# Orphanhood and the long-run impact on children [1]

Kathleen Beegle
The World Bank

Joachim De Weerdt
E.D.I., Tanzania

Stefan Dercon
Oxford University

September 2005

## Abstract

*This paper presents unique evidence that orphanhood matters in the long-run for health and education outcomes, in a region of Northwestern Tanzania, an area deeply affected by HIV-AIDS in Africa. We use a sample of non-orphans surveyed in 1991-94, who were traced and reinterviewed in 2004. A large proportion, 23 percent, lost one or more parents before the age of 15 in this period, allowing us to identify the impact of orphanhood shocks. Since a substantial proportion reaches adulthood by 2004, we can also assess permanent health and education impacts of orphanhood. In the analysis, we can control for a wide range of child and adult characteristics before orphanhood, as well as community fixed effects. We find that maternal death causes a permanent height deficit of about 2 cm (or 22 percent of one standard deviation) and a persistent impact on years of education of almost 1 year (or 25 percent of one standard deviation). We also find that paternal orphanhood has an impact on educational outcomes, but only for particular groups. We show evidence that living arrangements and whether the child was in school at the time of losing a parent strongly influence the impact of maternal and paternal death. We also illustrate the problems of inference on the impact of orphanhood if only children who remained in their baseline communities by 2004 had been reinterviewed.*

[1] Data collection was funded by the Knowledge for Change Partnership at the World Bank and DANIDA. This analysis was supported by the Trust Fund for Environmentally & Socially Sustainable Development at the World Bank, and the Economic and Social Research Council, United Kingdom. The views expressed here do not necessarily reflect those of the World Bank or its member countries. Please address correspondence to kbeegle@worldbank.org.

## 1. Introduction

Childhood orphanhood is considered a major risk factor for poverty in adulthood, through, among other channels, shortfalls in human capital investments in children. This paper provides unique evidence on the long-term impact of orphanhood in a region of Tanzania, near Lake Victoria in an area ravaged by HIV-AIDS. The underlying data set is a large 13-year panel data set in which individuals interviewed at baseline were traced irrespective of their current residence. This allows us to focus on non-orphaned children experiencing the loss of one or both parents during the survey period, controlling for their characteristics before becoming an orphan. Furthermore, we can focus on the permanent impact in terms of height and educational attainment once these children reach adulthood and catching-up is hardly possible anymore. We find significant permanent effects. Children who become maternal orphans before the age of 15 are 2 cm shorter in adulthood than similar children whose mother did not die during this age interval, representing 22 percent of one standard deviation of height in the sample. We also find that maternal and paternal orphanhood results in substantially lower educational attainment, each lowering years of education by adulthood by about a quarter of one standard deviation of educational attainment in the sample. This education effect is particularly pronounced for those children not in school by the time the parent dies, where schooling is lower by up to 3 years (or one standard deviation) compared to other children. We also find that the impacts of orphanhood vary according to living and fostering arrangements at time of the event.

In sub-Saharan Africa, the prevalence of orphanhood among children has been greatly exacerbated by the HIV/AIDS pandemic. While there are other, more prevalent diseases in Africa, the characteristics of HIV/AIDS suggest that its economic and demographic impact will be profound. Because HIV in Africa is transmitted primarily through heterosexual contact, the epidemic is having a major effect on the mortality of men and women in their prime childbearing and earning years; consequently, mortality rates have risen and life expectancies have fallen dramatically in Africa. As a consequence, orphanhood rates are increasing and putting a larger share of children at risk.

The available evidence of the impact of this trend typically relates to the impact of orphanhood from HIV/AIDS and other causes in the short-run, often by examining a sample of school-age children in cross-sectional survey data. A small number of studies use longitudinal data over a short-run (perhaps 1-2 years), periods in which household coping strategies may manage to mitigate impact. Studies of long-run impacts and outcomes are rare. Certainly, understanding short-run outcomes is important, but short-run effects may not ultimately translate into worse welfare outcomes in the long-run (that is, in adulthood). If the shock of an adult death (either a parent or other household member) is transitory, outcomes may be affected around the time of illness or during a period of funeral/mourning, but may recover over time. On the other hand, the lack of any short-run effects is not evidence that long-run impacts do not exist. Understanding these long-run impacts is critical for intergenerational models of the macroeconomic impact of AIDS which take into account the impact on human capital formation and its transmission between generations. Both Bell *et al.* (2003) and Corrigan *et al.* (2005) make assumptions about the magnitude of the impact of orphanhood without reference to any empirical studies of this link.

In this study, we will explore new data collected from Tanzania, which contain a sample of children surveyed in the early 1990s and then re-interviewed in 2004. Using unique features of the data, we try to assess the impact of orphanhood on schooling and health outcomes, with an emphasis on unpacking the long-run dynamics of this process. The unique features of our data set will allow us to examine the long-term, persistent impact of orphanhood shocks during childhood on adult health and educational attainment, based on a sample of children that were initially non-orphaned in our data, controlling for a wide variety of characteristics before orphanhood and allowing for migration since baseline. This work is part of a larger agenda to understand the impact of orphanhood which encompasses a range of welfare outcomes. Schooling and health are recognized as only two areas among many long-term welfare outcomes. Other relationships between orphans and long-run outcomes that will be explored using the longitudinal data include: access to networks to gain access to income (employment), inheritance of land, the role of assistance organizations, and the role of remittances from the extended family. In addition, we plan to

explore the fertility and marriage patterns of orphans and the nutritional status of the young offspring of our sample of orphans who have reached adulthood.

