Estimation and Projection of Dual Orphans in Populations with Generalized HIV Epidemics: Updated Methods

A Report to UNAIDS

Ian M. Timæus
Centre for Population Studies
London School of Hygiene & Tropical Medicine

Centre for Population Studies
London School of Hygiene
& Tropical Medicine
Keppel Street
London, WC1E 7HT
United Kingdom

Tel: +44-(0)20-7299 4689
email: ian.timaeus@lshtm.ac.uk

31st December, 2007
(Last revised: 24th April, 2008)

Estimation and Projection of Dual Orphans in Populations with Generalized HIV Epidemics: Updated Methods

Background

The death from AIDS of large numbers of young and middle-aged adults in Africa and elsewhere has produced a parallel rise in the number of orphaned children. Estimates and projections of the number of children with parents who have died from AIDS and other causes are needed to inform policy and programmatic decisions. Estimation of the number of children whose father has died (paternal orphans) or whose parents have both died (dual or double orphans) is relatively complex. Therefore, some of the early studies of the impact of AIDS in Africa on the prevalence of orphanhood only presented statistics on children whose mother had died (maternal orphans) (UNAIDS 2000). Other studies used simple assumptions to estimate paternal orphanhood that produced very approximate estimates (Hunter and Williamson 2000).

At the Rome meeting of the UNAIDS Reference Group on Estimates, Modelling and Projections in 2000, Timæus proposed that it should be possible to extend methods being used already to estimate numbers of maternal orphans in Africa (Gregson *et al.* 1994) to the estimation of paternal orphanhood (Grassly *et al.* 2000). A suitable model of men's fertility existed (Paget and Timæus 1994) and estimates made with it would be adequate even in populations with rather different age patterns of fertility among men. Estimation of dual orphans using this approach would be more difficult as it would require a very complex compartmental model that tracked whether each parent was infected with HIV, durations since infection, and parental ages according to their children's age. Instead, Timæus proposed that numbers of maternal and paternal orphans should be calculated separately from estimates of AIDS and non-AIDS mortality and fertility. Then the overlap between these two groups of orphans (i.e. the dual or double orphans) could be predicted using a regression model fitted to Demographic and Health Survey (DHS) data on the size of the three groups of orphans in different African populations. Suitable predictors would include the age of the child, HIV prevalence and indirect but widely available indices of the ages of adults bearing children such as the proportion of women age 15-19 who remain unmarried and the proportion of married women in polygynous unions.

These approaches to the estimation of paternal and dual orphanhood were subsequently developed collaboratively by Grassly (an epidemiologist) and Timæus (a demographer). Following their review at a meeting of members of the Reference Group on Estimates, Modelling and Projections in Talloires, France in July 2001, they were used to produce the estimates of AIDS and other orphans in countries with generalized epidemics published by UNAIDS, UNICEF and USAID for the *Report on the Global HIV/AIDS Epidemic* and the *Children on the Brink* volumes in 2002 (UNAIDS 2002; UNAIDS *et al.* 2002). The same methods have also been used to estimate numbers of orphans for the 2004 volumes with the same titles and for subsequent editions of the *AIDS Epidemic Update* and *Report on the Global AIDS Epidemic* volumes. A full description of these methods for estimating numbers of orphans and an illustrative application of them to data from Tanzania has been published in the scientific literature (Grassly and Timæus 2005).

When this earlier work on the estimation of dual orphanhood was undertaken, data were available from 36 DHS in Africa that had collected information on orphanhood of children.

However, the quality of the data in two of these surveys was too poor for them to be useful. The data from a further 24 such DHS are now available. Moreover, only eight of the surveys used in the earlier analysis were conducted in populations with a high prevalence of HIV infection (i.e. >6 per cent five years before the conduct of the survey). The results of a further ten such surveys have now been released, greatly increasing the evidence on the relationship between dual orphanhood and maternal and paternal orphanhood in populations with high AIDS mortality.

Terms of Reference

A meeting of the UNAIDS Reference Group on Estimates, Modelling and projections was held in Baltimore on July 12th, 2007 to reconsider the "Estimation of orphanhood due to AIDS and non-AIDS causes and the impact of intervention programmes". One recommendation that emerged from this meeting was that the regression analysis of dual orphanhood should be updated. The London School of Hygiene & Tropical Medicine were contracted by UNAIDS to conduct this study and the following terms of reference agreed:

i) Using Demographic and Health Survey (DHS) data, update an earlier analysis investigating the relationship between maternal and paternal orphanhood and the dual orphanhood of children in mainland sub-Saharan Africa.

ii) Using information from those recent DHS surveys that have collected orphanhood data on children aged 15-17, extend the analysis to investigate the relative size of the three groups of orphans in this age group.

iii) Develop and fit a regression model to predict dual orphanhood from the prevalence of maternal and paternal orphanhood by age, the lagged prevalence of HIV infection, and simple indices of the timing of parenthood in a country (i.e. the proportion married at 15-19; the proportion of married women in polygynous unions).

iv) Produce a short report on the analysis and present the results to the UNAIDS Reference Group on Estimates and Projections. If merited in the opinion of the Reference Group, produce a research note based on the report for publication in a scientific journal.

Data

Large numbers of nationally-representative household surveys conducted in sub-Saharan Africa and elsewhere since 1991 have collected data on orphaned children. Most of them have been part of the Demographic and Health Surveys (DHS) programme sponsored by USAID. The surveys use a tabular schedule to collect information on members of the households surveyed. In most African and some other surveys since 1991 this has included the questions "Is this child's mother alive?" and "Is this child's father alive?" for every child aged 0-14. Eight recently published DHS have also collected data on orphanhood at ages 15-17 years.

The analysis uses the orphanhood data from 55 African DHS conducted in 29 countries, each of which collected data on between 8000 and 33,000 children (Table 1). This data base does not represent a random sample of African countries or of points in time. Nevertheless, the surveys cover countries from all parts of the region. The only populous African country for which no data are available is the Democratic Republic of Congo.

Table 1: Percentage orphaned of children aged less than 15 in 58 Demographic and Health Surveys in mainland sub-Saharan Africa and UNAIDS estimates of the prevalence of HIV infection in adults.

