope of the regression line steepens, as already demonstrated by Timæus and Nunn (1997). They thus recalculated the intercept and slope of the regression line, but without introducing HIV as a covariate. This will lead to discontinuities in the trends when shifting from one set of coefficients to the other, due to the large difference in their intercepts (orange line in Fig. 4a). It has also become important to incorporate ART among the covariates

as the slope of the regression line declines again when ART coverage increases, as illustrated in Figure 4b.

To calculate new coefficients, we evaluated different regression models predicting the survivorship probabilities from the simulated proportions, testing combinations of predictors including HIV prevalence and ART coverage, measured at the midpoint of the two surveys or at the time of birth of respondents (Table S6). We also evaluated the predictive performance of models including $HIV \times (1 - ART)$, which captures the percentage of women who are seropositive and have not initiated treatment. Because the synthetic proportions are based on changes over time in the proportions of mothers surviving, in some models we included the absolute difference in HIV prevalence or in $HIV \times (1 - ART)$ between the two surveys. Some models were fitted over the full dataset, others used two regressions, one on the data points without any treatment, and one for data points with a least some women on ART. This helps capture potential changes in age patterns of mortality in growing and receding pandemics, as the early phases of the roll-out of ART correspond in most countries to the start of the decline in adult mortality (Reniers et al. 2014). Across the age groups, the best-performing approach used separate models for the pre-ART period and the period after treatment programmes were introduced:

$$
\begin{aligned}
{}_{n}p_{25} &= \{\beta_0(n) + \beta_1(n)\ M + \beta_2(n)\ {}_5S(h)_{n-5} + \beta_3\ HIV_t + \beta_4\ \Delta HIV\}\ (ART_t = 0) \\
&= \{\beta_5(n) + \beta_6(n)\ M + \beta_7(n)\ {}_5S(h)_{n-5} + \beta_8\ [HIV_t \times (1 - ART_t)] + \beta_9\ \Delta[HIV \times (1 - ART)]\} \\
&\quad (ART_t > 0)
\end{aligned}
\tag{11}
$$

Here HIV_t and ART_t refer to the HIV prevalence and ART coverage at the time of data collection and are obtained as the average of estimates from each survey (since two series are required to calculate synthetic proportions). The indicators of trends, Δ HIV and $\Delta[HIV \times (1 - ART)]$ refer to the absolute difference between the two surveys in the measures. The corresponding β coefficients are displayed in Table 3. An Excel template is available in the Supplementary

materials to facilitate the calculations.

As with other orphanhood-based methods, this new variant is based on approximating the proportion of surviving mothers by the proportion of children whose mothers are survivors. This entails assuming that maternal survival remains unaffected by both child survival and fertility, and that child survival is independent of sibship sizes. After correcting for biases associated specifically with HIV/AIDS, any resultant selection biases arising from violations of these assumptions are expected to be minimal (Palloni et al. 1984). When combining age-specific estimates of ${}_np_{25}$ into a summary index, an additional assumption is required: that the chosen model life table reflects the underlying age pattern of mortality. We suggest converting all age-specific estimates into ${}_{35}p_{15}$ using a model life table incorporating mortality attributable to AIDS (INDEPTH 2004) and averaging the resulting probabilities.

3.5 Orphanhood-based estimates of adult mortality for selected countries in Sub-Saharan Africa

In this section of the article, we test the different estimation approaches on real data for 16 countries where the peak HIV prevalence exceeded 5% (UNAIDS 2022). We extracted proportions of surviving mothers by age group from censuses, DHS and MICS, and other nationally-representative surveys (Appendix H) and constructed two sets of estimates:

- The first set considers each census or survey separately, and used standard coefficients when HIV prevalence was less than 5%, and those developed by Timæus and Nunn (1997) when it was greater than 5%. The proportions of surviving mothers were adjusted for HIV/AIDS bias based only on HIV prevalence. The reference periods were calculated using the method developed by Brass and Bamgboye (1981). This series corresponds to the procedure generally followed in previous research.

- The second set adjusts the proportions for HIV-related bias using the coefficients in Table 2, and chains successive sets of proportions together to construct synthetic cohorts. We combined inquiries separated by at least 3 years and less than 11 years to avoid irregularities

associated with very short intervals and to allow the combination of successive censuses, typically conducted every 10 years. Synthetic proportions were converted into life table estimates using the coefficients derived from microsimulations (Table 3), combined with estimates of HIV prevalence and ART coverage from UNAIDS (2022). The resulting mortality estimates are available in the supplementary materials.

In the absence of a gold standard, Figure 5 contrasts the orphanhood-based estimates with mortality rates from the *World Population Prospects* (WPP) (United Nations 2022), after interpolating these rates to obtain a value referring to the same time period. The country-specific estimates of the probabilities ${}_{10}q_{25}$ and ${}_{15}q_{25}$ calculated with the new approach are displayed in Figure 6 (see Figure S6 in appendix for the first set of estimates based on a single survey or census). Orphanhood data are one input into the existing WPP estimates of mortality. Thus, the WPP estimates and ours are not entirely independent. In most of the 16 countries though, the documentation available suggests that direct estimates calculated from sibling histories and questions about recent deaths in the household and the comparison of successive census counts had more influence on the WPP estimates than the orphanhood data. In Figure 6, mortality estimates are also compared to those extracted from sibling survival histories collected in DHS for the 6 years prior to each survey (Masquelier et al. 2014). We focus on the probabilities ${}_{10}q_{25}$ and ${}_{15}q_{25}$ because these are obtained from respondents aged 5-9 and 10-14. More observations are available on these age groups than older ones since DHS and MICS surveys do not ask about parental survival among adults.

According to the patterns observed in simulations, we expect the first set of mortality rates to be overestimated based on reports from children aged 5-9 when mortality is high, and underestimated when based on older respondents. This is indeed what is observed in Figure 5(a); the probability ${}_{10}q_{25}$ inferred from maternal survival tends to be below the levels predicted by the WPP when the risk of dying is lower than 0.07, while it tends to be higher when mortality increases above

this threshold. The overestimation is likely reduced because of the adoption bias, which was not introduced in the simulations, and will predominantly affect the estimates from a single survey or census with the proportions in synthetic cohorts being less affected. Overall, the median ratio between probabilities ${}_{10}q_{25}$ derived from orphanhood and WPP estimates is 1.26. This ratio drops below one (0.83) when mortality is lower than 0.07 in WPP, and reaches 1.28 when mortality is higher. This pattern is also visible in country-specific plots in Figure S6. There are fewer data points to evaluate the bias on the ${}_{15}q_{25}$ probability, but it appears to be consistently lower than the WPP estimates. The median ratio between orphanhood-based probabilities and WPP estimates is as low as 0.54 for this age group. This is likely due to a combination of recall and modelling biases.

Estimates obtained with the revised estimation approach do not refer to the same periods; there are fewer estimates available from respondents aged 5-9 years since it is necessary to combine surveys or censuses, but slightly more from those aged 10-14 years (since the Brass' dating method, which requires values up to age 19, is no longer used). Overall, there is a better congruence with WPP than in the previous set of orphanhood estimates: the median ratio between orphanhood-based rates and WPP is 0.99 for ${}_{10}q_{25}$ and 0.91 for ${}_{15}q_{25}$ (compared to 1.26 and 0.54 with the previous set of orphanhood estimates). Moreover, these median ratios remained stable over time: for the ${}_{10}q_{25}$ probability, the ratio is 1.05 for the pre-ART period and 0.98 for the post-ART period, while the corresponding indices for the ${}_{15}q_{25}$ probability were 0.91 at both periods. The country-specific trends suggest that the new estimates track the WPP probabilities of death remarkably well (Figure 6). This is the case in Malawi, Namibia, Rwanda, South Africa, Tanzania, Uganda, Zambia, and Zimbabwe. The rise and fall in probabilities are well reproduced, and the difference between the ${}_{10}q_{25}$ and ${}_{15}q_{25}$ probabilities is comparable to what is predicted by the WPP. In most countries (except Zimbabwe in the 1990s and 2000s), the orphanhood estimates are also in line with the probabilities of death from sibling histories.

4 DISCUSSION

Our results demonstrate the large impact of HIV-associated biases on mortality levels inferred from maternal survival data. We build on previous work by Timæus and Nunn (1997) by generalizing the series of simulations and relaxing the assumption of a stable population. We show that, without any adjustment, the conventional orphanhood method will overestimate mortality in the first few years following the onset of an HIV epidemic and then substantially underestimate mortality as the epidemic matures. Biases are considerably reduced when using the adjustments developed by Timæus and Nunn (1997), but they can be further reduced by incorporating information on HIV prevalence and PMTCT and ART coverage. The construction of synthetic cohorts also avoids the heavy smoothing of abrupt changes in the trend in mortality imposed by the basic orphanhood method. This new variant makes use of prevalence and treatment trends, so unlike estimates from sibling survival histories or recent household deaths, it does not provide a direct measure of mortality. The resulting estimates can, however, complement other series of primary estimates to help better reconstruct mortality trends during the epidemic. When applied to survey and census data from 16 countries in SSA, this new approach provides estimates that better reflect the timing of the epidemic and are more consistent with expectations based on the WPP and sibling histories. Once treatment programmes are scaled-up to the point that HIV-positive adults benefit from the same survival chances as HIV-negative adults, and vertical transmission is suppressed in the general population, no correction will be needed. However, reports on parental survival being collected currently are still affected by selection biases. Moreover, estimates based on existing data need to be made with methods that adequately reflect mortality patterns before, during and after the scale-up of ART.

The methods proposed in this study have some limitations. First, the adjustments proposed here are designed for estimating women's mortality only. Mortality estimates for men made from paternal orphanhood data will also be biased because fathers can infect or be infected by mothers,

who themselves can transmit the virus to their children. Unfortunately, due to the complexity of modelling the concordance of the HIV status of mothers and fathers and the impact of HIV on men's fertility, no adjustment has been developed yet for men's mortality. One avenue for future development in this regard might be to leverage the sophisticated modelling of paternal orphanhood in the Spectrum package (Grassly and Timæus 2005), but this would require revising this software to also produce outputs on orphanhood in youth and adults. Second, this study is limited by the assumptions made to model the demographic impact of HIV. For example, treatment allocation was done randomly at intervals of 5 years, and the mortality of orphans was assumed to be the same as that of children with living parents (apart from vertical transmission). Most of the model inputs for the HIV epidemic, such as HIV survival and the age pattern of incidence, were considered fixed. Third, and perhaps more importantly, we only focused here on selection biases, leaving aside other possible reporting biases. The most pervasive problem is the adoption effect. The potential magnitude of this adoption bias can be gauged from a cohort study conducted in Manicaland (Zimbabwe). Robertson et al. (2008) analysed the consistency of reporting of parent survival status across successive rounds and found that, out of 198 children reported as maternal orphans in the first round (and followed up to the third round), as many as one third were reported as non-orphans at least once in the next two rounds. A second problem is caused by non-responses. Although the proportions of missing data on questions about orphanhood are usually rather low, they can be of the same order of magnitude as the proportions of young children that are orphans. One thus needs to make assumptions about the orphanhood status of children with missing data. Finally, ages reported in censuses and surveys can be affected with inaccuracies, such as age exaggeration and heaping on round digits. Because of these different sources of error, estimates from orphan data should always be viewed with caution and compared to other sources. Nevertheless, in this study, when they were compared with sibling survival data, they provided comparable estimates of mortality after making the adjustments proposed here for HIV-related bias.