The next section reviews some of the existing empirical analysis. Section 3 discusses the data used in this study. Sections 4 and 5 present the empirical specifications and results. The final section concludes.

## 2. Background

Studies of the consequences of orphanhood usually focus on measuring the impact on education. The findings of the impact of orphanhood on schooling are mixed.[1] To the extent that orphans are found in relatively better-off households that also have higher demand for schooling, simple cross-sectional comparisons of enrollment rates between orphans and non-orphans may underestimate the true impact (for example, see Hargreaves and Glynn, 2002). On the other hand, in some settings, an orphan's household may be poorer prior to death, thus over-estimating the impact of orphanhood.

Several studies use large cross-sectional household survey data with controls for concurrent household characteristics to identify the impact of orphanhood on schooling. Ainsworth and Filmer (2002) find considerable diversity in the orphan/non-orphan differential across countries and conclude that generalizations about whether orphans are disadvantaged are not possible.[2] Case *et al*. (2004) use similar data but find that orphans are disadvantaged relative to other children *within* the same household. Although, if orphans are strategically placed in better-off households within the extended family, then the orphans in a household fixed-effects framework are compared to a non-random sample of non-orphan co-residents. It would be difficult then to interpret this result as evidence that orphanhood reduces schooling rather than the proposition that orphans are placed with better-off relatives.

[1] Even estimates of the *number* of orphans is debated, with some research showing that the most cited statistics from the UN may be seriously over-estimating the number of orphans (Bennell, 2005).

[2] Also using Demographic and Health Survey data, Bicego *et al.* (2003) combine household data across countries for West Africa and East Africa whereas Ainsworth and Filmer (2002) do not merge data sets across countries. In West Africa, they find evidence of a significant impact of among paternal orphans (6-14 years) and two-parent orphans (6-10 years). For East Africa, the findings are significant among maternal and two-parent orphans ages 11-14 years.

Chatterji *et al*. (2005) examine outcomes for orphans and non-orphans, including school attendance and health indicators in Rwanda and Zambia. The study finds that of the children 6-12 years old at the time of the survey, orphans have a higher probability of being in school than other children, although the difference is not statistically significant. In the age category 13-19 years old, orphans are less likely to be in school, but the difference is small in size and statistically insignificant. The authors are careful not to overinterpret these results, but it is useful to restate some of the difficulties of this analysis. In addition to concerns of omitted variables and endogeneity in a bivariate analysis based on a cross section. There could potentially be spurious positive correlation between orphan rates and school enrolment, as both are strongly positively correlated to age, especially around the ages of 6 to 12 in settings with delayed enrolment.

Evidence of other outcomes for orphans beyond schooling is much more scant but here, too, the evidence is mixed. Crampin *et al.* (2003) examine a cross-section of data from Malawi on child health linked with information on the HIV/AIDS status of parents measured 10 years prior.[3] They conclude that surviving children are not discriminated against as a result of parents having been ill or having died from HIV/AIDS. Lindblade *et al.* (2003) find that the health status of surviving orphans younger than 6 years is similar to their non-orphan counterparts in western Kenya.

Chatterji *et al.* (2005), noted above, also evaluate other human development indicators beyond schooling, including: number of meals the child had the day prior to the interview, immunization status, self-reported health status, and occurrence of illness. Depending on the country and the indicator the authors find some significant differences between orphans, children with chronically ill caregivers and other children. For example, in Zambia a lower proportion of orphans and children living with chronically ill caregivers were found to report having had three meals in the day prior to the interview, as well as having an immunization card. In both countries, significantly lower proportions of orphans and children with chronically ill caregivers were reported by their caregivers as having good or very good health status.

[3] Of the 2,250 offspring identified from the 593 individuals for whom they have HIV/AIDS status at baseline, they were able to trace 1,141. Of these, 761 were alive and interviewed, 167 were alive but left the district and not traced, and 213 were deceased. It is unclear what the implications of this sample selection are on the measured impacts.

The studies cited above use cross-sectional datasets which can only examine correlates of outcomes after a parental death, without controls for initial conditions, prior to the death of the parent(s). Depending on assumptions about the mortality patterns of adults, there could be omitted variables which bias results. In their study of children in Kenya, Evans and Miguel (2005) find that the direction of this bias is positive, meaning unobservable characteristics of orphans lead to underestimates of impact. Moreover, these studies only identify who is an orphan and not how recently a parent died or the status of the child prior to being orphaned, it is difficult to know whether the measured differentials are transitory (in which case the child may recover), permanent, or worsening as the orphan ages. This underscores the advantages of panel (longitudinal) data that follow children over long periods of time.