				Adult HIV prevalence	
Survey	Mother dead	Father dead	Both dead	1 year earlier	3 years earlier
Benin, 1996*	2.5	4.4	0.3	0.8	0.5
Benin, 2001	1.9	4.8	0.5	1.3	1.2
Burkina Faso, 1992*	3.2	5.5	1.0	1.8	1.8
Burkina Faso, 2003	2.6	5.1	0.7	2.0	2.1
Cameroon, 1991*	2.2	5.0	0.4	0.8	0.3
Cameroon, 1998*	2.9	6.8	0.6	5.8	4.6
Cameroon, 2004	2.9	6.8	0.7	5.6	5.9
Central African Rep. 1994-95*	3.8	7.5	0.8	3.9	2.6
Chad, 1996-97*	2.5	5.5	0.5	2.1	1.5
Chad, 2004†	2.6	5.0	0.6	3.4	3.3
Congo (Brazzaville), 2005†	3.2	5.5	1.0	3.7	4.1
Côte d'Ivoire, 1994*	2.0	4.4	0.4	4.3	2.7
Eritrea, 1995*	4.5	8.0	0.7	0.7	0.3
Eritrea, 2002	4.0	6.6	0.8	1.2	1.2
Ethiopia, 2000*	4.0	7.3	0.8	3.0	3.0
Ethiopia, 2005	3.7	7.0	1.0	2.2	2.6
Gabon, 2000*	2.5	3.8	0.4	5.0	3.8
Ghana, 1998*	2.3	4.0	0.4	2.1	1.7
Ghana, 2003	2.2	5.0	0.5	2.1	2.2
Guinea, 1999*	3.0	5.9	0.9	0.9	0.6
Guinea, 2005	2.9	5.6	1.1	1.4	1.3
Kenya, 1993*	1.7	5.7	0.3	7.3	4.6
Kenya, 1998*	2.7	7.7	0.9	9.9	9.8
Kenya, 2003	3.8	9.5	2.2	7.2	8.5
Lesotho, 2004†	7.7	22.9	4.4	23.4	23.7
Malawi, 1992*	4.2	5.7	1.1	6.2	3.6
Malawi, 2000*	4.9	8.4	1.9	12.7	12.5
Malawi, 2004†	5.4	10.8	2.9	11.9	12.4
Mali, 1995-96*	2.1	3.8	0.4	0.7	0.4
Mali, 2001	2.1	3.7	0.6	1.4	1.2
Mauritania, 2000-01	3.0	5.1	0.8	0.5	0.2
Mozambique, 1997*	5.3	7.9	1.0	5.4	3.6
Mozambique, 2003†	3.8	7.5	1.3	10.8	9.3
Namibia, 1992*	2.0	5.7	0.4	1.8	0.8
Namibia, 2000	3.5	9.0	1.1	13.5	10.8
Niger, 1992*	3.1	4.1	0.4	0.1	0.1
Niger, 1998*	2.3	3.5	0.3	0.5	0.3
Niger, 2006†	2.5	3.7	0.5	0.8	0.8
Nigeria, 2003	2.6	4.3	0.6	3.1	3.0
Rwanda, 1992*	3.0	7.3	0.7	8.0	7.3
Rwanda, 2005	5.7	14.8	3.0	3.4	4.2
Senegal, 1992-93*	2.1	4.6	0.4	0.1	0.1
Senegal, 2005	2.4	5.7	0.7	0.6	0.5
South Africa, 1998*	2.2	8.9	0.8	10.5	6.1
Tanzania, 1992*	2.3	5.2	0.4	5.4	3.8
Tanzania, 1996*	2.9	6.3	0.6	7.2	6.6
Tanzania, 1999*	3.4	6.5	1.1	7.2	7.3
Tanzania, 2004†	3.2	6.3	0.9	6.5	6.8
Togo, 1998*	2.9	6.6	0.6	3.3	2.9
Uganda, 1995*	4.9	10.1	1.9	12.2	13.3
Uganda, 2000-01*	5.2	9.6	2.3	8.5	9.8
Uganda, 2006†	5.4	10.5	2.6	5.8	6.6
Zambia, 1992*	2.8	5.7	0.6	11.7	6.0
Zambia, 1996*	4.4	8.9	1.5	16.1	15.3
Zambia, 2001-02	5.9	12.0	2.9	15.2	15.6
Zimbabwe, 1994*	2.6	7.4	0.7	25.9	18.5
Zimbabwe, 1999*	4.7	11.8	2.2	31.5	31.6
Zimbabwe, 2005-06†	8.5	19.0	5.7	21.0	25.0

* Surveys included in the 2002 analysis; † Surveys that collected data on orphanhood at ages 15-17.

Although DHS surveys in Haiti and Dominican Republic have included questions about orphanhood, ages at marriage and childbearing differ markedly between the Caribbean and sub-Saharan Africa and the main analysis was restricted to the latter region. In addition, for a variety of reasons, no use was made of the data from several African DHS that have asked about orphanhood. The 2000 survey of Rwanda documented exceptionally high levels of orphanhood that reflect the genocide of 1994. It was excluded from the analysis, as were the data on children aged 10-14 from the 2005 survey of the same country. The DHS conducted in Nigeria in 1999 was dropped because it had a 6.5 per cent non-response rate for questions about parental survival and internal consistency checks suggest that a disproportionate number of the non-responders had dead fathers. Moreover, the 1993 Ghana data were not used because this survey reported an implausibly high level of dual orphanhood that subsequent surveys of the same country have failed to confirm. Not only populations experiencing generalized HIV epidemics but those with an adult HIV prevalence of between 0.5 per cent and 1 per cent were included in the data set to which the final regression models were fitted (see Table 1). However, the 1992 and 1998 surveys in Niger and the 1992-93 survey in Senegal were omitted as adult HIV prevalence in these countries was still less than 0.5 per cent at the time that these enquiries were conducted.

Methods[1]

Dual orphanhood should be more common where the prevalence of maternal and/or paternal orphanhood is high. If the risk of a child's mother dying and father dying were independent, the expected proportion of dual orphans among children of any age would be simply the proportion with dead mothers multiplied by the proportion with dead fathers. In practice, the observed prevalence of dual orphanhood is always higher than this because the two parents of a child are usually of much the same socio-economic status, exposed to many of the same environmental risks, and at risk of dying together in an accident or episode of violence. Direct transmission of infections such as tuberculosis and HIV between a child's parents also occurs.

Dual orphanhood can be described as the product of its expected prevalence, given the prevalence of maternal and paternal orphanhood, and the excess risk of dual orphanhood relative to this expected risk. The excess risk at age *a* in a survey conducted at time *t*, $E_{t,a}$, is the ratio of the observed ($O_{t,a}$) to the expected ($\tilde{O}_{t,a}$) number of dual orphans aged *a* at time *t*. If $C_{t,a}$ is the total number of children aged *a* at time *t*, then,

$$E_{t,a} = \frac{O_{t,a}}{\tilde{O}_{t,a}} = \frac{\dfrac{O_{t,a}}{C_{t,a}}}{\dfrac{\omega_{t,a}}{C_{t,a}} \cdot \dfrac{\psi_{t,a}}{C_{t,a}}} \qquad \ldots(1)$$

where $\omega_{t,a}$ and $\psi_{t,a}$ represent maternal and paternal orphans. Once the determinants of the excess risk have been modelled using data from household surveys, forecasts of dual orphanhood can be made from those of maternal and paternal orphanhood within the *Spectrum* projections package (Stover *et al.* 2006) using methods described by Grassly and Timaeus (2005).