Despite these limitations, this study demonstrates that parental survival data remain useful for estimating mortality in countries lacking a complete death registration system. Provided disaggregated prevalence and treatment data are available, the new variant of the orphanhood method we developed could also be used to study mortality differentials in settings affected by HIV. More research is needed to investigate recall errors and develop ways of adjusting for bias in paternal orphanhood data. Statistical models could also be developed to combine orphanhood estimates with data from recent household deaths in censuses and from sibling histories, similarly to well-established models used for child mortality (Alkema and New 2014). Data on parental survival should be more systematically collected as they can help fill important data gaps in Sub-Saharan Africa and track progress against the HIV epidemic. In DHS and MICS surveys, the questions should be extended to adult respondents as well as children under 18. Additional questions could also be asked in surveys about the ages of living parents or ages at death or, most promisingly, dates of death for parents who have died, to allow more direct calculation (Chackiel and Orellana 1985). Although HIV/AIDS-related mortality has declined significantly in recent decades, the epidemic is far from over, and the true extent of excess mortality associated with this epidemic remains difficult to assess to this day.

References

Alkema, L. and New, J. 2014. Global estimation of child mortality using a Bayesian B-spline bias-reduction model. *The Annals of Applied Statistics*, 8(4):2122–2149. https://doi.org/10.1214/14-AOAS768.

Asiki, G., Murphy, G., Nakiyingi-Miiro, J., Seeley, J., Nsubuga, R. N., Karabarinde, A., Waswa, L., Biraro, S., Kasamba, I., Pomilla, C., Maher, D., Young, E. H., Kamali, A., and Sandhu, M. S. 2013. The general population cohort in rural south-western Uganda: a platform for

communicable and non-communicable disease studies. *International Journal of Epidemiology*, 42(1):129–141. https://doi.org/10.1093/ije/dys234.

Blacker, J. and Mukiza-Gapere, J. 1988. The indirect measurement of adult mortality in Africa: results and prospects. In International Union for the Scientific Study of Population, editor, *African Population Conference, Dakar, Senegal, November 7-12, 1988*, pages 23–38. IUSSP.

Booth, H. 1984. Transforming Gompertz's function for fertility analysis: the development of a standard for the relational Gompertz function. *Population Studies*, 38(3):495–506. https://doi.org/10.2307/2174137.

Brass, W. 1971. On the scale of mortality. In Brass, W., editor, *Biological Aspects of Demography*, pages 69–110. Taylor et Francis, London.

Brass, W. 1974. Perspectives in population prediction: illustrated by the statistics of England and Wales. *Journal of the Royal Statistical Society*, 137:532–583. https://doi.org/10.2307/2344713.

Brass, W. and Bamgboye, E. 1981. The time location of reports of survivorship: estimates for maternal and paternal orphanhood and the ever-widowed. Centre for Population Studies Research Paper, 81-1.

Brass, W. and Hill, K. 1973. Estimating adult mortality from orphanhood. In International Union for the Scientific Study of Population, editor, *Proceedings of the International Population Conference: Liège, Aug. 27 - Sept. 1, 1973*, volume 3, pages 111–123. IUSSP.

Central Statistical Office Zimbabwe and Macro International. 1995. *Zimbabwe Demographic and Health Survey 1994*. Calverton: Central Statistical Office/Zimbabwe and Macro International.

Central Statistical Office Zimbabwe and Macro International. 2007. *Zimbabwe Demographic and Health Survey 2005-06*. Calverton: Central Statistical Office/Zimbabwe and Macro International.

Chackiel, J. and Orellana, H. 1985. Adult female mortality trends from retrospective questions about maternal orphanhood included in census and surveys. In International Union for the Scientific Study of Population, editor, *International Population Conference, Florence, Italy, June 5-12, 1985*, volume 4, pages 39–51. IUSSP.

Chen, W.-J. and Walker, N. 2010. Fertility of HIV-infected women: insights from Demographic and Health Surveys. *Sexually Transmitted Infections*, 86(Suppl 2):ii22–ii27. https://doi.org/10.1136/sti.2010.043620.

Chisumpa, V. and Dorrington, R. 2011. Estimating adult mortality in Zambia using information on survival of parents from surveys. *African Population Studies*, 25(1):113–126. https://doi.org/10.11564/25-1-259.

Clark, S., Thomas, J., and Bao, L. 2012. Estimates of age-specific reductions in HIV prevalence in Uganda: Bayesian melding estimation and probabilistic population forecast with an HIV-enabled cohort component projection model. *Demographic Research*, 27(26):743–774. https://doi.org/10.4054/demres.2012.27.26.

Coale, A., Demeny, P., and Vaughan, B. 1983. *Regional Model Life Tables and Stable Populations (2nd ed.)*. New York: Academic Press.

De Cock, K., Fowler, M., Mercier, E., de Vincenzi, I., Saba, J., Hoff, E., Alnwick, D., Rogers, M., and Shaffer, N. 2000. Prevention of mother-to-child HIV transmission in resource-poor countries: translating research into policy and practice. *Journal of the American Medical Association*, 283(9):1175–1182. https://doi.org/10.1001/jama.283.9.1175.

Dorrington, R., Moultrie, T., and Timæus, I. M. 2004. *Estimation of Mortality using the South African Census 2001 Data*. Cape Town: Centre for Actual Research.

Ewbank, D. 1981. *Age Misreporting and Age-selective Underenumeration: Sources, Patterns and*

Consequences for Demographic Analysis. Washington, DC: The National Academies Press. https://doi.org/10.17226/19649.

Feeney, G. 2001. The impact of HIV/AIDS on adult mortality in Zimbabwe. *Population and Development Review*, 27, 4:771–780. https://doi.org/10.1111/j.1728-4457.2001.00771.x.

Grassly, N. and Timæus, I. M. 2005. Methods to estimate the number of orphans as a result of AIDS and other causes in sub-Saharan Africa. *Journal of Acquired Immune Deficiency Syndromes*, 39:635–675. https://doi.org/10.1097/01.qai.0000156393.80809.fd.

Henry, L. 1960. Mesure indirecte de la mortalité des adultes. *Population*, 15, 3:457–466. https://doi.org/10.3917/popu.p1960.15n3.0466.

Heuveline, P. 2003. HIV and population dynamics: a general model and maximum-likelihood standards for East Africa. *Demography*, 40(2):217–245. https://doi.org/10.2307/3180799.

Hill, K. and Trussell, J. 1977. Further developments in indirect mortality estimation. *Population Studies*, 31(2):313–334. https://doi.org/10.2307/2173920.

Hill, K., Choi, Y., and Timæus, I. M. 2005. Unconventional approaches to mortality estimation. *Demographic Research*, 13(12):281–300. https://doi.org/10.4054/demres.2005.13.12.

Hosegood, V., Vanneste, A., and Timæus, I. M. 2004. Levels and causes of adult mortality in rural South Africa: the impact of AIDS. *AIDS*, 18(4):663–671. https://doi.org/10.1097/00002030-200403050-00011.

INDEPTH. 2004. *INDEPTH Model Life Tables for Sub-Saharan Africa*. Aldershot: Ashgate.

Joubert, J., Rao, C., Bradshaw, D., Vos, T., and Lopez, A. D. 2013. Evaluating the quality of national mortality statistics from civil registration in South Africa, 1997-2007. *PLoS ONE*, 8 (5):e64592. https://doi.org/10.1371/journal.pone.0064592.

Kasamba, I., Baisley, K., Mayanja, B. N., Maher, D., and Grosskurth, H. 2012. The impact of antiretroviral treatment on mortality trends of HIV-positive adults in rural Uganda: a longitudinal population-based study, 1999-2009. *Tropical Medicine and International Health*, 17(8):e66–73. https://doi.org/10.1111/j.1365-3156.2012.02841.x.

Keyfitz, N. and Caswell, H. 2005. The demographic theory of kinship. In Keyfitz, N. and Caswell, H., editors, *Applied Mathematical Demography (3rd ed.)*, chapter 15, pages 370–398. Springer.

Lesotho NSO. 2009. *2006 Lesotho Population Census Analytical Report, Volume IIIA, Population Dynamics*. Ministry of Finance and Development Planning, Bureau of Statistics.

Lewis, J., Ronsmans, C., Ezeh, A., and Gregson, S. 2004. The population impact of HIV on fertility in sub-Saharan Africa. *AIDS*, 18 (S2):35–43. https://doi.org/10.1097/00002030-200406002-00005.

Lotka, A. 1931. Orphanhood in relation to demographic factors. *Metron*, 9(2):37–109.

Masquelier, B. 2010. *Estimation de la mortalité adulte en Afrique subsaharienne à partir de la survie des proches. Apports de la microsimulation*. Louvain-la-Neuve: Presses Universitaires de Louvain.

Masquelier, B., Reniers, G., and Pison, G. 2014. Divergences in mortality trends in sub-Saharan Africa: survey evidence on the survival of children and siblings. *Population Studies*, 68(2): 161–177. https://doi.org/10.1080/00324728.2013.856458.

Masquelier, B., Eaton, J. W., Gerland, P., Pelletier, F., and Mutai, K. 2017. Age patterns and sex ratios of adult mortality in countries with high HIV prevalence. *AIDS*, 31 Suppl 1:S77–S85. https://doi.org/10.1097/QAD.0000000000001332.

Masquelier, B., Pison, G., Rakotonirina, J., and Rasoanomenjanahary, A. 2019. Estimating cause-specific mortality in Madagascar: an evaluation of death notification data from the capital city. *Population Health Metrics*, 8(17). https://doi.org/10.1186/s12963-019-0190-z.

Menashe Oren, A. and Stecklov, G. 2018. Urban–rural disparities in adult mortality in sub-Saharan Africa. *Demographic Research*, 39(5):136–176. https://doi.org/10.4054/DemRes.2018.39.5.

Nhacolo, A., Nhalungo, D., Sacoor, C., Aponte, J., Thompson, R., and Alonso, P. 2006. Levels and trends of demographic indices in southern rural Mozambique: evidence from demographic surveillance in Manhica district. *BMC Public Health*, 6(1):291. https://doi.org/10.1186/1471-2458-6-291.

Odimegwu, C., Chisumpa, V., and Somefun, O. 2018. Adult mortality in sub-Saharan Africa using 2001—2009 census data: does estimation method matter? *Genus*, 74(1). https://doi.org/10.1186/s41118-017-0025-3.

Palloni, A. and Heligman, L. 1985. Re-estimation of structural parameters to obtain estimates of mortality in developing countries. *Population Bulletin of the United Nations*, 18:10–33.

Palloni, A., Massagli, M., and Marcotte, J. 1984. Estimating adult mortality with maternal orphanhood data: analysis of sensitivity of the techniques. *Population Studies*, 38(2):255–279. https://doi.org/10.2307/2174076.

Preston, S. 1987. Estimation of certain measures in family demography based upon generalized stable population relations. In Bongaarts, J., Burch, T., and Wachter, K., editors, *Family Demography. Methods and Their Application*, pages 40–62. Oxford: Clarendon Press.

Reniers, G., Slaymaker, E., Nakiyingi-Miiro, J., Nyamukapa, C., Crampin, A. C., Herbst, K., Urassa, M., Otieno, F., Gregson, S., Sewe, M., Michael, D., Lutalo, T., Hosegood, V., Kasamba, I., Price, A., Nabukalu, D., Mclean, E., and Zaba, B. 2014. Mortality trends in the era of antiretroviral therapy: evidence from the Network for Analysing Longitudinal Population based HIV/AIDS data on Africa (ALPHA). *AIDS*, 28 Suppl 4:S533–542. https://doi.org/10.1097/qad.0000000000000496.

Robertson, L., Gregson, S., Madanhire, C., Walker, N., Mushati, P., Garnett, G., and Nyamukapa, C. 2008. Discrepancies between UN models and DHS survey estimates of maternal orphan prevalence: insights from analyses of survey data from Zimbabwe. *Sexually Transmitted Infections*, 84 S(1):57–62. https://doi.org/10.1136/sti.2008.029926.

Stover, J. 2009. AIM: A Computer Program for Making HIV/AIDS Projections and Examining the Demographic and Social Impacts of AIDS. Technical report, Washington, DC.