Bhargava (2005) analyzes a sample of orphans only and finds a positive relationship between enrollment prior to orphanhood and subsequent enrollment among orphans. This may be because of social pressure on the host family not to deprive the orphan in obvious ways. On the other hand, orphans who are fostered from a young age may be more likely to be treated more like the children of the head, if foster children when fostered at young ages display similar bonding with foster parents as biological children. Foster children whose parents are still alive may be better off than those whose parents are not, as parents may pay for schooling or reciprocate for their schooling expenditures (the child may also reciprocate later in life, but his/her parent can reciprocate more immediately). Also, if host households are typically richer than the original households, then this could explain a positive (or insignificant) effect of orphanhood on education. This suggests that evaluating the orphan effect on schooling will require moving beyond simple indicators of orphan status.

Using a 2-3 year panel survey from Tanzania (the baseline for this study, described in Section 2 below), Ainsworth *et al.* (2005) find that adult deaths are associated with delayed enrollment among younger children (7-10 years). Among orphans, younger maternal orphans are held back whereas other orphans are not found to be disadvantaged. Case and Ardington (2004) study schooling outcomes using a 2-3 year panel survey from South Africa. They find evidence of a causal effect of mother's deaths on children's education outcomes, whereas the father's death is not associated with less education. Rather,

father's deaths are found to be indicative of the socio-economic conditions of the household prior to the death, and not causal. Their test for causality is evaluated in a latter section of this paper using a much longer panel.

Ainsworth and Semali (2000) use the KHDS 1991/94 data (the baseline for this study) to present random and fixed child effects model estimates of adult death on children's height-for-age and weight-for-height. While random effects results show lower height for all maternal orphans and reduced height paternal orphans in poor households, the fixed effects results are not significant possibly due to the small number of children who have any change in their orphan status in between the survey rounds (up to about 21 months from first to last interview). They find no association between orphanhood nor recent adult death in the households and weight-for-height. Overall the authors conclude that policy interventions should target poor households in general, among which the households hardest hit by adult mortality are likely to be found. They identify population-wide policy interventions, like universal availability of ORS at health facilities, measles vaccination and improved physical access to medical care as being most appropriate.

Using unique panel data from Kenya, Evans and Miguel (2005) study a large sample of non-orphans enrolled in grades 1-7 in 1998 and reinterviewed in 2002. They evaluate the impact of orphanhood transitions on schooling participation (measured as the fraction of visits in the school year in which the child was in school on the day of an unannounced check). Maternal deaths lead to lower participation after the death but also in the 1-2 years before the death. Paternal orphans did not have lower school participation. It is unclear how the measure of school participation translates into completed school years but presumably it implies lower overall attainment.

This paper contributes to the existing set of studies by presenting unique evidence of the extent to which orphanhood matters in the long-run for health and education outcomes. The data are from a region of Northwestern Tanzania deeply affected by HIV-AIDS in Africa. We use a sample of non-orphans surveyed in 1991-94, who were traced and reinterviewed in 2004. A large proportion, 23 percent, lost one or more parents before the age of 15 in this period, allowing us to identify the impact of orphanhood shocks. Since a substantial proportion reaches adulthood by 2004, we can also assess permanent health and education

impacts of orphanhood. In the analysis, we can control for a wide range of child and adult characteristics before orphanhood, as well as community fixed effects.

## 3. Data

The Kagera Region of Tanzania is located on the western shore of Lake Victoria, bordering Uganda to the north and Rwanda and Burundi to the west. The population (1.3 million in 1988, about 2 million in 2004) is overwhelmingly rural and primarily engaged in producing bananas and coffee in the north and rain-fed annual crops (maize, sorghum, cotton) in the south. This study uses baseline data from the Kagera Health and Development Survey (KHDS), a longitudinal socioeconomic survey conducted from September 1991 to January 1994 covering the entire Kagera region in northwest Tanzania (for details, see World Bank, 2004 and http://www.worldbank.org/lsms/). Because adult mortality of the working age population (15-50) is a relatively rare event and HIV/AIDS was unevenly distributed in Kagera, the KHDS household sample was stratified based on the agro-climatic features of the region, levels of adult mortality from the 1988 Census (including both high and low mortality areas), and household-level indicators thought to be predictive of elevated adult illness or mortality, in order to capture a higher percentage of households with a death while retaining a control group of households without a death. As a result of the sampling scheme, orphan rates in the KHDS 1991-1994 are more than twice as high as would be observed in a non-stratified random sample.

In 2004, another round of data collection was completed (World Bank, 2005). The KHDS 2004 was supported by funds from DANIDA and the Knowledge for Change Trust Fund at the World Bank. The goal of the KHDS 2004 was to re-interview the sample of about 6,200 respondents from the 1991-1994 survey (which excludes 169 individuals who died over the course of the baseline rounds). While the questionnaires for the KHDS 2004 were revised to take into account the 10-year retrospective, where possible, comparability is maintained with the baseline KHDS survey instruments.