---

[1] This section is reproduced with revisions from Grassly, N. C. and I. M. Timæus. 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(3): 365-375.

The number of dual orphans of a given age in a particular population has been assumed to be a Poisson random variable with mean equal to the number of children of that age multiplied by an underlying risk of being a dual orphan, $\lambda_{t,a}C_{t,a}$. If the equivalent risk of dual orphanhood assuming independent risks of maternal and paternal orphanhood is $\tilde{\lambda}_{t,a}$, then:

$$E_{t,a} = \frac{\lambda_{t,a}}{\tilde{\lambda}_{t,a}} \qquad \ldots(2)$$

and, taking the logarithm of each side, Equation 1 can be rearranged to give:

$$\ln O_{t,a} = \ln E_{t,a} + \ln C_{t,a} + \ln \tilde{\lambda}_{t,a} \qquad \ldots(3)$$

The last term of the latter equation is the log excess risk of dual orphanhood relative to the risk if the probabilities of a child's mother and father being dead were independent. To establish whether the fitted model can be used to make projections, it is important to establish the extent to which the excess risk of dual orphanhood by age, between countries or over time. Therefore, log dual orphanhood was modelled using additive multilevel or random-intercept Poisson regression:

$$\ln O_{t,a} = \theta_0 + \ln C_{t,a} + \theta_1 \ln \tilde{\lambda}_{t,a} + \theta_2 Age_i + \sum_x \theta_{x,t,k} X_{t,k} + \upsilon_k + u_t + \gamma_i \qquad \ldots(4)$$

where the subscripts *i* and *k* refer to individual children and countries respectively, the θs are coefficients to be estimated, the $X_{t,k}$ are covariates measured at the population level, υ_k and u_t are country-level and survey-level distributions of a random intercept term (assumed to be normally distributed with variances σ^2_k and σ^2_t), indicating the degree to which the excess risk of dual orphanhood varies between countries and over time within countries, and γ_i is an individual-level residual that is permitted to have extra-Poisson variation, ω.

The model was fitted to DHS data on the percentages of children orphaned in the age groups 0-2, 3-5, 6-9, and 10-14 years, together with 15-17 years where available. However, age is modelled as a continuous covariate so that the fitted model can be used to predict dual orphanhood at any age or for any age group. The model was fitted using the package *MLWIN 2.02* using restricted iterative generalized least squares. We experimented with various specifications of the regression model, including different assumptions about both the population-level and individual error distributions, and fitted it using the negative-binomial random-effects regression procedure in *Stata 10* as well as in *MLWIN*. All the approaches yielded similar conclusions and parameter estimates to those presented here.

Covariates that seem likely to affect the excess risk of dual orphanhood in a population include the ages of the children's parents and the severity of the AIDS epidemic in the country. The model includes only simple aggregate indicators of these factors that it should be possible to estimate for any African population. Two indices of marriage patterns were used as proxies for parents' ages at the birth of their children - the proportion of women aged 15-19 who are married and the proportion of married women in monogamous unions. (The prevalence of polygyny is a major determinant of the average age difference between husbands and wives in Africa). These indices were calculated from data collected in the DHS surveys that also measured orphanhood.

Estimates of log adult HIV prevalence in the 12 years preceding each survey that underlie the statistics published by UNAIDS were assessed as possible indices of the severity of the AIDS

			Mother		
Father			Dead		Alive
			AIDS	Other	
	Dead	AIDS	Dual AIDS orphan	Dual AIDS orphan	Paternal AIDS orphan
		Other	Dual AIDS orphan	Other dual orphan	Other paternal orphan
	Alive		Maternal AIDS orphan	Other maternal orphan	

Figure 1: Classification of children according to the survival status of each parent, and the corresponding definitions of a dual AIDS orphan, dual non-AIDS orphan, maternal AIDS orphan, and paternal AIDS orphan.

epidemic affecting each population. Of course, the prevalence of dual orphanhood is not directly determined by adult HIV prevalence. Rather, the numbers of both orphaned children and infected adults reflect past trends in the age-specific incidence of HIV infections and AIDS deaths. With 12 lagged prevalence rates as explanatory variables, each of which has a relationship with dual orphanhood that potentially varies by age, and a sample of aggregated observations (age groups by surveys) of only 227 cases, overfitting of the data is a potential problem. Exploratory work showed that measuring HIV prevalence at two points in time is sufficient to capture those features of the epidemic that affect dual orphanhood. In particular, although the recent level and trend in HIV prevalence are determined by three parameters in *EPP* (Brown *et al.* 2006), they are described sufficiently well to forecast dual orphanhood by a single prevalence measure. More complex parameterizations of the epidemic have not been used to predict children's risk of dual orphanhood even when they provided small, statistically significant improvements in the fit of the regression model to the set of available surveys.

UNAIDS defines a dual AIDS orphan as a child whose parents are both dead, at least one due to AIDS. This statistic can be calculated from the estimated number of dual orphans due to all causes by subtracting the number of dual orphans whose parents both died from causes other than AIDS (Figure 1). Thus, the number of dual AIDS orphans aged a at last birthday at time t is:

$$\hat{O}'_{t,a} = \hat{O}_{t,a} - \hat{O}''_{t,a} = \left(\omega_{t,a}.\psi_{t,a}.\hat{E}_{t,a} - \omega''_{t,a}.\psi''_{t,a}.\hat{E}''_{t,a}\right)/\,C_{t,a} \qquad \ldots(5)$$

where the double primes distinguish non-AIDS from total orphans.