Stover, J., Johnson, P., Zaba, B., Zwahlen, M., Dabis, F., and Ekpini, R. 2008. The Spectrum projection package: improvements in estimating mortality, ART needs, PMTCT impact and uncertainty bounds. *Sexually Transmitted Infections*, 84(Suppl 1):i24–i30. https://doi.org/10.1136/sti.2008.029868.

Stover, J., Johnson, P., Hallett, T., Marston, M., Becquet, R., and Timæus, I. M. 2010. The Spectrum projection package: improvements in estimating incidence by age and sex, mother-to-child transmission, HIV progression in children and double orphans. *Sexually Transmitted Infections*, 86(Suppl 2):16–21. https://doi.org/10.1136/sti.2010.044222.

Stover, J., Brown, T., and Marston, M. 2012. Updates to the Spectrum/Estimation and Projection Package (EPP) model to estimate HIV trends for adults and children. *Sexually Transmitted Infections*, 88 Suppl 2:i11–16. https://doi.org/10.1136/sextrans-2012-050640.

Stover, J., Brown, T., Puckett, R., and Peerapatanapokin, W. 2017. Updates to the spectrum/estimations and projections package model for estimating trends and current values for key HIV indicators. *AIDS*, 31(Supplement 1):S5–S11. https://doi.org/10.1097/qad.0000000000001322.

Timæus, I. M. 1986. An assessment of methods for estimating adult mortality from two sets of data on maternal orphanhood. *Demography*, 23, 3:435–450. https://doi.org/10.2307/2061440.

Timæus, I. M. 1991a. Estimation of mortality from orphanhood in adulthood. *Demography*, 28 (2):213–227. https://doi.org/10.2307/2061276.

Timæus, I. M. 1991b. Estimation of adult mortality from orphanhood before and since marriage. *Population Studies*, 45(3):455–472. https://doi.org/10.1080/0032472031000145636.

Timæus, I. M. 1992. Estimation of adult mortality from paternal orphanhood: a reassessment and a new approach. *Population Bulletin of the United Nations*, 33:47–63.

Timæus, I. M. 2013. Indirect estimation of adult mortality from orphanhood. In Moultrie, T., Dorrington, R., Hill, A., Hill, K., Timæus, I. M., and Zaba, B., editors, *Tools for Demographic Estimation*. Paris: International Union for the Scientific Study of Population.

Timæus, I. M. and Jasseh, M. 2004. Adult mortality in Sub-Saharan Africa: evidence from demographic and health surveys. *Demography*, 41, 4:757–772. https://doi.org/10.1353/dem.2004.0037.

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

Tollman, S., Kahna, K., Garenne, M., and Gear, J. 1999. Reversal in mortality trends: evidence from the Agincourt field site, South Africa, 1992- 1995. *AIDS*, 13(9):1091–1097. https://doi.org/10.1097/00002030-199906180-00013.

UNAIDS. 2022. *UNAIDS DATA 2022*. UNAIDS. https://aidsinfo.unaids.org/.

United Nations. 1983. *Manual X: Indirect Techniques for Demographic Estimation*. New York: Population Division, Dept. of International Economic and Social Affairs, United Nations.

United Nations. 2022. *World Population Prospects: The 2022 Revision*. United Nations. https://esa.un.org/unpd/wpp/.

Verdery, A. M., Smith-Greenaway, E., Margolis, R., and Daw, J. 2020. Tracking the reach of COVID-19 kin loss with a bereavement multiplier applied to the united states. *Proceedings of the National Academy of Sciences*, 117(30):17695–17701. https://doi.org/10.1073/pnas.2007476117.

Wachter, K., Blackwell, D., and Hammel, E. 1997. Testing the validity of kinship microsimulation. *Journal of Mathematical and Computer Modeling*, 26:89–104. https://doi.org/10.1016/s0895-7177(97)00172-6.

Wang, H. et al. 2020. Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950–2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019. *The Lancet*, 396(10258):1160–1203. https://doi.org/10.1016/S0140-6736(20)30977-6.

Zagheni, E. 2011. The impact of the HIV/AIDS epidemic on kinship resources for orphans in Zimbabwe. *Population and Development Review*, 37(4):761–783. https://doi.org/10.1111/j.1728-4457.2011.00456.x.

Zagheni, E. 2015. Microsimulation in Demographic Research. In Wright, J., editor, *International Encyclopedia of the Social and Behavioral Sciences*, volume Vol 15., pages 343–346. Oxford: Elsevier. https://doi.org/10.1016/b978-0-08-097086-8.31018-2.

Zlotnik, H. and Hill, K. 1981. The use of hypothetical cohorts in estimating demographic parameters under conditions of changing fertility and mortality. *Demography*, 18(1):103–122. https://doi.org/10.2307/2061052.

TABLES AND FIGURES

Background mortality		Incidence curve	
α_m	0, -0.4, -0.8	α_{hiv}	5, 7
β_m	0.8, 1.1	β_{hiv}	3, 5
		H	0.08, 0.13
Fertility			
(seronegative women)		HIV settings	
α_f	-0.35, 0.25	Treatments	No, Slow, Rapid scale-up
β_f	0.85, 1.15	Reduced fert./vertical trans.	Yes, No

Table 1 – *Parameters used to set up the microsimulations*

n	β_0	β_1	β_2	RMSE	R^2
0	1.0000	-0.0813	-0.0013	0.0015	0.8401
5	1.0000	-0.2296	0.0032	0.0031	0.8975
10	1.0000	-0.3594	0.0164	0.0042	0.9121
15	1.0000	-0.4672	0.0264	0.0048	0.9144
20	0.9999	-0.5200	0.0154	0.0053	0.9026
25	0.9999	-0.5500	0.0054	0.0056	0.8538
30	1.0000	-0.5550	0.0007	0.0065	0.6352

Table 2 – *Coefficients for adjusting proportions of mothers surviving for HIV-related bias*

	Pre-ART period					Post-ART period				
n	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9
10	-0.2742	0.0010	1.2455	-0.2039	0.2609	-0.3532	0.0021	1.3030	-0.1336	-0.0017
15	-0.1488	0.0016	1.1053	-0.2611	0.3539	-0.2393	0.0026	1.1732	-0.1056	0.0183
20	-0.1123	0.0027	1.0387	-0.2747	0.3711	-0.1693	0.0031	1.0854	-0.1071	-0.0330
25	-0.1336	0.0047	1.0082	-0.2841	0.3386	-0.1559	0.0045	1.0442	-0.1866	-0.1147
30	-0.2051	0.0075	1.0078	-0.2701	0.2995	-0.1915	0.0069	1.0223	-0.2466	-0.0751
35	-0.3074	0.0113	1.0182	-0.2213	0.2649	-0.2708	0.0100	1.0241	-0.2220	-0.0228
40	-0.4580	0.0162	1.0530	-0.1323	0.1770	-0.3801	0.0136	1.0536	-0.1708	-0.0691
45	-0.6244	0.0215	1.1031	-0.0343	0.0671	-0.5119	0.0176	1.1132	-0.1318	-0.2382

Table 3 – *Coefficients for converting proportions of mothers alive into life table survivorship in the presence of HIV and ART*

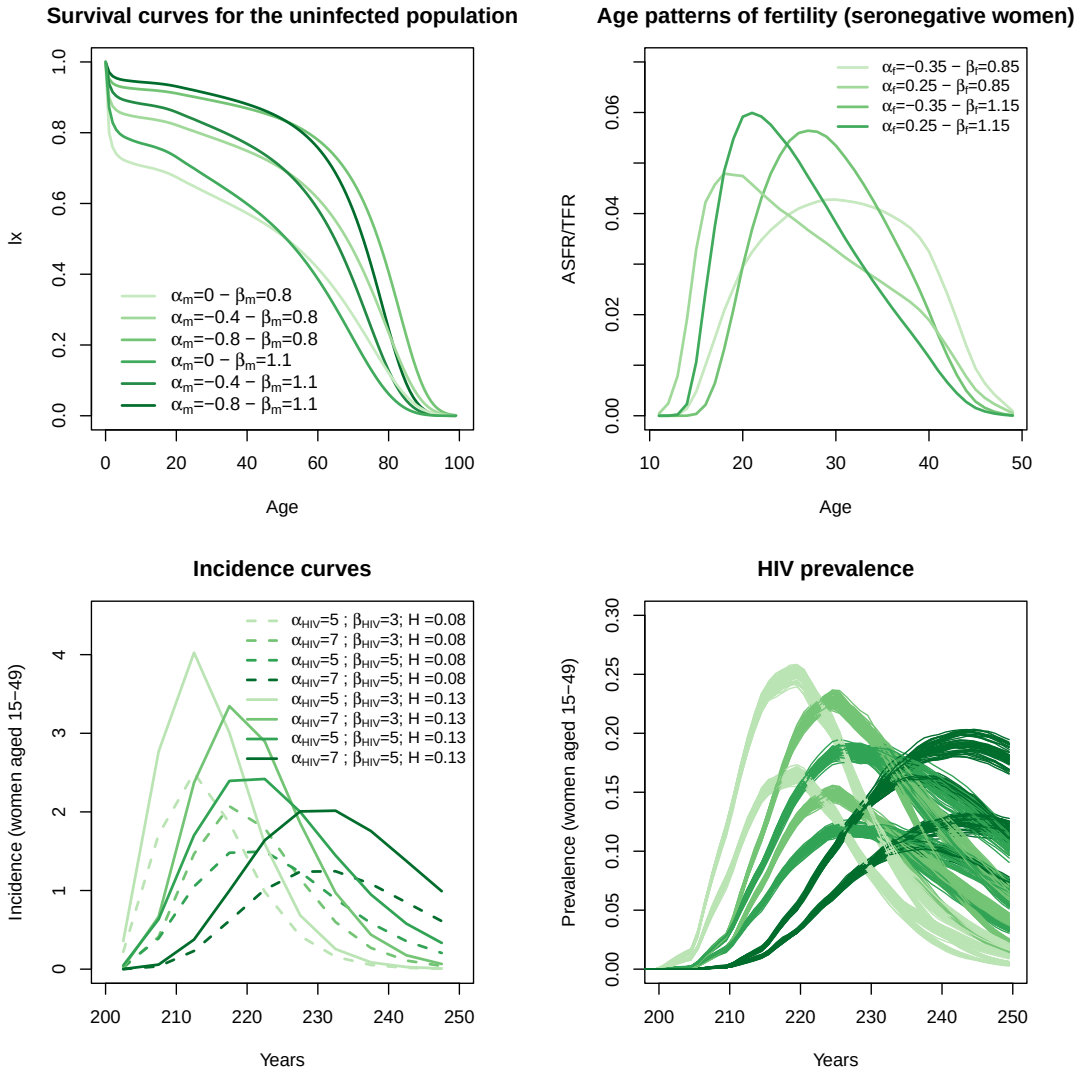

Figure 1 – *(a) Age patterns of non-AIDS mortality, (b) fertility schedules in the absence of HIV, (c) incidence curves and, (d) trends in HIV prevalence*