Over the course of 10-13 years, we anticipated that a significant number of individuals would have migrated from the dwelling occupied in 1991-1994. A considerable effort was made to track surviving respondents to their current location, be it in the same community (typically a village), a nearby community, within the region, or even outside the region. The success of panel surveys is often measured in terms of re-contact of households, rather than individuals therein. By this measure, excluding households in which all previous members are deceased (17 households with 31 people), the KHDS 2004 survey re-contacted 93% of the baseline households (835 out of 895 households). This is an excellent rate of re-contact compared to panel surveys in low-income countries *and* high-income countries. The KHDS panel has an attrition rate that is much lower than that of other well-known panel survey summarized in Alderman *et al*. (2001) in which the rates ranged from 17.5% attrition *per year* to the lowest rate of 1.5%. Most of these surveys in Alderman *et al* (2001) covered considerably shorter time periods (two to five years). Refusals in the KHDS were quite uncommon; the main reason for not reinterviewing surviving panel respondents was failure to locate that person.

Figure 1 shows preliminary statistics on the relocation of the 1991-1994 households. Because people have moved out of their original household, the new sample in KHDS 2004 consists of over 2,700 households from the baseline 832 which were recontacted. Much of the success in recontacting respondents was due to the effort to track people who had moved out of the baseline communities. One-half of all households interviewed were tracking cases, meaning they did not reside in the baseline communities. Of those households tracked, only 38% were located nearby the baseline community. Overall, 31% of all households were not located near the baseline communities. While tracking is costly, it is an important exercise because migration and dissolution of households are often hypothesized to be important responses to hardship, and this may be most relevant for seriously affected individuals like orphans. Excluding these households in the sample raises obvious concerns regarding the selectivity of attrition. In particular out-migration from the community (for example, due to dissolution of households or marriage), may be responses to adult mortality. The importance of following household members who move out of the original dwelling, especially when economic mobility or the assessment of changes in

welfare outcomes over time is of interest, is shown to substantially bias results (See the analysis of other panel data surveys in Beegle, 2000, and Rosenzweig, 2003).

Focusing on the sample of interest for this paper, children, Table 1 shows the rate of re-interview of children ages 0-15 at baseline (See Appendix 1 for the re-interview rates for entire sample). Among the surviving children, over 80 percent were located and re-interviewed. Of those, 42 percent were residing outside of the baseline community. Thus, without making an effort to track individuals who left the village, the re-contact rate would have fallen to 47 percent of the surviving children. Relying on the location information reported by other re-contacted household members, movement out of the region appears to be an important factor affecting inability to re-trace children. Among re-interviewed children, less than 10 percent were located outside Kagera. Among their non-traced counterparts, one-third were reported to reside outside Kagera and 9 percent of children did not have reliable location information reported from their previous household members. The sample of children who were re-interviewed, 2,499, resided in 735 households during the baseline. By 2004, these children (at least 10 years older) were residing in 1,806 households.[4]

Table 2 takes a closer look at correlates of survival and re-interview. Survivorship is correlated with several baseline characteristics of the sample of children. The youngest (under age 5) were least likely to survive as were maternal orphans. Paternal orphanhood status is not associated with higher mortality. Children in the Ngara district were significantly less likely to survive than their counterparts in Bukoba rural district. Ngara is the district which borders Rwanda and Burundi and experienced massive economic and social upheaval starting in early 1994 with the genocides in those countries and the influx of refugees

[4] It is interesting to note that the KHDS 2004 sample is remarkably similar to a random sample of households from Kagera, despite the fact that the baseline sample was not a simple random sample and given the moderate attrition since the baseline. The households in the KHDS 2004 had similar characteristics to the CWIQ survey in 2004 (including characteristics such as: household size, female headship, head's education, head's age, land holdings, livestock holdings, and dwelling characteristics). Among children under 18 years, 14% in the KHDS 2004 are single or double-parent orphans and 63% of children were living with both parents. These numbers are strikingly close to results from the DHS 2003. DHS 2003 survey found that 11% of the children under 18 years old in Kagera had lost one or both parents (a number equal to what they found in the Tanzanian national sample). In the DHS, 60% of children in Kagera who were under 18 years of age live with both parents.

and operations by donor organizations. Along with Biharamulo, Ngara was the least HIV/AIDS-affected of the six districts in the region.

Among the surviving children, the youngest (0-4 years at baseline) were most likely to be re-interviewed. Children who were living with either parent at baseline were statistically also more likely to be traced. Household wealth indicators and residence in Bukoba urban (the largest urban center and the capital of the region) were associated with lower probabilities of re-interview. This is consistent with these traits being associated with increased mobility which then results in more difficult successful tracing of the individual.