Results

The proportion of all children aged less than 15 who were maternal orphans ranges from 1.7 per cent in Kenya in 1993 to 8.5 per cent in Zimbabwe at the time of the 2005-06 survey. The proportions of children whose father had died are much higher, ranging from 3.5 per cent in

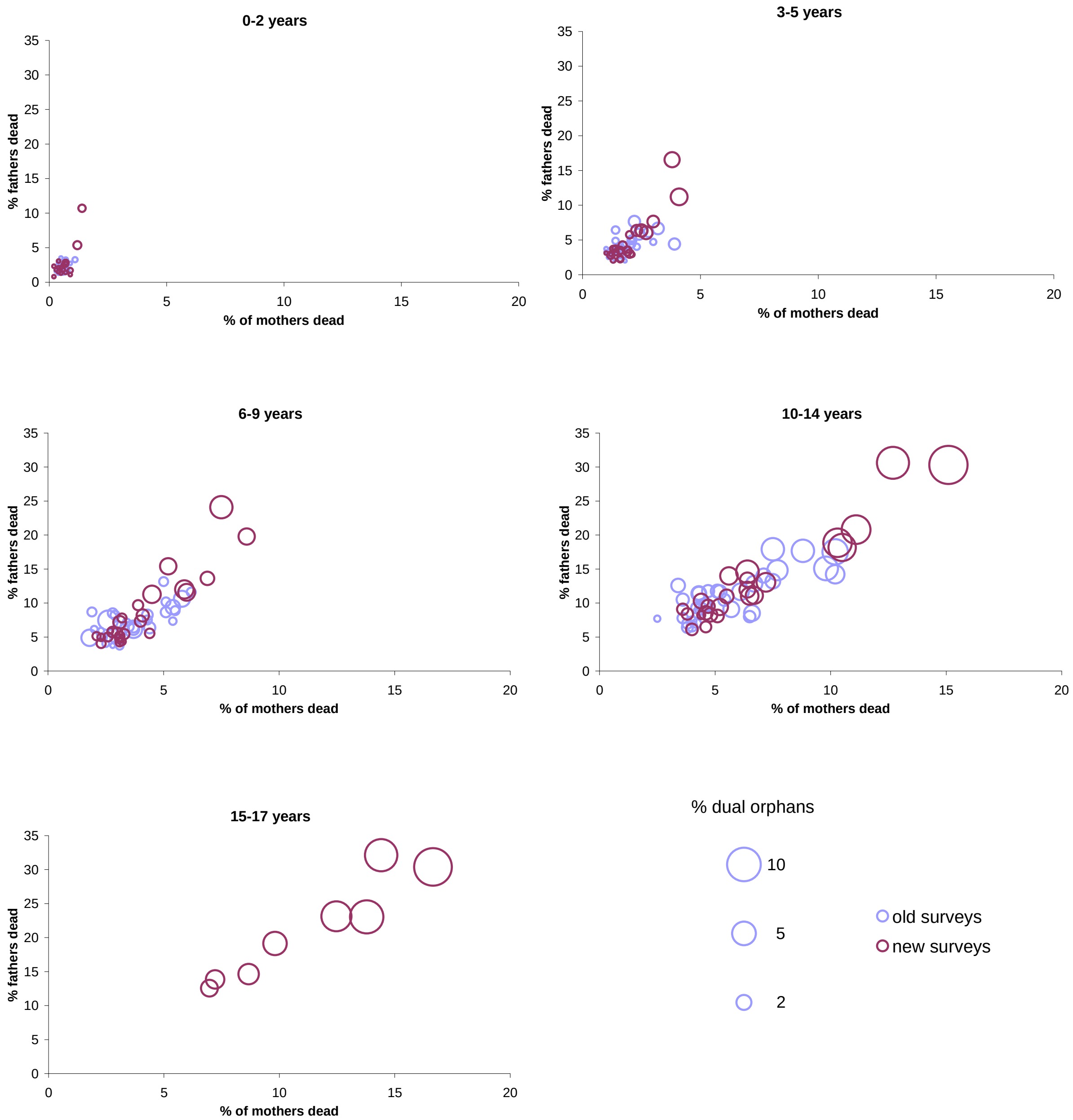

Figure 2: Percentage by age group of children that are dual orphans according to the percentages that are maternal and paternal orphans, 58 Demographic and Health Surveys in mainland sub-Saharan Africa.

Niger in 1998 to 22.9 per cent in Lesotho in 2004. The relatively high prevalence of paternal mortality reflects the facts that men in sub-Saharan Africa tend to be considerably older than their wives and that in most of the region men suffer considerably higher mortality than women from causes other than AIDS. Most orphaned children had lost only one of their parents. The proportion of children aged less than 15 whose parents were both dead ranges from 0.3 per cent in Kenya in 1993 to 5.7 per cent in Zimbabwe in 2005-06.

Figure 2 examines the relationship between maternal, paternal and dual orphanhood by the age group of the children. The area of the bubbles indicates the prevalence of dual orphanhood. The first important point that this Figure 2 emphasizes is that the proportion of children who have been orphaned rises rapidly with age. On average, 0.5 per cent of babies and toddlers aged 0-2 are maternal orphans, 1.8 per cent of them are paternal orphans and only 0.1 per cent of them are

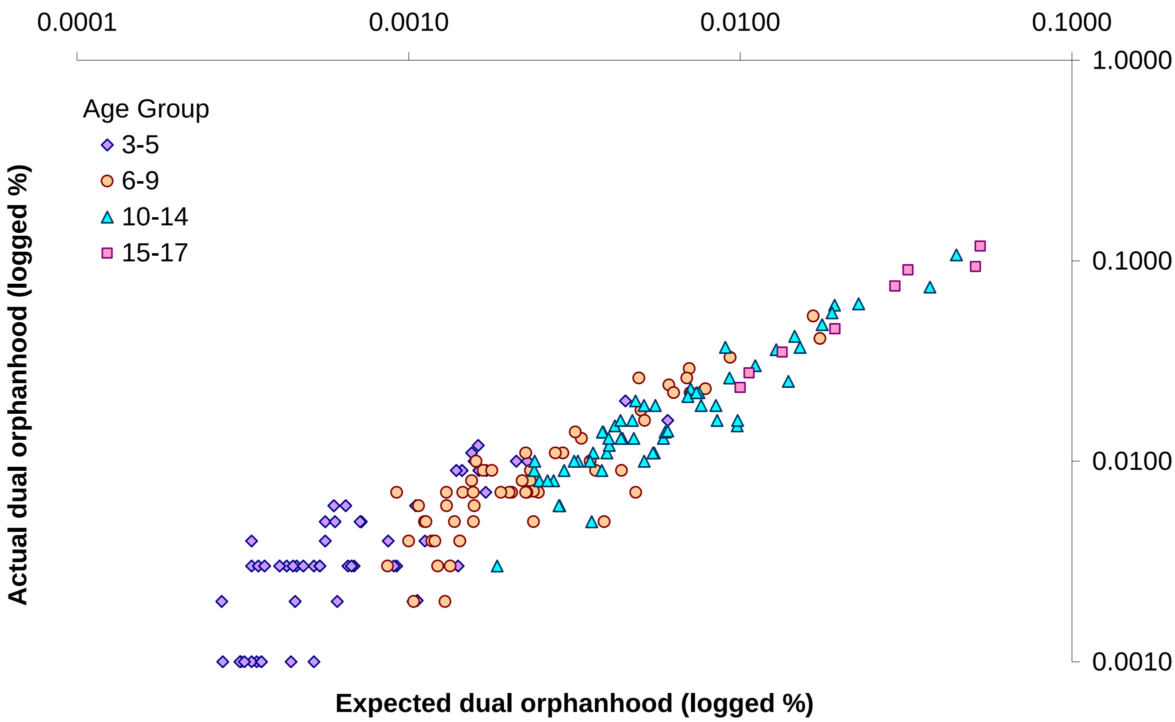

Figure 3: Relationship by age group between the proportion of children whose parents have both died and the equivalent proportion if the risks of dying of the mother and father were independent, 58 Demographic and Health Surveys in mainland sub-Saharan Africa.