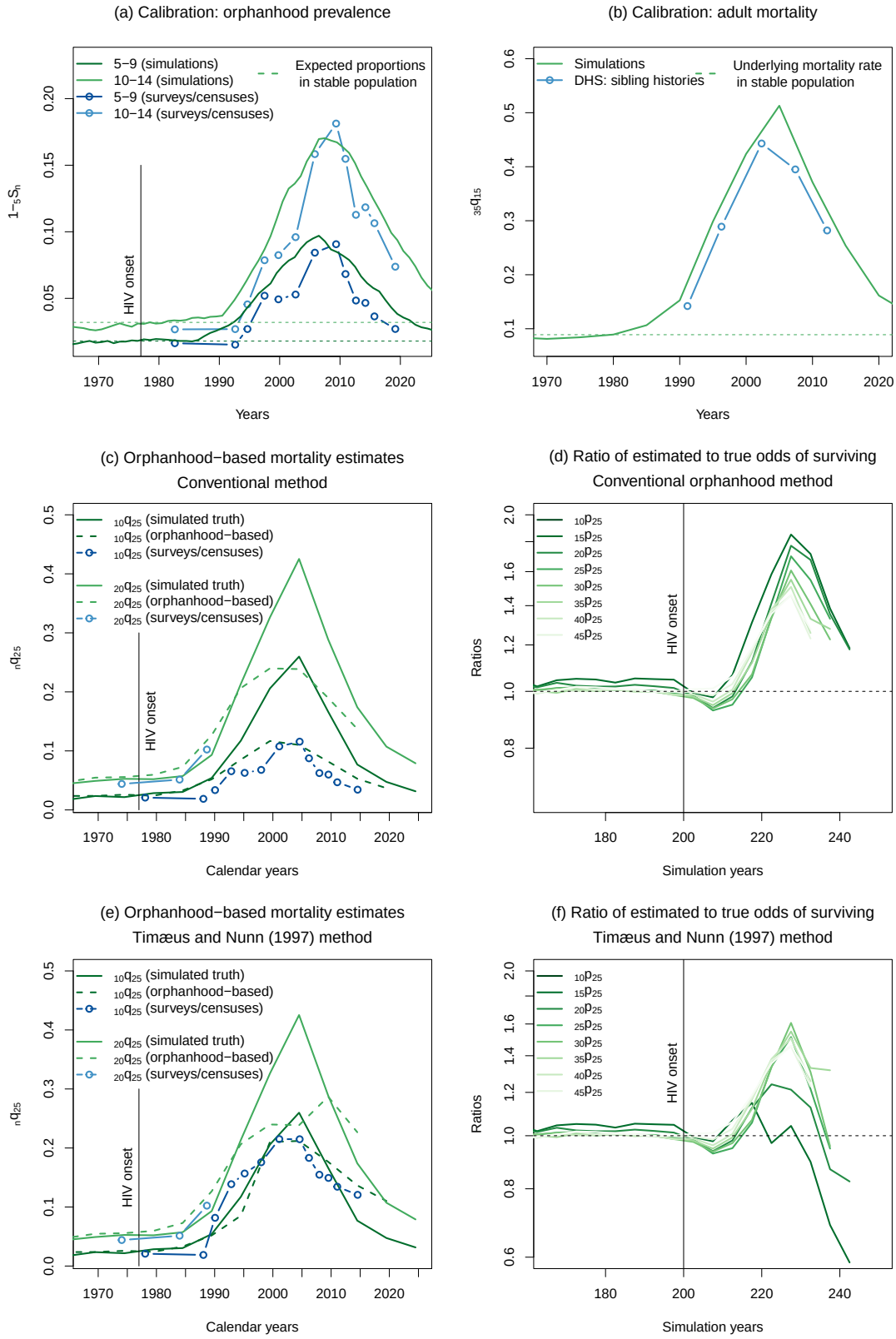

Figure 2 – *(a) Orphanhood prevalence in children aged 5-14 in one simulation resembling Zimbabwe and in surveys and censuses from this country, (b) Adult mortality ($_{35}q_{15}$) in females in the same simulation set and in sibling histories from DHS from Zimbabwe, (c and e) Probabilities $_{10}q_{25}$ and $_{20}q_{25}$ estimated from orphanhood in surveys and censuses from Zimbabwe and in the simulation resembling Zimbabwe, compared to the simulated truth, (d and f) Median ratios of estimated to true odds of surviving across the 576 simulations with reduced fertility and vertical transmission*

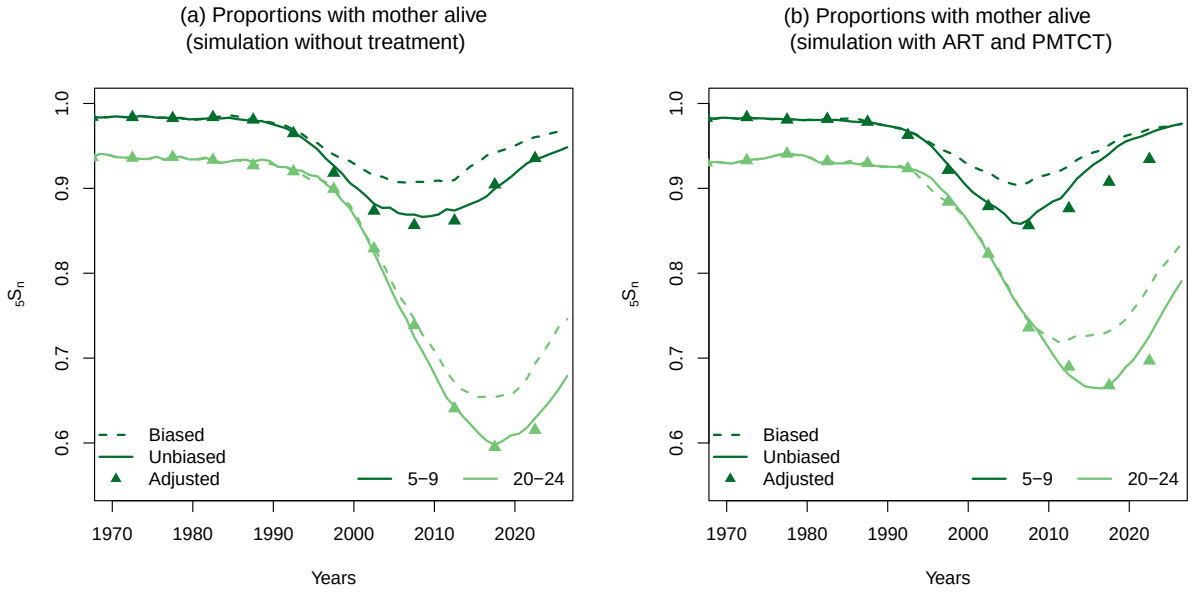

Figure 3 – *Effect of the vertical transmission and reduced fertility in trends in proportions of respondents aged 5-9 and 20-24 with a surviving mother: (a) simulation without treatment, (b) simulation with treatment*

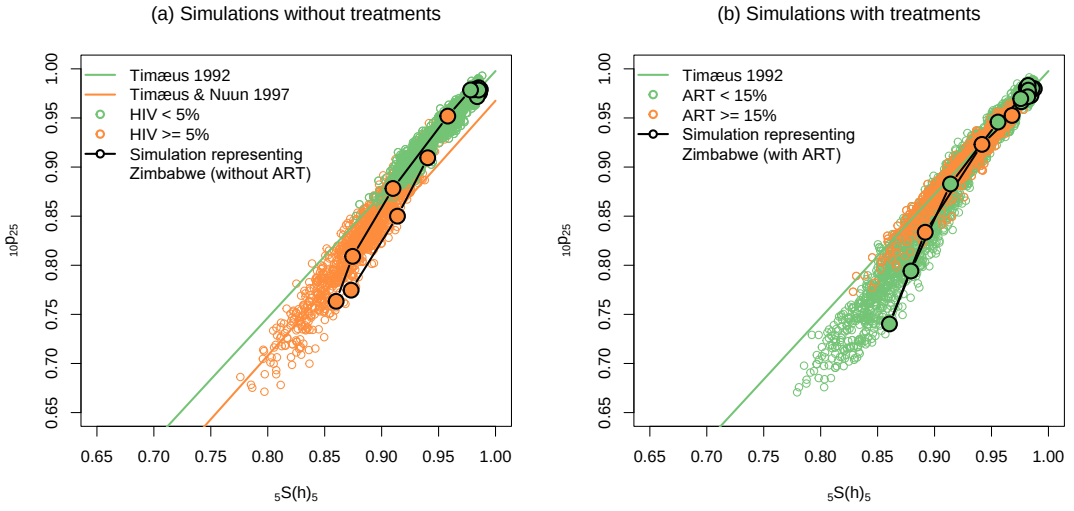

Figure 4 – *Relationship between growth-corrected proportions* $_5S(h)_n$ *and life table survivorship* $_{10}p_{25}$

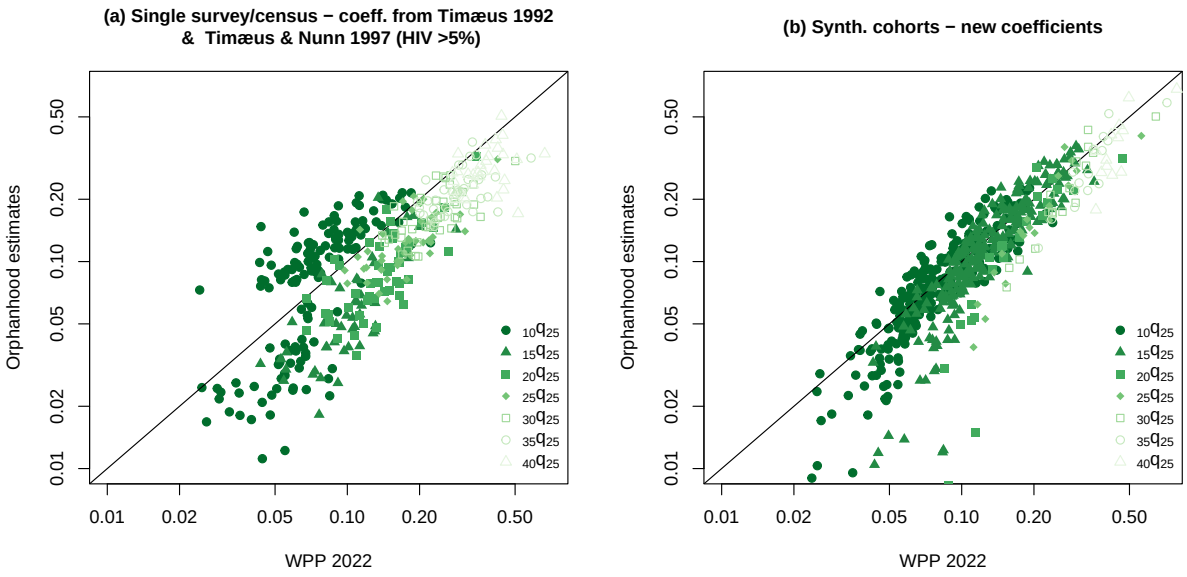

Figure 5 – *Comparison of probabilities ${}_nq_{25}$ obtained from maternal orphanhood and WPP estimates using two estimation approaches*