As mentioned before, in our analysis, we will focus on a sample of (panel) individuals who were 0-15 years old in the baseline survey (the first observation of the child during the KHDS1 years 1991-94). Furthermore, the sample is restricted to those who were not orphaned at that time. The full sample consists of 1,485 individuals. We do not include children who were already orphaned at baseline, since we do not have pre-orphanhood characteristics of these children. These characteristics may be correlated with current orphanhood status - the standard problem encountered by those studies only using cross-sectional data, leading to endogeneity and omitted variable bias as noted above. For both health and education outcomes, we can then assess the impact of orphanhood shocks between 1991-94 and 2004 on outcomes, for a sample of non-orphans at baseline and controlling for a wide range of child and adult characteristics. It could be argued that at younger ages, the impact of orphanhood does not need to be permanent: catch-up in health (measured via height) and in education (measured as completed years of education) after a slowdown in its accumulation is plausible. To ensure our results are reflecting *permanent* long-term losses due to orphanhood, we will also restrict our sample to those children aged at least 19 at the time of KHDS2. In this way, we are focusing on *final* adult height as well as years of education at an age where there is unlikely to be catch-up in schooling attainment. The second, alternative sample is the subset of these 1,485 panel respondents who have reached at least 19 years of age by 2004 (N=681).

Orphanhood is expected to influence health outcomes and schooling although there are multiple potential pathways of this effect. Obviously, income effects are a strong candidate, but in addition, the

conditions of the child's orphanhood play a role, such as the age at which the child is orphaned, the living conditions and arrangements prior to orphanhood, and (for schooling) the enrollment status when orphaned. In the next two sections, we will discuss the impact of orphanhood on health and education in turn. First, we examine some descriptive statistics from the two samples of children.

The advantage of panel data is the opportunity to observe the correlates of orphanhood, conditional on initial conditions. Thus, it is important to consider the number of transitions in orphan status to be able to apply this approach. In the analysis, we use the sample of non-orphaned children ages 0-15 in baseline. For this second sample, we can interpet the orphanhood impacts as permanent since the outcomes are unlikely to be altered in adulthood. The orphanhood transitions for these two samples are presented in Table 3, for parental deaths before the panel respondent reached age 15. Twenty three percent of children in the full sample experienced the loss of at least one parent between the first contact during the baseline survey and re-interview in 2004; for those that reached adulthood this percentage is about 19 percent. The most common shock was losing a father, experienced by about 18 percent in the full sample. Few lost both parents in this period.

Even when parents are alive, children do not necessarily reside with them. As is the case in other Sub-Saharan countries, the rate of fostering of children is high. In fact, fostering may be associated with orphanhood. In both rounds of the KHDS, one-parent orphans are significantly less likely to be residing with surviving parent than children with both parents alive. In baseline, more than 80 percent of non-orphans were residing with at least one parent. Among one-parent orphans, however, 62-67 percent were residing with the surviving parent. For the 2004 round, when children are between 11-28 years, fewer non-orphans are living with at least one parent (57 percent), but they are still more likely to be living with a parent than one-parent orphans. Among one-parent orphans, paternal orphans are more likely to live with their mother than maternal orphans are to live with their father. For children who became orphaned between survey rounds, we are not necessarily able to identify the living arrangement immediately following the parent's death. This would be known only in cases when orphanhood occurs close to the 2004

survey round. For others, those orphaned in the mid 1990s, we know their living arrangements prior to orphanhood and subsequent arrangements by 2004.

Before modeling the link between human development indicators and orphanhood, we start by looking at the household characteristics and socio-economic status during the baseline survey of those who became an orphan relative to their non-orphan counterparts, with simple significance tests of the differences of the means. In Table 4, we find that children experiencing the death of one or more parents during the survey period are on average younger than those who did not and are less likely to live with either their mother or father at baseline. The head of the household in which they were then living was older, less educated, and more likely to be female headed. However, they do not appear to be wealthier in terms of household consumption per capita or quality of flooring in the dwelling. The height of their mothers is also not significantly different. This would suggest that socio-economic differences determining selection into orphanhood are possibly not as strong as in some other studies.[5] Nevertheless, these descriptive statistics suggest the need to carefully control for baseline characteristics of children and parents for inference about the impact of subsequent orphanhood. Finally, orphans are more likely to be found outside the community, illustrating the importance of our tracking exercise. Its implications for inference about the impact of orphanhood will be examined below further.

## 4. Health and Education Investments

The outcomes used in this study relate to investments which affect long-term economic prospects of children, health and education. Health investments are often proxied for using anthropometric measures (such as height and weight). We will focus on height, as the key long-term health measure, since it has been shown to affect wage-earning capacity as well as participation in the labor force for men and women (see: Haddad and Bouis, 1991; Thomas and Strauss, 1997). We model height in a multivariate context as influenced by several factors, including the characteristics of the child (sex and age, via a full set of age

dummies), genetic background of the family (using height at baseline for both the child and the mother[6]) as well as the socio-economic environment in the household at baseline which is assumed to influence human development investment (including whether the child was living with the mother, whether the child was living with the father, years of education of the head, sex and age of the head of the household and two indicators of wealth: cemented floor in dwelling and log per capita household consumption). By controlling for child height at baseline, which reflects investment in child health before becoming orphaned, we further isolate the effect of orphanhood on height, net of its correlation with unobserved background characteristics influencing health.