dual orphans. These statistics rise to 1.7 per cent, 3.6 per cent and 0.3 per cent respectively among children aged 3-5; and to 3.3 per cent, 6.5 per cent and 1.7 per cent at ages 6-9. On average, 5.2 per cent of children in the age group 10-14 years are maternal orphans, 11.0 per cent paternal orphans and 1.4 per cent dual orphans. While the prevalence of orphanhood at ages 15-17 is very high in those countries in which a recent DHS has collected data on this age group, this finding probably reflects the fact that these are high mortality populations as much as it does the continuing rise in the prevalence of orphanhood with age.

Figure 2 also reveals that levels of maternal, paternal, and dual orphanhood are closely associated in every age group. The countries in which many children's mothers had died are also those in which many children's fathers had died. In the age groups 0-2 and 3-5 years, two countries stand out as having exceptionally high levels of both paternal and dual orphanhood. They are Lesotho in 2004 and Zimbabwe in 2005-06. These two populations had higher levels of maternal, paternal, and dual orphanhood than any of the other populations in every age group.

Comparison of the blue and purple bubbles on the plots in Figure 2 provides some indication of what has been gained by updating this analysis to include the results of recent DHS. Many of the additional surveys included in the analysis were conducted in populations in which a relatively high proportion of children were orphaned. In several of them, orphanhood was more prevalent than in any of the surveys that had been published by 2002.

Figure 3 examines the relationship between the natural logs of the number of children that were dual orphans and the number of them that one would expect to be dual orphans if the risks of dying of the children's two parents were independent. A strong association exists between the observed prevalence of dual orphanhood and the levels of maternal and paternal orphanhood as captured in the index "expected dual orphanhood". A straight line fitted to the points in Figure 3

Table 2: Parameter estimates (with standard errors) for random-intercept Poisson regression models of the percentage of dual orphans among children aged 0-17, Demographic and Health Surveys in mainland African countries with generalized epidemics of HIV.

Variables	Coefficients	Standard errors
Intercept	1.3478	0.1563
ln(% of children who are expected dual orphans)	0.7539	0.0283
Children's age	-0.02108	0.0131
ln(% HIV prevalence in year *t*-1)*	-0.1506	0.0540
Children's age × ln(% HIV prevalence in year *t*-1)*	0.007225	0.0050
ln(% HIV prevalence in year *t*-12)*	0.09887	0.0113
Inter-country level variance of the intercept (υ_k)	0.0196	0.0059
Inter-survey level variance of the intercept (u_l)	0.0000	0.0000
Extra-Poisson variation (ω)	1.7701	0.2071

*Past adult HIV prevalence rates of <0.5% should be set to 0.5% to predict dual orphanhood.

will have a positive intercept. In other words, more children were dual orphans than would be expected if their parents' chances of dying were independent. The slope of the fitted line will be less than one, a pattern that was not evident in the results of the surveys available in 2002. This pattern suggests that the actual number of dual orphans is a greater multiple of the expected number when maternal and paternal orphanhood are rare than when they are common. It implies that log expected orphanhood should be incorporated into a model of dual orphanhood as an explanatory variable, rather than an offset term with a coefficient set to one.

The points for the different age groups of children included in Figure 3 fall on approximately the same line. Thus, while young children are far less likely to be orphaned than older children, the variation in the percentage of children who are dual orphans by age is largely a function of the variation by age in the proportions of them that are maternal and paternal orphans. Age has a limited independent effect on the probability that a child is a dual orphan after controlling for expected dual orphanhood.

Table 2 presents a Poisson regression model that predicts the percentage of children that are dual orphans from the expected percentage and other variables. The model has been fitted to DHS data on populations experiencing a generalized epidemic of HIV infection and is designed to predict dual orphanhood in such populations. Because dual orphanhood is predicted from the logarithms of adult HIV prevalence the model can produce implausible out-of-sample predictions as HIV prevalence tends to zero. Nevertheless, it should produce reasonable estimates of dual orphanhood in low HIV prevalence populations so long as estimates of adult HIV prevalence of less than 0.5 per cent are replaced by a lower-bound figure of 0.5 per cent.

As one would expect from the exploratory analysis, expected dual orphanhood (i.e. the proportion of children in an age group who have dead mothers multiplied by the proportion of them who have dead fathers) is less common than observed dual orphanhood. Nevertheless, expected dual orphanhood is strongly predictive of the actual percentage of children who are dual orphans.

Experimentation with different specifications of the model (not shown) revealed that the indices of the ages of parents relative to their children, that is the proportions of 15-19 year old women who are married and the proportions of married women who are in polygynous unions, only