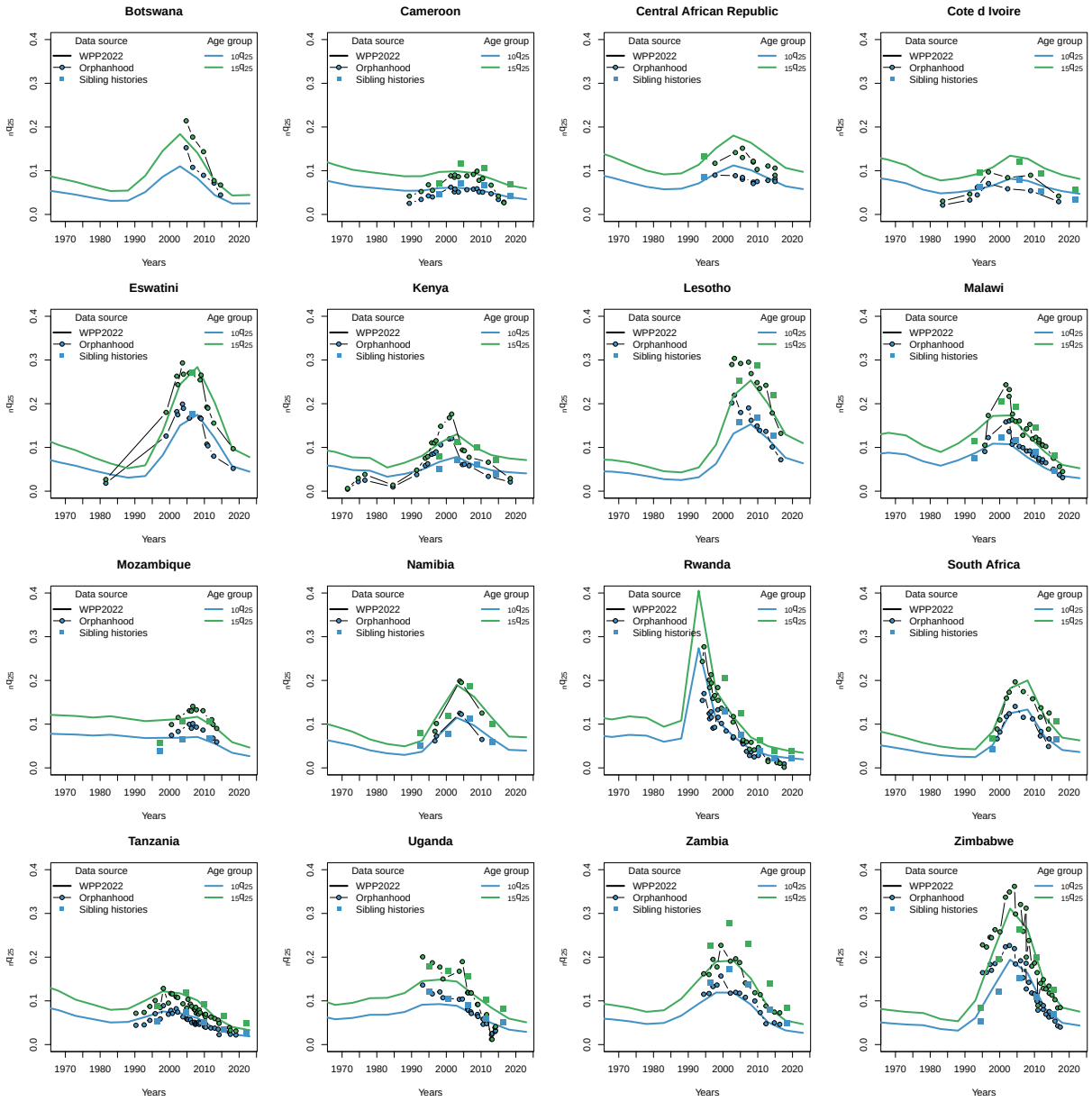

Figure 6 – *Trends in the probabilities* ${}_{10}q_{25}$ *and* ${}_{15}q_{25}$ *using the new coefficients on surveys/censuses reports, estimates from the World Population Prospects 2022 and sibling histories*

Appendix

Appendix A List of censuses conducted in Sub-Saharan Africa with orphanhood data

Western Africa	1970	1980	1990	2000	2010	2020
Benin	1979		1992	**2002**	**2013**	
Burkina Faso	1975	1985	**1996**	**2006**	**2019**	
Cabo Verde	1970	1980	1990	2000	**2010**	2021
Côte d'Iv.	1975	**1988**	**1998**		2014	2021(?)
Gambia	**1973**	**1983**	**1993**	**2003**	**2013**	
Ghana	1970	1984		2000	2010	2021
Guinea		1983	1996		**2014**	
Guinea-Bissau	1970 1979		1991	2009		
Liberia	1974	**1984**		**2008**		2022
Mali	1976	**1987**	**1998**	**2009**		**2022**
Mauritania	1976	1988		2000	2013	
Niger	1977	**1988**		**2001**	**2012**	
Nigeria	1973		1991	**2006**		
Senegal	1976	**1988**		2002	**2013**	**2023**
Sierra Leone	**1974**	**1985**		**2004**	2015	2021
Togo	1970	1981			**2010**	2022(?)

Middle Africa	1970	1980	1990	2000	2010	2020
Angola	1970				**2014**	
Cameroon	1976	**1987**		**2005**		
Central African Republic	1975	**1988**		**2003**		
Chad		**1989**	1993	**2009**		
Congo	**1974**	1984	1996	**2007**		
DR of the Congo		**1984**				
Equat. Guinea		1983	1994	2002	2015	
Gabon	1970	1980	1993	**2003**	2013	
Sao Tome and Principe		1981	1991		2012	

Eastern Africa	1970	1980	1990	2000	2010	2020
Burundi	1979		**1990**	**2008**		
Comoros		**1980**	**1991**	2003	**2017**	
Djibouti		**1983**		**2009**		
Eritrea		1984				
Ethiopia		1984	1994	**2007**		
Kenya	**1979**	**1989**	**1999**	**2009**	2019(?)	
Madagascar	1975		**1993**		2018	
Malawi	**1977**	1987	**1998**	**2008**	2018(?)	
Mauritius	1972	1983	1990	2000	2011	2022

Mozambique	1970	1980	**1997**	**2007**	2017(?)	
Rwanda	1978		**1991**	**2002**	**2012**	2022
Seychelles	1971;1977	1987	1994	2002	2010	2022
Somalia	1975	1987				
South Sudan	**1973**	1983	**1993**	**2008**		
Uganda		**1980**	**1991**	**2002**	**2014**	
United Republic of Tanzania	**1978**	**1988**		**2002**	**2012**	2022(?)
Zambia		1980	1990	2000	**2010**	2022(?)
Zimbabwe		**1982**	**1992**	**2002**	**2012**	2022
Southern Africa	1970	1980	1990	2000	2010	2020
Botswana	**1971**	1981	1991	**2001**	**2011**	**2022**
Eswatini	**1976**	**1986**	**1997**	**2007**	2017(?)	
Lesotho	**1976**	**1986**	**1996**	**2006**	**2016**	
Namibia	1970	1981	1991	**2001**	**2011**	2023(?)
South Africa		1980;1985	1991; **1996**	**2001**	**2011**	2022(?)

Table S1 – *Censuses conducted in Sub-Saharan Africa (censuses in which maternal orphanhood data were collected are in bold)*
Source: United Nations, Department of Economic and Social Affairs, Population Division 2021. Collected Data: List of selected demographic topics collected in specific data sources by country or area, Available from https://population.un.org/DataArchiveWeb/. Censuses for which it is unclear whether orphanhood questions were asked are identified with a question mark.

Appendix B Coefficients used to convert proportions of surviving mothers to survival probabilities

n	β_0	β_1	β_2	R^2	CV
10	-0.2894	0.00125	1.2559	0.997	0.0015
15	-0.1718	0.00222	1.1123	0.996	0.0031
20	-0.1513	0.00372	1.0525	0.995	0.0058
25	-0.1808	0.00586	1.0267	0.993	0.0088
30	-0.2511	0.00885	1.0219	0.992	0.0126
35	-0.3644	0.01287	1.0380	0.992	0.0172
40	-0.5181	0.01795	1.0753	0.992	0.0222
45	-0.6880	0.02342	1.1276	0.993	0.0271
50	-0.8054	0.02721	1.1678	0.992	0.0400

Table S2 – *Coefficients used to convert proportions of surviving mothers to survival probabilities - Sc: Timæus (1992)*

n	β_0	β_1	β_2
10	-0.3611	0.00125	1.2974
15	-0.4030	0.00222	1.3732
20	-0.2120	0.00372	1.1342
25	-0.2389	0.00586	1.1131
30	-0.2513	0.00885	1.0223

Table S3 – *Coefficients provided by Timæus and Nunn (1997) for converting proportions of surviving mothers into survival probabilities in HIV/AIDS-disrupted settings*

Appendix C Parametrization and calibration of the microsimulations

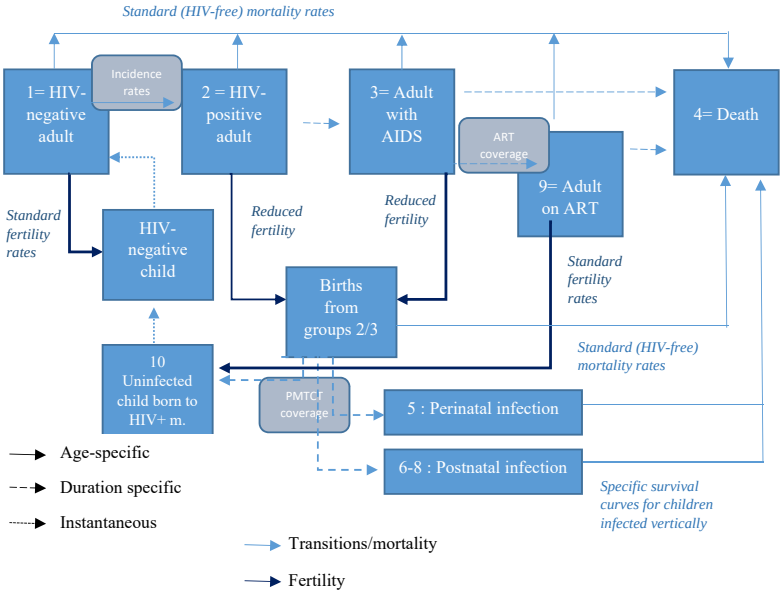

Figure S1 – *Flow chart of population subgroups identified in the microsimulations to incorporate HIV/AIDS*

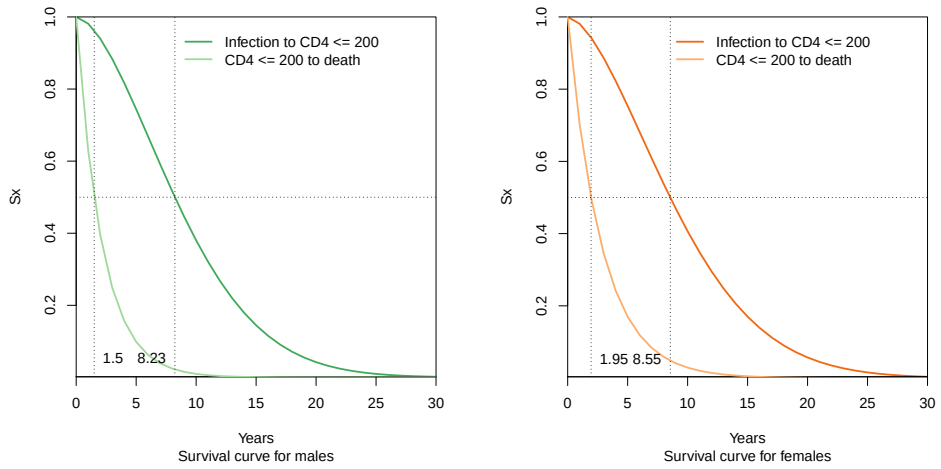

Figure S2 – *Survival curves for infected individuals progressing from infection to clinical stage, then from clinical stage to death, in the absence of treatment - Sc : Stover (2009)*

To build the ART and PMTCT uptake scenarios, we examined treatment coverage in 16 countries where HIV prevalence reached at least 5%. We distinguished between two groups: those where the maximum coverage of ART had reached 83% (9 countries), and those where it remained lower (7 countries). Trends in ART coverage for these countries are presented in Figure S3. A logistic growth curve fitted to these two sets of ART coverage trends helped to build the first two scenarios: rapid and slow increases in ART take-up. For these same countries, another logistic curve was fitted on PMTCT treatment coverage. This helps reflect the faster increase in

PMTCT coverage and the time lag between the two curves. In addition to the scenarios with rapid and slow scale-up of ART, a third scenario without any treatment was included.