Educational attainment can be used as a proxy for human capital levels. We focus on the educational attainment in terms of the years of education completed, counting each grade completed as a year. We use a similar specification as above, whereby years of education in 2004 are determined in a multivariate context as influenced by several factors, including child characteristics, years of education completed, whether at school at baseline, and the same household characteristics as above. As a further check on the interpretation of the results, we also include interaction terms whether the child was already enrolled when it became a paternal and maternal orphan. By controlling for years of education completed and whether at school at baseline, we further isolate the effect of orphanhood on education, net of its correlation with other unobserved background characteristics of these families influencing education. Finally, all our results are controlling for community fixed effects at baseline, isolating further the impact of orphanhood on health and education from factors such as access to schools and health services.

In this section, we present the results for the two samples of children non-orphaned at baseline, the first including all of these children 0-15 in baseline and the second restricting the full sample to those aged at least 19 in 2004, in order to ensure that the results reflect final height attainment, and (plausibly) permanent educational effects. By focusing on the sample of non-orphans in the baseline, we can focus on

[5] In the panel data examined by Case and Ardington (2003), they found that paternal orphans had lower socio-economic status. Their subsequent work concludes that the death of the father does not cause lower socio-economic outcomes but that poverty may have contributed to these deaths.
[6] A dummy is included for those mothers for whom height at baseline is unavailable.

the impact of the subsequent *transition* to maternal or paternal orphan status. The combined ability to look at final (permanent) impacts of orphanhood for a fully tracked sample of children, and the ability to control for a wide variety of characteristics before orphanhood makes these results exceptional. The orphanhood variable is defined as losing a parent before the age of 16. We allow for separate effects for losing a father and a mother, as well as the possible additive effect of losing both parents before the age of 16. Below we explore further whether the age at which one becomes an orphan matters.

Table 5 first presents the descriptive statistics for the two samples, distinguishing those not experiencing an orphanhood shock compared to those experiencing losing at least one parent before the age of 15 on the outcome variables of interest. We find significant differences in 2004 in both educational and health attainment, with orphans having significantly lower levels. However, we also find that at the time of the baseline survey (before orphanhood), subsequent orphans had lower significantly education and height. Another feature is relevant as well. The standard error of years of schooling of the two samples by 2004 is similar, around 3, irrespective of the transition in orphan status, but the standard error of the natural logarithm of height is substantially lower for the sample of those having reached adulthood. This could be expected, partly because the larger sample includes children still naturally growing (so that a wider spread of ages are present) but also because of natural processes of regression to the mean so that by adulthood height variation in any population is relatively smaller than during child and teenage growth. Indeed, our data confirm this when considering standard errors by age group – with the standard deviation remaining relatively stable around 0.5 between the ages of 12 and 17, and then becoming near 0.05 beyond this. This has implications for the interpretation of our findings

Table 6 gives the basic regressions, with full controls for baseline and child characteristics. To make the results between health and education fully comparable, we use effectively identical specification and samples for both outcomes.

For both height and educational attainment, the regressions suggest a significant effect of losing one's mother across all samples. It is striking that the effects are strong for maternal orphans, despite controlling for baseline characteristics (including initial height and education, and a full set of age controls),

which were all significantly different between orphans and non-orphans. Relative to the mean levels of non-orphans and in the full sample, maternal orphans have 11 percent lower height (or 16 cm), equivalent to 29 percent of one standard deviation of ln height in this sample. They have 0.4 school years (or 8 percent) lower years of education completed, equivalent to about 13 percent of the standard deviation of years of schooling.

These figures may potentially be misleading, since some growth catch-up is possible after a period of stunting, while children may only be temporarily affected in schooling terms. Our adult sample confirms permanent effects. The coefficient of the height regression suggests a permanent impact of about 1 percent (or 2 cm), but this still represents 22 percent of one standard deviation in this sample. In other words, for all children in the sample, we find substantial regression to the mean but only partial catch-up of the maternal orphans, leaving the larger part of the orphanhood impact persistent.[7] The 'final' educational loss is nevertheless even higher: maternal orphanhood is not just delaying education, but reducing it by almost a year (0.8 years) by adulthood. For those affected by maternal orphanhood before the age of 15, this represents 25 percent of one standard deviation of years of education in the sample. Paternal orphanhood is not associated with height and only in the full sample associated with less schooling of about 0.3 years, or 12 percent of one standard deviation of years of education. The fact that the effect is not significant in the adult sample suggest that this is may not be a persistent effect on average.

These regressions only control for the gender of the child using a dummy. Exploring further whether the impact of orphanhood had any gender dimension showed nevertheless that the impact was never significantly sex-dependent in any of the samples used for health and education. The potential additive effect of losing both parents was also tested but never to be found significant. It also did not affect the size of the effects of orphanhood episodes so it is not considered further. Finally, we explored whether the timing of orphanhood mattered, by splitting the orphanhood variables in whether the particular maternal

[7] The result that maternal orphanhood has initially large height effects during childhood, is confirmed by running the same regression as in Table 6 on the full sample but excluding those age 19+ in 2004 (the restricted sample). The effect is 14 percent lower height (or about 20 cm relative to the mean), which is equivalent of about 30 percent of one

or paternal death occurred when the child was below 12 or above 12 years of age.[8] We find no significant differences between these age categories for paternal death and for maternal death in the full sample. We find that for maternal orphanhood shocks, the impact is significantly different for both age groups in the height regressions, with the entire effect taking place before the age of 12 and no effect subsequently of an orphan shock occurring after the age of 12. Given the general crucial role of mothers at younger age, this is consistent with the larger impact of maternal death, resulting in stunting at relatively younger age, from which there is only partial recovery. In the next section, we explore the evidence further and will try to unpack these findings by investigating sources of heterogeneity in the impact of orphanhood.