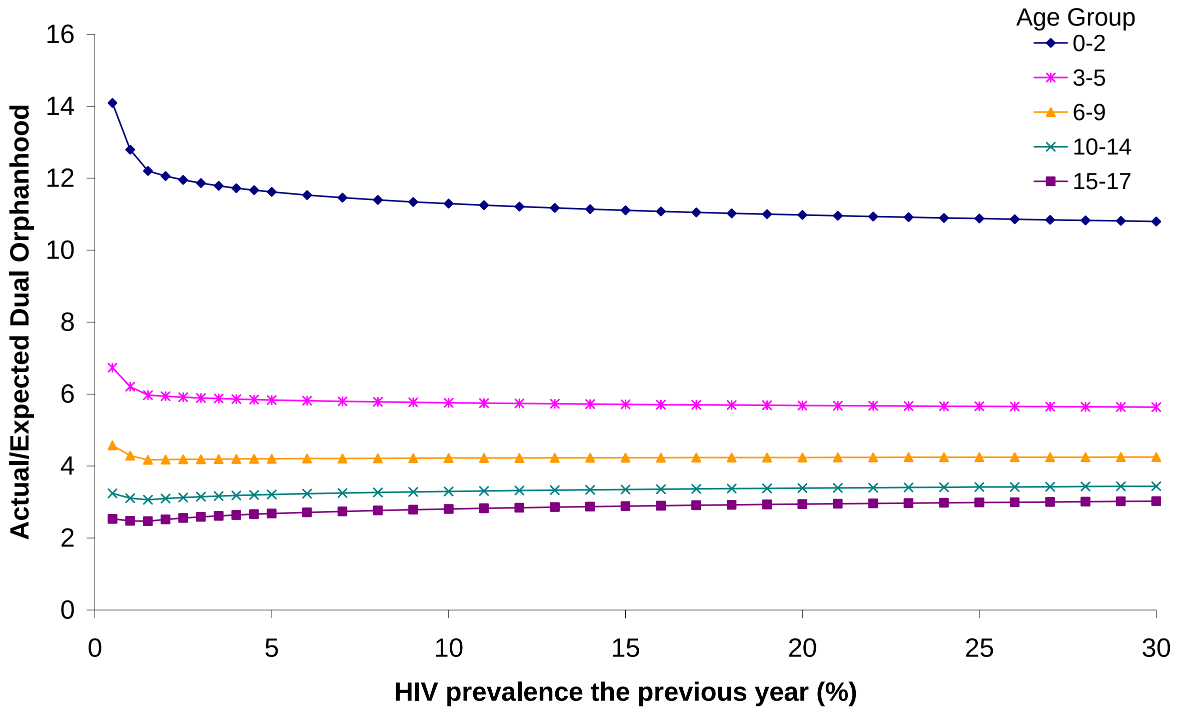

Figure 4: Fitted estimates of the ratio of the excess to the expected risk of dual orphanhood according to HIV prevalence by age group of children, mainland African countries with generalized HIV epidemics.

predict dual orphanhood in models that omit the indicator of expected dual orphanhood. In other words, all forms of orphanhood tend to be relatively common in populations in which adults bear children late in life. However, the ages at which men and women bear children do not modify to any extent the relationship between the prevalence of dual orphanhood and prevalence of maternal and paternal orphanhood.

The risk of dual orphanhood is associated in a rather complex way with the history of the AIDS epidemic in the population concerned. The negative coefficient on HIV prevalence indicates that the excess risk of dual orphanhood is highest when adult mortality and the expected risk of dual orphanhood are low. As AIDS mortality rises and pushes up maternal and paternal orphanhood, a more muted rise occurs in the percentage of children that are dual orphans than would be predicted from the multiplicative rise in expected dual orphanhood alone. The positive coefficient on HIV prevalence 12 years earlier implies that other things being equal, countries with rapidly growing HIV epidemics will have a higher excess risk of dual orphanhood than those where the rise in HIV prevalence has been slower or HIV prevalence has stabilised.

Age has a negative effect on the percentage of children that are dual orphans after controlling for expected dual orphanhood, particularly in populations in which HIV infection is uncommon. The inter-related effects of age and HIV prevalence on the risk of dual orphanhood are shown graphically in Figure 4. This figure plots the ratio of actual to expected dual orphans according to adult HIV prevalence 1 year earlier for five different age groups of children. In order to calculate these estimates, it is assumed that the population concerned has experienced a typical HIV epidemic with the prevalence of infection among adults rising at a gradually slowing rate for the previous 12 years. It is also assumed that the prevalence of maternal and paternal orphanhood by age, and therefore expected dual orphanhood, is typical of the data collected in the 58 surveys.

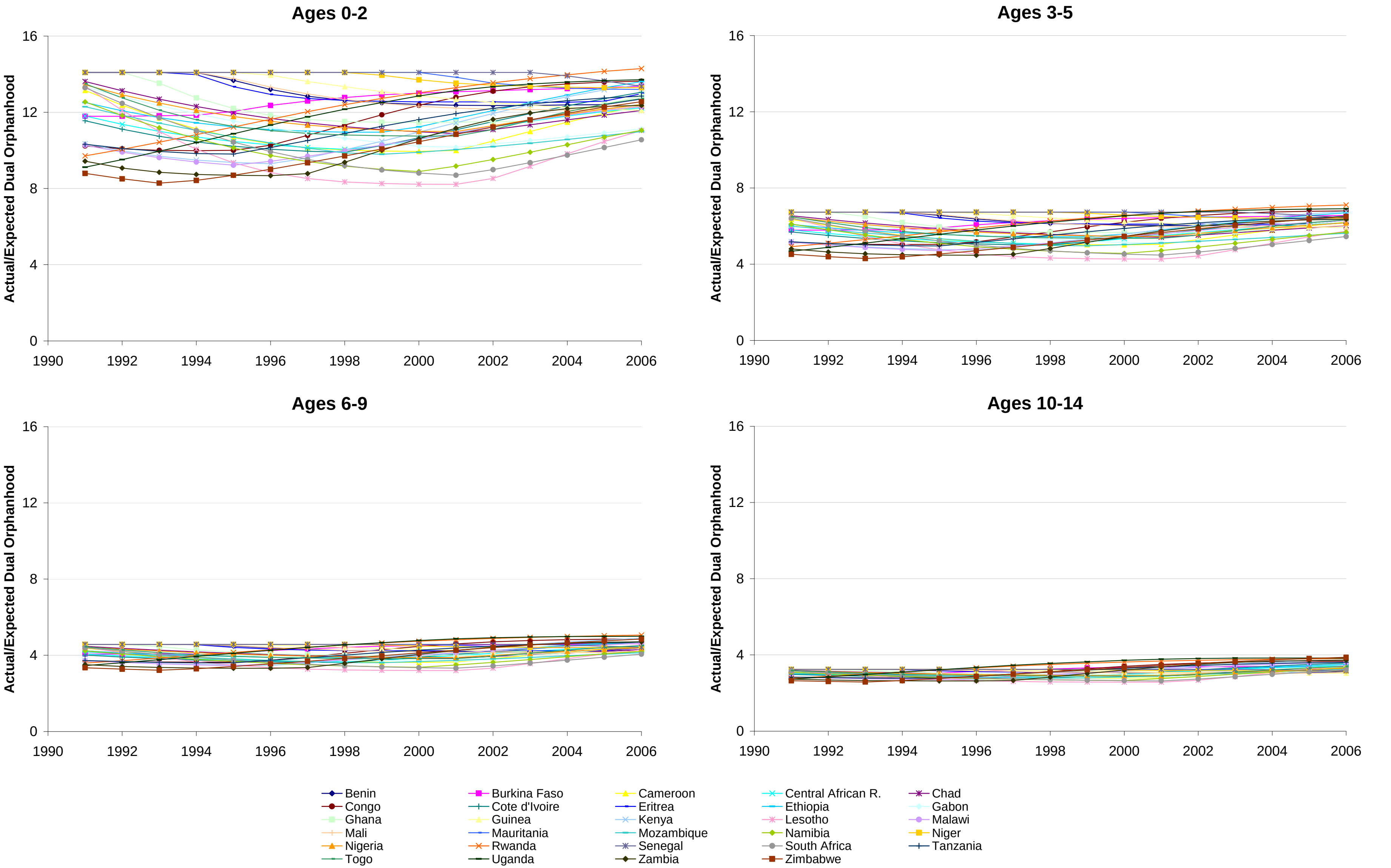

Figure 5: Fitted estimates of the trend in the ratio of the excess to the expected risk of dual orphanhood in 29 mainland African countries that have conducted Demographic and Health Surveys, by age group of children.