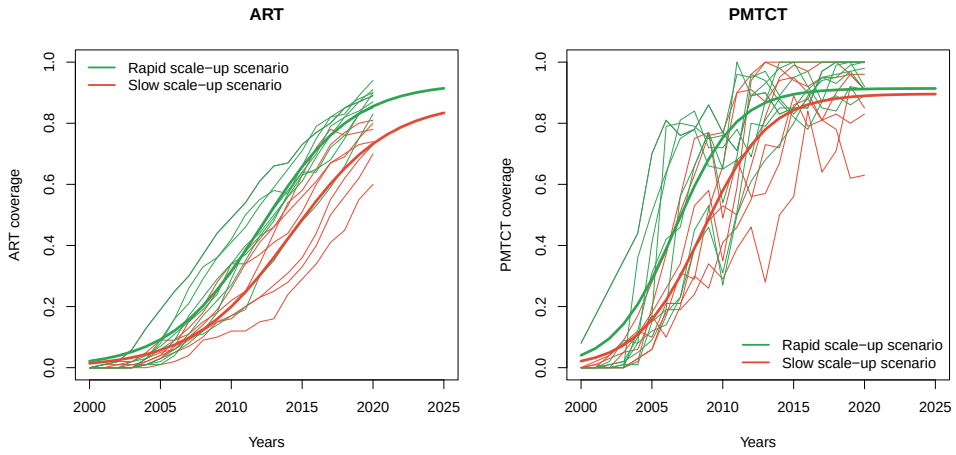

Figure S3 – *ART and PMTCT treatment coverage trends in 16 sub-Saharan African countries where prevalence has reached 5%.*
Sc: UNAIDS estimates extracted from https://data.worldbank.org/

Women were randomly recruited to the subgroup on ART at the beginning of each 5-year period, so as to achieve the desired coverage. To reflect the fact that historically pregnant women identified at antenatal clinics tended to initiate ART before other women, we prioritized women who were soon to give birth. Priority was given to women aged 15-49 who have reached the clinical stage and for whom SOCSIM has scheduled a birth as the next event. Women who have reached the clinical stage without a scheduled birth were second in the order of priority. They were followed by HIV-positive women who have not yet reached the clinical stage but have a birth scheduled in the next few months, and finally, other HIV-positive women aged 15-49, until we reached the expected ART coverage. This procedure ensures that PMTCT coverage increases faster than ART coverage.

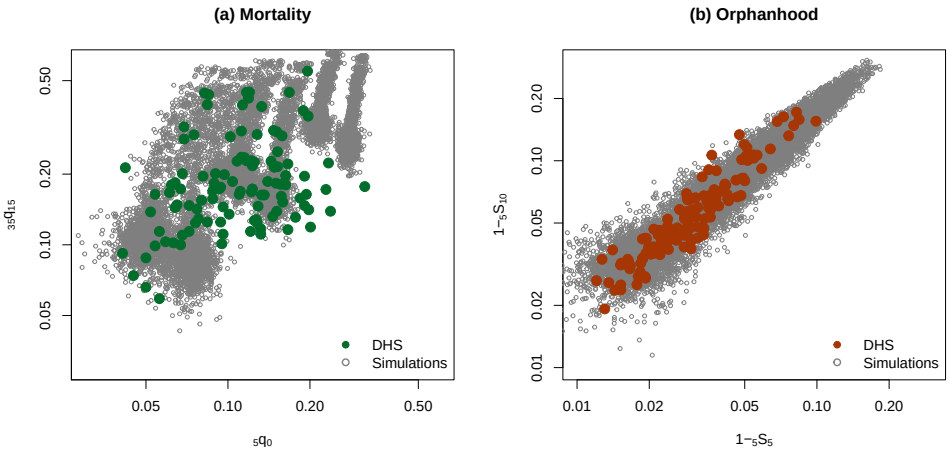

Figure S4 – *(a) Adult ($_{35}q_{15}$) and child ($_5q_0$) mortality in DHS surveys and in the set of simulations, (b) Orphanhood prevalence in DHS surveys and in the simulation set (in children aged 5-9 and 10-14).*

Appendix D Ratios of unbiased to observed proportions of mothers surviving

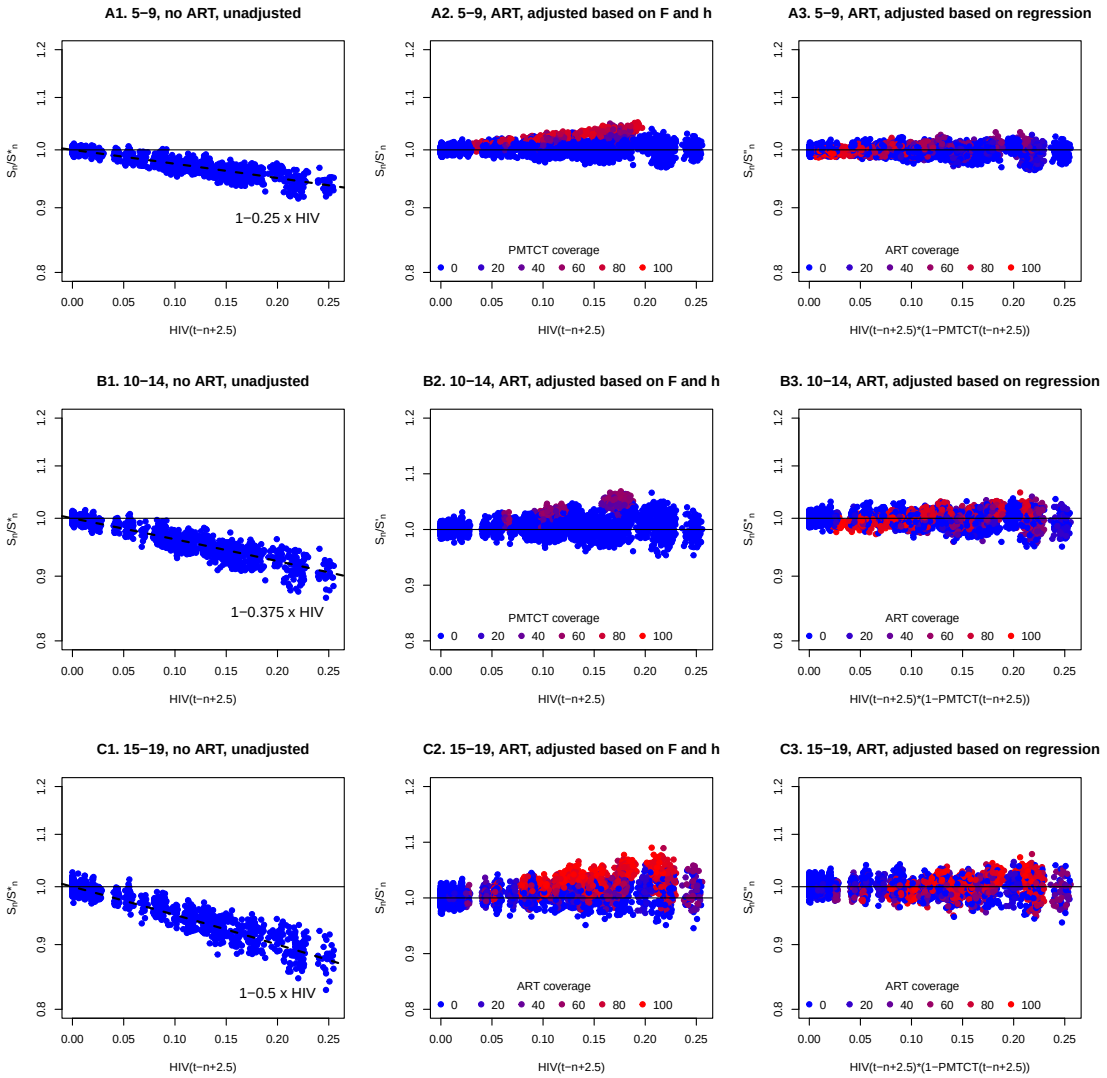

Figure S5 – *Ratios of unbiased to observed proportions of mothers surviving, in simulations without treatments (${}_5S_n/{}_5S_n^*$, left panel), in all simulations with adjustments developed by Timæus and Nunn (1997) (${}_5S_n/{}_5S_n'$, middle panel) and with adjustments obtained through regression (${}_5S_n/{}_5S_n''$, right panel), for three age groups (5-9, 10-14 and 15-19).*

Appendix E Median ratios of estimated to true odds of surviving in simulations

	Pre-HIV		HIV epidemic								
Year	192	197	202	207	212	217	222	227	232	237	242
Coeff.: Timæus (1992) - Time location: Brass and Bamgboye (1981)											
${}_{10}p_{25}$	1.05	1.05	0.99	0.98	1.07	1.31	1.58	1.85	1.72	1.38	1.18
${}_{15}p_{25}$	1.02	1.01	0.98	0.94	0.98	1.12	1.41	1.77	1.68	1.35	1.18
${}_{20}p_{25}$	1.00	0.99	0.98	0.93	0.95	1.06	1.34	1.70	1.55	1.33	
${}_{25}p_{25}$	1.00	0.99	0.97	0.94	0.97	1.07	1.33	1.61	1.41	1.23	
${}_{30}p_{25}$	1.00	0.99	0.98	0.95	0.99	1.12	1.36	1.55	1.33	1.28	
${}_{35}p_{25}$	1.00	0.99	0.99	0.96	1.01	1.17	1.38	1.51	1.26		
${}_{40}p_{25}$	1.00	1.00	1.00	0.98	1.03	1.17	1.37	1.46	1.23		
${}_{45}p_{25}$	1.00	1.00	1.01	1.01	1.06	1.19	1.36	1.46	1.28		
Coeff.: Timæus and Nunn (1997) when HIV >= 5%, Timæus (1992) when HIV < 5% Time location: Brass and Bamgboye (1981)											
${}_{10}p_{25}$	1.05	1.05	0.99	0.98	1.07	1.15	0.97	1.04	0.90	0.69	0.59
${}_{15}p_{25}$	1.02	1.01	0.98	0.94	0.98	1.12	1.24	1.21	1.13	0.87	0.82
${}_{20}p_{25}$	1.00	0.99	0.98	0.93	0.95	1.06	1.34	1.51	1.22	0.95	
${}_{25}p_{25}$	1.00	0.99	0.97	0.94	0.97	1.07	1.33	1.61	1.31	0.96	
${}_{30}p_{25}$	1.00	0.99	0.98	0.95	0.99	1.12	1.36	1.55	1.33	1.32	
${}_{35}p_{25}$	1.00	0.99	0.99	0.96	1.01	1.17	1.38	1.51	1.26		
${}_{40}p_{25}$	1.00	1.00	1.00	0.98	1.03	1.17	1.37	1.46	1.23		
${}_{45}p_{25}$	1.00	1.00	1.01	1.01	1.06	1.19	1.36	1.46	1.28		
New coefficients on adjusted proportions Time location: synthetic cohorts											
${}_{10}p_{25}$	1.01	1.01	0.98	1.01	0.99	0.98	1.01	1.03	1.01	0.95	0.98
${}_{15}p_{25}$	1.00	1.00	0.99	1.03	1.01	0.98	0.99	1.04	1.02	0.95	1.00
${}_{20}p_{25}$	1.00	1.00	1.00	1.03	1.02	0.98	0.98	1.04	1.02	0.99	0.99
${}_{25}p_{25}$	1.00	0.99	1.00	1.02	1.01	0.99	0.98	1.03	1.02	1.00	1.00
${}_{30}p_{25}$	1.00	1.00	1.01	1.02	1.00	0.99	0.98	1.03	1.01	1.00	1.00
${}_{35}p_{25}$	1.00	1.00	1.01	1.02	1.00	0.99	0.98	1.03	1.01	0.99	0.99
${}_{40}p_{25}$	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.02	1.01	0.99	0.99
${}_{45}p_{25}$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.03	1.01	1.00	0.99

Table S4 – *Median ratios of estimated to true odds of surviving in simulations with vertical transmission and reduced fertility*

Appendix F Prediction errors for models to correct bias in proportions

Model performance was evaluated using the root-mean-square error, and the maximum and minimum values of the median ratio of the estimated to the true odds of surviving. These metrics were first computed in a sample of 80% of all simulations, which served to obtain the coefficients. Then out-of-sample metrics were calculated based on values predicted from these coefficients in the remaining simulations.