## 5. Interpreting the Findings

One of the key problems in assessing the impact of shocks such as orphanhood is that unobservables correlated with orphanhood may bias the results. Finding that orphanhood is significant may not actually determine a causal impact, but rather may rather capture effects of unobservable covariates. For example, if particular attitudes towards health may lead to parental illness and death as well as to poor health outcomes of the child, then orphanhood in itself is not the cause of lower height attainment in the child. An alternative narrative could also be that it is the illness of the parent that really causes the health or educational deprivation of the child, and not the actual death of the parent.

One avenue of investigating this is to classify the sample of non-orphans used in the previous section by their *future* orphan status. That is, we know in later survey rounds which non-orphans in the sample *became* orphans in the second round (2004). We would expect that future orphan status would not be correlated with current outcomes unless (i) orphanhood captures some unobservable characteristics for which we are unable to control or (ii) morbidities associated with deaths affect outcomes before becoming an orphan. This approach follows Case and Ardington (2004) who make a causal interpretation of maternal orphanhood in a panel (2001 to 2003/04) by showing no effect of future orphan status on current schooling.

---

standard deviation in that sample (with larger standard deviation of ln height relative to the small sample, as in a healthy population).

Likewise, Evans and Miguel (2005) make note of the same issue in the context of censoring of the last round of their panel data since they include post-orphan variables which by definition are unknown for non-orphaned children in the final survey round.

Table 7 presents the regressions results of the correlates of height and years of education completed at the time of the baseline survey for the full and the adult panel sample. The child and baseline control variables are identical to Table 6, including a control for community fixed effects. We find that future orphanhood is not correlated with baseline height or education, suggesting that unobservables correlated with health and education are not correlated with becoming an orphan in the future.

The specifications in Table 6 treat the orphanhood impact as a homogenous impact on all children, such that these results are simple average effects across different types of children. By including a number of interaction effects with orphanhood, we can unpack these effects rather than treat the impact as homogenous. We focus on four sets of results: the role of wealth variables, living arrangements, migration status and, for schooling, educational status at the time of orphanhood. For educational status, we simply interact the orphan effect with a variable describing whether the child was already enrolled at the time of the parental death. This is the only variable available at the time of parental death which we could construct from the education histories in the KHDS 2004 data. For wealth variables and living arrangements, we do not have information at the time of orphanhood and instead use characteristics at the baseline survey as interactions. In particular, we define a dummy to signify the wealthiest quartile of households in terms of household consumption per capita in the baseline survey. For living arrangements, we interact the orphanhood shock with whether at baseline the child was living with the respective parent that subsequently died. All regressions control for the un-interacted effect as well, i.e. for wealth and living arrangements.

Table 8 gives the results for schooling, first each interaction category separately and then all four categories together. These results are very revealing. When these interactions are included, the impact of both maternal and paternal death are strongly significant, at least for particular groups of children. For

[8] Introducing a cut-off at a younger age did not significantly affect the results. The number of cases of parents dying at a younger age than 12 becomes relatively small, especially in the sample of children having reached adulthood by

schooling status, the interaction terms suggest that the results only hold if the child was not yet enrolled at the time of the parental death (columns 1 & 2). Otherwise, schooling outcomes of orphans are indistinguishable from other children. When the child was not yet enrolled at the time of the parent's death, the effects are very substantial. In this case, if a child was not enrolled at the time of maternal orphanhood, about 54 percent less schooling is completed (slightly more than 3 years, or one standard deviation) in the adult sample, and, for the full sample, 21 percent (or about 34percent of one standard deviation). Paternal death results in 22 percent less schooling in the full sample and about 33 percent in the adult sample. When all other dummies are included, the differential effect is still strongly significant even though the effect for orphans not enrolled somewhat smaller and less clearly identified for the death of a father. Clearly, these circumstances at the time of the parental death matter for children, resulting in substantial inertial effects, so that children do not necessarily lose school although a number of explanations for this effect are plausible. It could reflect wealth effects (since richer children are more likely to be enrolled without delay) or unobservable family attitudes towards education overriding the impact of the serious mortality shocks affecting the family.