As can be seen from Figure 4, the excess risk over the expected risk of dual orphanhood tends to decline both with age and as adult HIV prevalence rises. However, the excess risk of dual orphanhood varies hardly at all with HIV prevalence for children aged 10 or more. Even for young children, the excess risk of dual orphanhood varies only across populations with relatively small HIV epidemics by African standards. About three times more teenaged children are dual orphans in most African populations than if the risks of dying of their mothers and fathers were independent. The proportion of young children who are dual orphans is much higher, however, than would be the case if those who have lost one parent were no more likely than other children to have lost their other parent. In a typical country in which adult HIV prevalence has risen to 6 per cent, for example, the children's predicted risk of being dual orphans is 11.5 times the expected risk in the age group 0-2 years, 5.8 times the expected risk at ages 3-5, and 4.2 times the expected risk at ages 6-9.

Figure 5 presents fitted estimates for four different age groups of children of trends in the ratio of the excess to the expected risk of dual orphanhood in the 29 mainland countries that have collected orphanhood data in one or more Demographic and Health Surveys. It can be seen that the excess risk of dual orphanhood is not only higher at younger ages than older ones, but also both absolutely and relatively more variable between countries. Apart from its dampening effect, the interaction term between age and HIV prevalence in the fitted regression model has only a minor effect on the trends in the ratios. Thus, the evolution over time in the excess risk of dual orphanhood in a particular country follows broadly the same pattern in each age group. In Figure 5, it can be observed most readily in the plots for children aged 0-2 years. These trends vary between countries depending on history of the HIV epidemic in that country as estimated in *EPP*. In general though, the excess risk of dual orphanhood tends to fall fairly sharply in the early years of the epidemic but then to stabilize and to begin to rise again as adult mortality increases and the rise in the prevalence of HIV infection slows or is reversed. Only Rwanda is the *excess* percentage of children who are dual orphans higher than before the country first developed a generalized epidemic of HIV. Nevertheless, because the increases in the prevalence of maternal and paternal orphanhood that result from high AIDS mortality have a multiplicative effect on the *expected* risk of dual orphanhood, the number of dual orphans has grown faster than that of other orphans.

Having controlled for HIV prevalence, the residual variance in the prevalence of dual orphanhood between different surveys in the same country is tiny and statistically insignificant. However, the residual variance between countries in the data remains appreciable and significant. This indicates that differences in the prevalence of dual orphanhood exist between countries that are not explained entirely by the either the prevalence of maternal and paternal orphanhood, the history of the HIV epidemic in that country, or the marriage variables (which failed to improve the predictive power of the model at all). The size of the errors that are likely to occur in predictions of excess dual orphanhood as a result of this unexplained variation between countries is indicated by Table 3. Care should be taken in interpretation of these indicators because, as has been emphasized already, the DHS surveys were not conducted on a probability sample of the populations of all African countries at all dates since the onset of the AIDS epidemic. Bearing this in mind, one can tentatively conclude that, if the assumptions made about the prevalence of maternal and paternal orphanhood and the history of the HIV epidemic are correct, the regression model should usually predict the numbers of dual orphans to within ±4 per cent.

Table 3: Estimates of the excess risk of dual orphanhood and their confidence intervals from the random-intercept Poisson regression model at two levels of HIV prevalence, mainland African countries with generalized epidemics of HIV.

	1% HIV prevalence			25% HIV prevalence		
Children's Age Group	Lower 95% Confidence Interval	Mean	Upper 95% Confidence Interval	Lower 95% Confidence Interval	Mean	Upper 95% Confidence Interval
0-2	12.3	12.8	13.3	10.5	10.9	11.3
3-5	6.0	6.2	6.4	5.4	5.7	5.9
6-9	4.1	4.3	4.5	4.1	4.2	4.4
10-14	3.0	3.1	3.2	3.3	3.4	3.5
15-17	2.4	2.5	2.6	2.9	3.0	3.1

Discussion

This report updates the regression model that UNAIDS and other agencies use to calculate their estimates of dual orphans in mainland African countries experiencing generalized HIV epidemics. The overlap between the groups of maternal and paternal orphans comprises orphans whose parents have both died. Multiplying together for a particular country and year, the proportions of children of each age that are maternal and paternal orphans gives the proportion of children of that age who would be dual orphans if the risks of each child's two parents dying were independent. The regression model presented here predicts the percentage of children that are dual orphans as a function of this independent or expected risk of dual orphanhood and the severity and rate of spread of the HIV and AIDS epidemic in the country as measured by the percentage of the adult population infected with HIV.

This regression model is not intended to predict how likely particular children are to be dual orphans, but the population prevalence of dual orphanhood. Bearing in mind the caveat that the DHS are not a random sample of African populations, with accurate data on maternal and paternal orphanhood and past HIV prevalence, the regression model will usually predict what percentage of children are dual orphans to within ±4 per cent. This uncertainty reflects the inability of the model to completely explain the variation between countries in the prevalence of dual orphanhood. This degree of imprecision in the model is likely to have less impact on predictions produced using it than errors in the estimates of the input variables. Moreover, the excess risk of dual orphanhood, relative to the independent risk, only varies significantly according to the severity of the HIV epidemic in a country at very young ages (Figure 4) and very few babies and toddlers are dual orphans. Thus, most of the error in the estimates of dual orphanhood produced using this regression model will arise from the estimates of maternal and paternal orphanhood used to generate them. The latter estimates, in turn, depend on a large number of assumptions. The major source of error in them, however, is likely to be inaccurate estimates of the number of adults dying from AIDS and other causes each year (Grassly and Timæus 2005).

Given the same input data, the updated model will produce different estimates of the number of dual orphans from the existing one. These differences reflect the joint impact of three sets of changes. First, the model has been fitted to data from 21 more DHS than previously (24 surveys have been added to the dataset and 3 surveys conducted in countries where adult HIV prevalence was less than 0.5 per cent have been dropped from it). The new dataset has a wider geographical coverage than the previous set. Moreover, the additional surveys have all been conducted since 2000 and have added substantially to the information on orphanhood in populations with severe HIV epidemics and high adult mortality. Second, the specification of the regression model has been revised. In particular, exploratory analysis of the enlarged dataset revealed that the expected prevalence of dual orphanhood (maternal orphanhood multiplied by paternal orphanhood) should be included in the model as a variable rather than an offset term with a coefficient of one. This has produced a simpler yet more powerful model. Third, the model has been fitted using UNAIDS current (i.e. 2007) estimates of the growth of the HIV epidemic among adults rather than their 2002 estimates of past HIV prevalence. Using these estimates, one would obtain different regression coefficients even if one was to refit the previous model to the original set of DHS surveys. All three of these factors have contributed to both changing the estimates of excess risk and the narrowing of the confidence interval around the predictions.