	0-4 (n =5)		5-9 (n =10)		10-14 (n =15)	
RMSE	in-sample	out-of-sample	in-sample	out-of-sample	in-sample	out-of-sample
Model 1	0.00206	0.00207	0.00639	0.00667	0.01141	0.01038
Model 2	0.00186	0.00186	0.00429	0.00445	0.00555	0.00529
Model 3	0.00158	0.00157	0.00332	0.00346	0.00445	0.00418
Model 4	0.00157	0.00155	0.00333	0.00351	0.00465	0.00441
Model 5	0.00156	0.00155	0.00327	0.00343	0.00431	0.00406
Model 6	0.00148	0.00147	0.00309	0.00333	0.00462	0.00439
Model 7	0.00147	0.00146	0.00306	0.00330	0.00420	0.00399
	15-19 (n =20)		20-24 (n =25)		25-29 (n =30)	
RMSE	in-sample	out-of-sample	in-sample	out-of-sample	in-sample	out-of-sample
Model 1	0.01508	0.01419	0.01653	0.01460	0.01446	0.01471
Model 2	0.00548	0.00562	0.00552	0.00531	0.00574	0.00534
Model 3	0.00481	0.00496	0.00530	0.00515	0.00571	0.00535
Model 4	0.00548	0.00562	0.00552	0.00531	0.00574	0.00534
Model 5	0.00481	0.00496	0.00530	0.00515	0.00571	0.00535
Model 6	0.00548	0.00562	0.00552	0.00531	0.00574	0.00534
Model 7	0.00481	0.00496	0.00530	0.00515	0.00571	0.00535
	30-34 (n =35)		35-39 (n =40)		40-44 (n =45)	
RMSE	in-sample	out-of-sample	in-sample	out-of-sample	in-sample	out-of-sample
Model 1	0.01088	0.00996	0.00854	0.00917	0.01023	0.01005
Model 2	0.00652	0.00634	0.00782	0.00850	0.01023	0.01004
Model 3	0.00652	0.00634	0.00782	0.00849	0.01023	0.01004
Model 4	0.00652	0.00634	0.00782	0.00850	0.01023	0.01004
Model 5	0.00652	0.00634	0.00782	0.00849	0.01023	0.01004
Model 6	0.00652	0.00634	0.00782	0.00850	0.01023	0.01004
Model 7	0.00652	0.00634	0.00782	0.00849	0.01023	0.01004

Table S5 – *Prediction errors for candidate models for bias in proportions of mothers surviving*

Model 1: $\frac{{}_5S_n}{{}_5S_n^*} \sim \beta_0 + \beta_1\ HIV_t$

Model 2: $\frac{{}_5S_n}{{}_5S_n^*} \sim \beta_0 + \beta_1\ HIV_{t-n+2.5}$

Model 3: $\frac{{}_5S_n}{{}_5S_n^*} \sim \beta_0 + \beta_1\ HIV_{t-n+2.5} + \beta_2\ ART_{t-n+2.5}$

Model 4: $\frac{{}_5S_n}{{}_5S_n^*} \sim \beta_0 + \beta_1\ HIV_{t-n+2.5} + \beta_2\ PMTCT_{t-n+2.5}$

Model 5: $\frac{{}_5S_n}{{}_5S_n^*} \sim \beta_0 + \beta_1\ HIV_{t-n+2.5} + \beta_2\ PMTCT_{t-n+2.5} + \beta_3\ ART_{t-n+2.5}$

Model 6: $\frac{{}_5S_n}{{}_5S_n^*} \sim \beta_0 + \beta_1\ [\ HIV_{t-n+2.5} \times (1 - PMTCT_{t-n+2.5})]$

Model 7: $\frac{{}_5S_n}{{}_5S_n^*} \sim \beta_0 + \beta_1\ [\ HIV_{t-n+2.5} \times (1 - PMTCT_{t-n+2.5})\ ] + \beta_2\ ART_t$

Appendix G Prediction errors for models to convert proportions into life table survivorship probabilities

Model performance was again evaluated using the root-mean-square error, and the maximum and minimum values of the median ratio of the estimated to the true odds of surviving.

5-9 (n =10)	In-sample			Out-of-sample		
	RMSE	Max ratio	Min ratio	RMSE	Max ratio	Min ratio
Model 1	0.0110	1.1265	0.9063	0.0111	1.1236	0.9030
Model 2	0.0100	1.0781	0.8598	0.0101	1.0858	0.8635
Model 3	0.0090	1.0421	0.9226	0.0089	1.0449	0.9354
Model 4	0.0087	1.0555	0.9576	0.0087	1.0514	0.9570
Model 5	0.0100	1.0786	0.8609	0.0100	1.0884	0.8629
Model 6	0.0097	1.0825	0.8983	0.0097	1.0894	0.8979
Model 7	0.0093	1.0598	0.8724	0.0092	1.0679	0.8758
Model 8	0.0099	1.0682	0.8530	0.0099	1.0773	0.8536
Model 9	0.0086	1.0346	0.9475	0.0086	1.0352	0.9505
10-14 (n =15)	In-sample			Out-of-sample		
	RMSE	Max ratio	Min ratio	RMSE	Max ratio	Min ratio
Model 1	0.0134	1.1330	0.9272	0.0137	1.1125	0.9194
Model 2	0.0121	1.0852	0.8862	0.0125	1.0789	0.8809
Model 3	0.0108	1.0630	0.9172	0.0110	1.0548	0.9113
Model 4	0.0103	1.0621	0.9718	0.0105	1.0574	0.9692
Model 5	0.0120	1.0819	0.8828	0.0124	1.0725	0.8763
Model 6	0.0116	1.0804	0.9199	0.0121	1.0701	0.9116
Model 7	0.0108	1.0610	0.9002	0.0110	1.0525	0.8867
Model 8	0.0119	1.0763	0.8786	0.0123	1.0653	0.8732
Model 9	0.0100	1.0414	0.9547	0.0103	1.0365	0.9457
15-19 (n =20)	In-sample			Out-of-sample		
	RMSE	Max ratio	Min ratio	RMSE	Max ratio	Min ratio
Model 1	0.0150	1.1276	0.9499	0.0148	1.1300	0.9443
Model 2	0.0134	1.0843	0.9087	0.0136	1.0878	0.8931
Model 3	0.0128	1.0875	0.9160	0.0131	1.0848	0.9027
Model 4	0.0117	1.0584	0.9792	0.0120	1.0591	0.9583
Model 5	0.0133	1.0810	0.9061	0.0135	1.0844	0.8895
Model 6	0.0130	1.0770	0.9321	0.0133	1.0800	0.9284
Model 7	0.0120	1.0610	0.9390	0.0123	1.0612	0.9544
Model 8	0.0132	1.0742	0.9033	0.0134	1.0783	0.8875
Model 9	0.0113	1.0410	0.9806	0.0115	1.0402	0.9582
20-24 (n =25)	In-sample			Out-of-sample		
	RMSE	Max ratio	Min ratio	RMSE	Max ratio	Min ratio
Model 1	0.0168	1.1311	0.9670	0.0168	1.1446	0.9567
Model 2	0.0147	1.0752	0.9289	0.0148	1.0762	0.9143

Model 3	0.0144	1.0792	0.9309	0.0144	1.0800	0.9163
Model 4	0.0135	1.0520	0.9698	0.0135	1.0460	0.9535
Model 5	0.0147	1.0776	0.9288	0.0148	1.0771	0.9155
Model 6	0.0144	1.0730	0.9484	0.0146	1.0695	0.9455
Model 7	0.0139	1.0597	0.9562	0.0140	1.0583	0.9479
Model 8	0.0145	1.0686	0.9217	0.0146	1.0676	0.9081
Model 9	0.0131	1.0363	0.9706	0.0131	1.0272	0.9541
25-29 (n =30)		In-sample			Out-of-sample	
	RMSE	Max ratio	Min ratio	RMSE	Max ratio	Min ratio
Model 1	0.0190	1.1283	0.9713	0.0189	1.1299	0.9638
Model 2	0.0167	1.0667	0.9400	0.0168	1.0697	0.9376
Model 3	0.0165	1.0668	0.9401	0.0166	1.0688	0.9398
Model 4	0.0158	1.0437	0.9780	0.0159	1.0374	0.9789
Model 5	0.0167	1.0678	0.9411	0.0168	1.0697	0.9409
Model 6	0.0165	1.0615	0.9533	0.0167	1.0597	0.9538
Model 7	0.0161	1.0500	0.9606	0.0163	1.0474	0.9497
Model 8	0.0166	1.0603	0.9366	0.0167	1.0621	0.9377
Model 9	0.0156	1.0354	0.9790	0.0157	1.0317	0.9797
30-34 (n =35)		In-sample			Out-of-sample	
	RMSE	Max ratio	Min ratio	RMSE	Max ratio	Min ratio
Model 1	0.0209	1.1168	0.9748	0.0204	1.1032	0.9644
Model 2	0.0192	1.0527	0.9483	0.0188	1.0484	0.9374
Model 3	0.0192	1.0519	0.9480	0.0188	1.0477	0.9360
Model 4	0.0186	1.0348	0.9891	0.0182	1.0227	0.9736
Model 5	0.0192	1.0529	0.9488	0.0189	1.0479	0.9360
Model 6	0.0192	1.0486	0.9557	0.0188	1.0484	0.9438
Model 7	0.0187	1.0411	0.9672	0.0184	1.0288	0.9651
Model 8	0.0192	1.0502	0.9475	0.0188	1.0467	0.9355
Model 9	0.0185	1.0309	0.9855	0.0181	1.0139	0.9752
35-39 (n =40)		In-sample			Out-of-sample	
	RMSE	Max ratio	Min ratio	RMSE	Max ratio	Min ratio
Model 1	0.0220	1.0734	0.9717	0.0229	1.0821	0.9701
Model 2	0.0214	1.0348	0.9654	0.0221	1.0302	0.9668
Model 3	0.0214	1.0346	0.9654	0.0221	1.0300	0.9667
Model 4	0.0212	1.0245	0.9873	0.0218	1.0310	0.9775
Model 5	0.0214	1.0369	0.9672	0.0221	1.0313	0.9688
Model 6	0.0214	1.0367	0.9685	0.0221	1.0282	0.9727
Model 7	0.0213	1.0323	0.9634	0.0219	1.0353	0.9682
Model 8	0.0214	1.0331	0.9654	0.0220	1.0312	0.9646
Model 9	0.0211	1.0217	0.9880	0.0217	1.0296	0.9786
40-44 (n =45)		In-sample			Out-of-sample	
	RMSE	Max ratio	Min ratio	RMSE	Max ratio	Min ratio
Model 1	0.0240	1.0416	0.9574	0.0240	1.0363	0.9305

Model 2	0.0238	1.0314	0.9860	0.0237	1.0259	0.9709
Model 3	0.0238	1.0316	0.9860	0.0237	1.0258	0.9708
Model 4	0.0238	1.0320	0.9858	0.0237	1.0298	0.9709
Model 5	0.0240	1.0344	0.9571	0.0240	1.0294	0.9295
Model 6	0.0240	1.0361	0.9581	0.0239	1.0263	0.9299
Model 7	0.0240	1.0349	0.9562	0.0240	1.0298	0.9292
Model 8	0.0238	1.0288	0.9848	0.0237	1.0247	0.9693
Model 9	0.0237	1.0317	0.9902	0.0236	1.0267	0.9692

Table S6 – *Prediction errors for candidate models for converting proportions of mothers surviving into survival probabilities*

Model 1: ${}_np_{25} = \beta_0(n) + \beta_1(n)\ M + \beta_2(n)\ {}_5S(h)_{n-5}$

Model 2: ${}_np_{25} = \beta_0(n) + \beta_1(n)\ M + \beta_2(n)\ {}_5S(h)_{n-5} + \beta_3\ HIV_t + \beta_4\ ART_t$

Model 3: ${}_np_{25} = \beta_0(n) + \beta_1(n)\ M + \beta_2(n)\ {}_5S(h)_{n-5} + \beta_3\ HIV_t + \beta_4\ ART_t + \beta_5\ HIV_{t-n+2.5} + \beta_6\ ART_{t-n+2.5}$

Model 4: ${}_np_{25} = \beta_0(n) + \beta_1(n)\ M + \beta_2(n)\ {}_5S(h)_{n-5} + \beta_3\ HIV_t + \beta_4\ ART_t + \beta_5\ \Delta HIV + \beta_6\ \Delta ART$

Model 5: ${}_np_{25} = \beta_0(n) + \beta_1(n)\ M + \beta_2(n)\ {}_5S(h)_{n-5} + \beta_3\ [HIV_t \times (1 - ART_t)]$

Model 6: ${}_np_{25} = \beta_0(n) + \beta_1(n)\ M + \beta_2(n)\ {}_5S(h)_{n-5} \times [HIV_t \times (1 - ART_t)]$

Model 7: ${}_np_{25} = \beta_0(n) + \beta_1(n)\ M + \beta_2(n)\ {}_5S(h)_{n-5} + \beta_3\ [HIV_t \times (1 - ART_t)] + \beta_4\ \Delta[HIV \times (1 - ART)]$

Model 8 (two equations):

$\{{}_np_{25} = \beta_0(n) + \beta_1(n)\ M + \beta_2(n)\ {}_5S(h)_{n-5} + \beta_3\ HIV_t\}\ (ART_t = 0)$

$\{{}_np_{25} = \beta_4(n) + \beta_5(n)\ M + \beta_6(n)\ {}_5S(h)_{n-5} + \beta_7\ [HIV_t \times (1 - ART_t)]\}\ (ART_t > 0)$

Model 9 (two equations):

$\{{}_np_{25} = \beta_0(n) + \beta_1(n)\ M + \beta_2(n)\ {}_5S(h)_{n-5} + \beta_3\ HIV_t + \beta_4\ \Delta HIV\}\ (ART_t = 0)$

$\{{}_np_{25} = \beta_5(n) + \beta_6(n)\ M + \beta_7(n)\ {}_5S(h)_{n-5} + \beta_8\ [HIV_t \times (1 - ART_t)] + \beta_9\ \Delta[HIV \times (1 - ART)]\}\ (ART_t > 0)$

where

HIV_t *or* ART_t *= HIV prevalence or ART coverage at the time of data collection (obtained as the average of estimates from each survey as we use two series to construct the synthetic proportion)*

$HIV_{t-n+2.5}$ *or* $ART_{t-n+2.5}$ *= HIV prevalence or ART coverage at the time of birth (average of estimates from each survey)*

Δ *HIV or* Δ *ART = difference between the two surveys in HIV prevalence or ART coverage at the time of data collection*

Appendix H Data sources on maternal orphanhood prevalence

Proportions of orphans by age were calculated based on microdata in the IPUMS database, DHS and MICS, based on census reports available online, and on the DemoData database of the United Nations Population Division:

- Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 7.3 [dataset]. Minneapolis, MN: IPUMS, 2020. https://doi.org/10.18128/D020.V7.3
- ICF. 2004-2017. Demographic and Health Surveys (various) [Datasets]. Funded by USAID. Rockville, Maryland: ICF [Distributor].
- Anna Bolgrien, Elizabeth Heger Boyle, Matthew Sobek, and Miriam King. IPUMS MICS Data Harmonization Code. Version 1.1 [Stata syntax]. IPUMS: Minneapolis, MN. , 2024. https://doi.org/10.18128/D082.V1.1
- United Nations DemoData: https://popdiv.dfs.un.org/DemoData/web/ - see Gerland, P. (2023, December). What's Beneath the Future: World Population Prospects. In Semaine Data-SHS, Dec 2023, Aix-en-Provence, France.

When the data could be disaggregated by gender of the respondent, we retained the proportions of surviving mothers computed from female respondents, as they were on average lower than in reports from males, possibly due to some age exaggeration in men (Ewbank 1981).

The mean age at childbearing was calculated based on the World Population Prospects to have time-varying estimates (United Nations 2022).

Botswana	1971 Census, 2001 Census, 2007 Family Health Survey (MICS), 2011 Census, 2017 Survey
Cameroon	1960-1965 Enquête Démographique, 1978 World Fertility Survey, 1987 Census, 1991 DHS, 1998 DHS, 2000 MICS, 2004 DHS, 2005 Census, 2006 MICS, 2011 DHS, 2014 MICS, 2018 DHS
Central African Rep.	1988 Census, 1994-1995 DHS, 2000 MICS, 2006 MICS, 2010 MICS, 2018-2019 MICS
Côte d Ivoire	1978-1979 Enquête démographique à passages répétés, 1988 Census, 1994 DHS, 1998 Census, 2000 MICS, 2005 AIS, 2006 MICS, 2011-2012 DHS, 2016 MICS, 2021 DHS
Eswatini	1976 Census, 1986 Census, 1997 Census, 2000 MICS, 2006-2007 DHS, 2007 Census, 2010 MICS, 2014 MICS, 2021-2022 MICS
Kenya	1969 Census, 1973 Demographic Baseline Survey, 1977 National Demographic Survey, 1979 Census, 1983 National Demographic Survey, 1989 Census, 1993 DHS, 1998 DHS, 1999 Census, 2000 MICS, 2003 DHS, 2009 Census, 2014 DHS, 2022 DHS
Lesotho	1971-1973 Demographic Survey, 1976 Census, 1977 WFS, 1986 Census, 2000 MICS, 2001 Demographic Survey, 2004 DHS, 2006 Census, 2009 DHS, 2014 DHS, 2016 Census, 2018 MICS
Malawi	1966 Census, 1970-1972 Population Change Survey, 1977 Census, 1982 Demographic Survey, 1992 DHS, 1998 Census, 2000 DHS, 2004 DHS, 2006 MICS, 2008 Census, 2010 DHS, 2013-2014 MICS, 2015-2016 DHS, 2019-2020 MICS
Mozambique	1997 DHS, 1997 Census, 2003 DHS, 2007 Census, 2008 MICS, 2009 HIV-AIDS Indicator Survey, 2011 DHS, 2015 HIV-AIDS Indicator Survey
Namibia	1992 DHS, 2000 DHS, 2001 Census, 2006-2007 DHS, 2013 DHS
Rwanda	1991 Census, 1992 DHS, 1996 Socio-demographic Survey, 2000 DHS, 2000 MICS, 2002 Census, 2005 DHS, 2010 DHS, 2012 Census, 2014-2015 DHS, 2019 DHS
South Africa	1996 Census, 1998 DHS, 2001 Census, 2007 Community Survey, 2011 Census, 2016 Community Survey, 2016 DHS
Tanzania	1973 National Demographic Survey, 1978 Census, 1988 Census, 1991-1992 DHS, 1996 DHS, 1999 Reproductive and Child Health Survey, 2002 Census, 2003-2004 AIS, 2004-2005 DHS, 2007-2008 AIS/MIS, 2010 DHS, 2011 HIV-AIDS Indicator Survey, 2012 Census, 2015-2016 DHS, 2022 DHS
Uganda	1969 Census, 1988-1989 DHS, 1991 Census, 1995 DHS, 2000-2001 DHS, 2002 Census, 2006 DHS, 2011 DHS, 2014 Census, 2016 DHS
Zambia	1992 DHS, 1996 DHS, 1999 MICS, 2001-2002 DHS, 2007 DHS, 2010 Census, 2013-2014 DHS, 2018 DHS
Zimbabwe	1982 Census, 1992 Census, 1994 DHS, 1997 Inter-Censal Demographic Survey, 1999 DHS, 2002 Census, 2005-2006 DHS, 2009 MICS, 2010-2011 DHS, 2012 Census, 2014 MICS, 2015 DHS, 2019 MICS

Table S7 – *Data sources on maternal orphanhood prevalence used in this study*

Appendix I Trends in the probabilities ${}_{10}q_{25}$ based on orphanhood, using coefficients from Timæus and Nunn (1997)

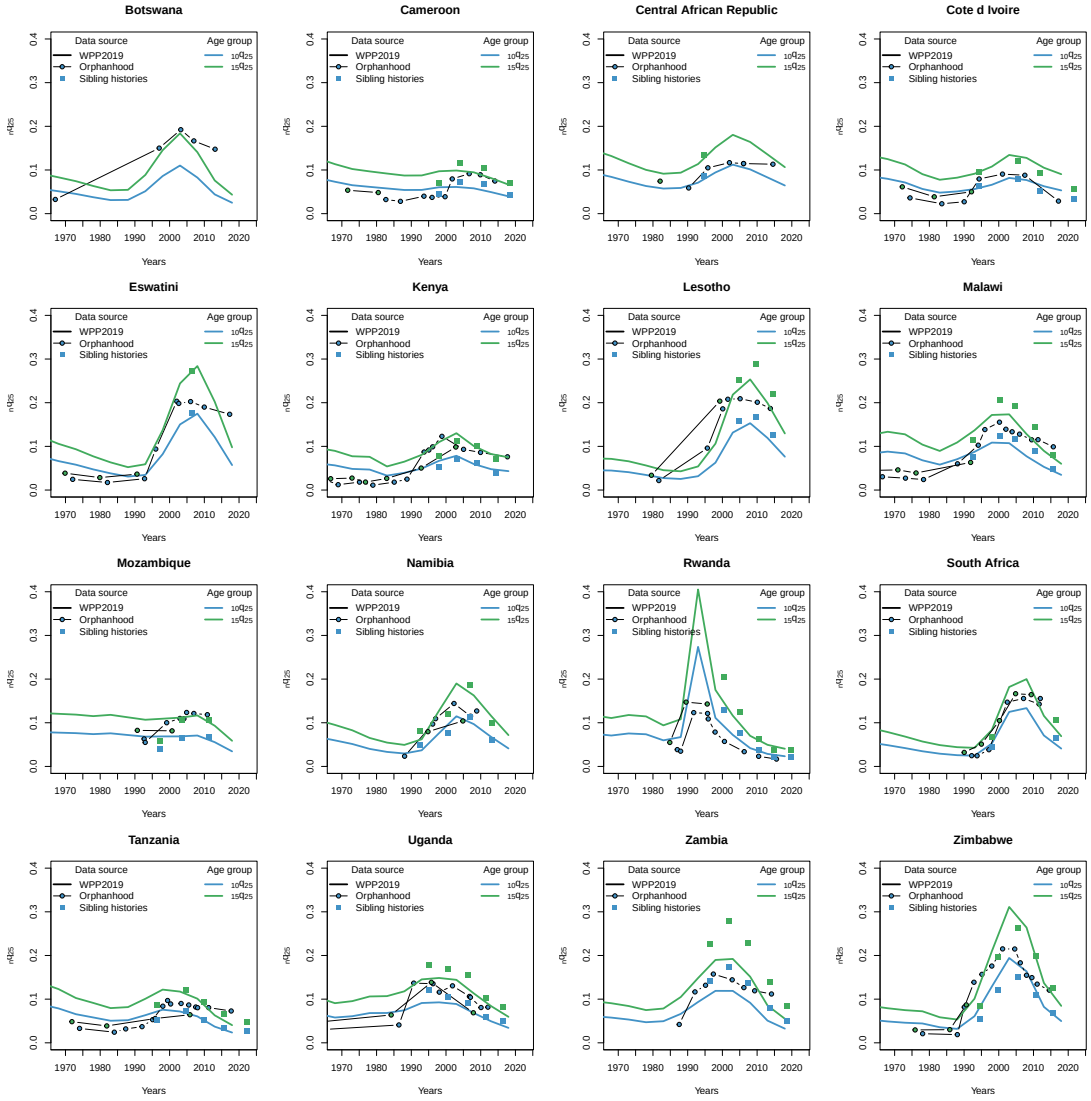

Figure S6 – *Trends in the probabilities ${}_{10}q_{25}$ based on orphanhood, using coefficients from Timæus and Nunn (1997) when HIV >= 5% and Timæus (1992) when HIV < 5%, from a single survey or census, estimates from WPP and sibling histories*