Table 4 presents equivalent statistics to those in Table 3 but for the previous version of this regression model. At low levels of HIV prevalence, the revised estimates of the excess relative to the independent risk of dual orphanhood are considerably higher than the previous ones. The discrepancy is greatest at young ages. For example, at 1 per cent prevalence, the predicted number of dual orphans aged less than 3 is 86 per cent higher (i.e. 12.8/6.7). The overall impact on this on estimates of dual orphans will be small, however, as less than one in a thousand of these very young children are dual orphans in all Africa countries except those with very severe HIV epidemics. In countries with more severe epidemics the discrepancies between the predictions from the revised and earlier models are smaller, particularly after the first few years of life. When adult HIV prevalence is 25 per cent, for example, as shown in Tables 3 and 4, estimates of dual orphans made using the new model will be about 10 per cent higher than the previous ones. Thus, the relative increase in estimates of numbers of dual orphans resulting from the use of the new model is small when the absolute number of dual orphans is large.

It is somewhat perturbing that the previous model predicted a greater excess risk of dual orphanhood in populations where HIV infection was highly prevalent than in those where it was not, but the revised model predicts the reverse pattern. Thus, in the current model the coefficient on HIV prevalence in year *t*-1 is negative whereas, in the previous, one the coefficient on HIV prevalence in year *t*-5 was positive. Investigation of the reason for this reveals that neither the incorporation into the analysis of additional data on populations with severe HIV epidemics, nor revisions to UNAIDS estimates of HIV prevalence in particular countries are responsible.

Table 4: Estimates of the excess risk of dual orphanhood and their confidence intervals from the random-intercept Poisson regression model for two levels of HIV prevalence, based on 34 Demographic Surveys conducted by 2000 in mainland sub-Saharan Africa.

	1% HIV prevalence			25% HIV prevalence		
Children's Age Group	Lower 95% Confidence Interval	Mean	Upper 95% Confidence Interval	Lower 95% Confidence Interval	Mean	Upper 95% Confidence Interval
0-2	6.3	6.9	7.5	8.4	9.2	10.0
3-5	3.8	4.1	4.5	5.0	5.5	6.0
6-9	2.5	2.8	3.0	3.4	3.7	4.0
10-14	2.1	2.3	2.5	2.8	3.1	3.4

Source: Grassly and Timæus (2005).

Moreover, neither allowing the coefficient of the expected dual orphanhood term to vary from one nor dropping the marriage variables from the regression equation greatly affects the estimated effect of HIV prevalence. Instead, the crucial revision turns out to be taking into account more of the variability across Africa in the unfolding of HIV epidemics by using HIV prevalence in year *t*-12, as well as more recently, to predict how many of the maternal and paternal orphans are also dual orphans.

Although the regression model was fitted to African data, the relationship between observed and expected dual orphanhood in the two other countries with generalized epidemics of HIV where the DHS have collected orphanhood data, Haiti and Dominican Republic, is similar to that depicted for African countries in Figure 3. As indices of the ages of parents are no longer included among the predictors, there seems no reason not to use this regression model to model the percentage of children that are dual orphans in such countries.

The revised method for estimating dual orphanhood described in this report has benefited from an enlarged database on orphanhood in populations with severe AIDS mortality, from improvements in the quality of African HIV surveillance data and in the functionality of *EPP*, and from changes to the specification of the regression model. The new model produces substantially more precise estimates of the what percentage of children are dual orphans than the old one without any need to use indices of marriage patterns in a country as predictors.

The validity of the predictions made by this regression model is contingent on the accuracy of the DHS data on orphans. As household surveys, the DHS are not designed to enumerate the small number of orphans living in institutions or “on the street”. Interviewers may also fail to find out about all of the orphans among the children living in foster families. On the other hand, various attempts to assess these data have not found evidence of widespread gross biases in them (Grassly *et al.* 2004). Moreover, the regression model is not the most important determinant of the estimated prevalence of dual orphanhood. It is the estimates of AIDS deaths made from surveillance and survey data on HIV prevalence, and the estimates of maternal and paternal orphanhood that are derived from them, that largely determine how many children are dual orphans.

References

Brown, T., N. C. Grassly, G. Garnett, and K. Stanecki. 2006. Improving projections at the country level: the UNAIDS Estimation and Projection Package 2005, *Sexually Transmitted Infections* 82(supplement 3): iii34-iii40.

Grassly, N., T. Boerma, G. Garnett, S. Gregson, and B. Zaba. 2000. Recommendations from the UNAIDS Epidemiology Reference Group Meeting, Rome, October 8th-10th 2000. Available: http://www.epidem.org/Publications/Rome2000rec.pdf (accessed: 1st January, 2008).

Grassly, N. C., J. J. C. Lewis, M. Mahy, N. Walker, and I. M. Timæus. 2004. Comparison of survey estimates with UNAIDS/WHO projections of mortality and orphan numbers in sub-Saharan Africa, *Population Studies* 58(2): 207-217.

Grassly, N. C. and I. M. Timæus. 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(3): 365-375.

Gregson, S., G. P. Garnett, and R. M. Anderson. 1994. Assessing the potential impact of the HIV-1 epidemic on orphanhood and the demographic structure of populations in sub-Saharan Africa, *Population Studies* 48(3): 435-458.

Hunter, S. and J. Williamson. 2000. Children on the Brink 2000: Executive Summary, Updated Estimates and Recommendations for Intervention. USAID. Available: http://www.usaid.gov/pubs/hiv_aids/childrenreport.pdf (accessed: 1st January, 2008).

Paget, W. J. and I. M. Timæus. 1994. A relational Gompertz model of male fertility: development and assessment, *Population Studies* 48(2): 333-340.

Stover, J., N. Walker, N. C. Grassly, and M. Marston. 2006. Projecting the demographic impact of AIDS and the number of people in need of treatment: updates to the Spectrum projection package, *Sexually Transmitted Infections* 82(Supplement 3): iii45-iii50.

UNAIDS. 2000. *Report on the Global HIV/AIDS Epidemic, June 2000*. Geneva: United Nations Joint Programme on HIV/AIDS.

UNAIDS. 2002. *Report on the Global HIV/AIDS Epidemic, July 2002*. Geneva: United Nations Joint Programme on HIV/AIDS.

UNAIDS, UNICEF, and USAID. 2002. *Children on the Brink 2002*. Washington, D. C.: UNAIDS, UNICEF, USAID.