Exploring this further, the shocks were also interacted with whether the household was richer at baseline – using the bottom quartile as the base group (columns 3 & 4). For children in poorer households, the maternal orphan shock leads to about 25 percent less schooling in the full sample (or 40 percent of one standard deviations) and 43 percent less schooling in the adult sample. This also suggests that at least to some extent, wealth effects may help to explain the results related to being enrolled at the time of the parent dying (but not entirely given that paternal effects are not identified using wealth variables). Once all interaction effects are jointly included, this effect disappears, possibly because even if the enrolment effect is a wealth effect, the wealth dummy is not precisely measuring the circumstances at the time of death but in the baseline period. Other definitions for the wealth dummy, such as using terciles, the median or alternative variables, such as floor material did not affect this interpretation.

---

2004, since relatively few of them would have been below 10 in 1991-94, affecting the identification of the effect.

Living arrangements (whether the child was fostered or living with the parent at the time of the baseline interview) also feature strongly. Among the sample of children reaching adulthood, those who were living with the respective parent at baseline are affected by the subsequent orphanhood shock, suggesting the importance of fostering arrangements before the death of the mother. The impact remains even if controlling for all the other interactions. As shown in Table 8 column 6, beyond other effects, those living with their mother and subsequently losing her have an extra 33 percent less schooling, relative to the mean non-orphan while those living with their father before his death have an additional 21 percent less schooling (about 42 percent of one standard deviation). Migration status is not significantly associated with differential orphanhood effects, although later in this section we will show that tracing migrants is important for assessing orphanhood impacts in general.

Table 8 Columns 9 & 10 combine the sets of interactions (excluding the migration indicator) and results show the types of orphans particularly affected: for example, maternal orphans, not in school when their mother died and growing up with their mother, have by adulthood about two-thirds less schooling linked to their becoming orphan (or more than 1.4 standard deviations), even after controlling for a wide set of household and community characteristics at baseline. Children who become paternal orphans when not in school and residing with their father at the time of the event similarly suffer, a finding robust to other the inclusion of other interactions.

Table 9 explores the issue of heterogeneity in impacts for the height regressions. Interacting with consumption appears to dilute all effects (columns 1 & 2), while in the full sample, we find that living with the mother at baseline reduces the impact of orphanhood (column 3). Since in this sample, some of these children would have been very young at baseline this may pick up some other latent effect and its longer term effect, possibly related to early childhood experience, with children not growing up at home being affected nutritionally or otherwise, rather than a specific orphanhood effect. Note that this effect disappears by adulthood in the restricted sample of respondents reaching adulthood, suggesting a transient issue for health (column 4). But the main lessons from this table appears to be that there is no self-evident way of unpacking maternal orphanhood effects observed in Table 6 in the short and long-run, although some of the

large short-run effects may be related to issues of living arrangements and conditions pre-orphanhood, implying that the long-term effects were relative smaller. Still, the result in Table 6 of a permanent loss of 2 cm due to maternal orphanhood, or 22 percent of one standard deviation, appears a robust result.

The final set of analyses focuses on the composition of the samples. As discussed in Section 3, the KHDS 2004 attempted to track *all* individuals from the baseline survey, with a high reinterview rate. That is, the KHDS 2004 data include individuals still living in the same communities as during the baseline survey (10-13 years earlier) as well as those who had migrated to new locations by 2004. The final table (Table 10) provides evidence that this effort was crucial for inference about the overall impact of orphanhood in a sample of initially non-orphaned children. We present results for the sample of respondents still residing in the same baseline community; this is the sample that would have been collected using more standard panel data collection practices (only tracing individuals who continue to reside in their original location). We find that the effect of mother's death on height would have been somewhat overestimated without the tracking exercise (column 1). More strikingly, the impact of maternal orphanhood on education would not have been picked up without inclusion of the individuals tracked outside the community of origin (column 3). Furthermore, without such tracking the impact of paternal death on education would have been overestimated (column 4).

## 6. Conclusions

This paper has provided unique and robust evidence on the long-term effects of orphanhood on children, into adulthood, based on a sample of children first interviewed in 1991-94 and reinterviewed in 2004. Focusing the analysis on a sample of initially non-orphaned children in 1991-94, this study has investigated the impact of observed orphanhood shocks on height and educational attainment in 2004, controlling for a wide range of household and child conditions before orphanhood and for community fixed effects. By further restricting the sample to those already reaching adulthood in 2004, we provide evidence on the persistent impacts, from which little or no recovery is possible.

We find strong effects of maternal and paternal orphanhood on education. Maternal orphans permanently lose on average close to one year of schooling (or 25 percent of one standard deviation of the non-orphan schooling levels), or 13 percent relative to the mean of non-orphans. By exploring the interaction of the orphanhood shock with particular characteristics, we find that these effects are far larger for some groups. For example, children not enrolled at the time of orphanhood and cared for at home with the parent subsequently dying tend to lose about two-thirds of schooling in comparison to non-orphans – or about 1.3 standard deviations. Similarly large shocks can be found for paternal orphans in these circumstances.

We also find robust evidence on the long-term impact on height of maternal orphanhood into adulthood, even though there is some evidence of partial catch-up (by about one third) of the impact during childhood. The permanent effect of maternal orphanhood appears to be in the order of magnitude of 2 cm (or 22 percent of one standard deviation). We found some evidence that the fostering arrangements at early childhood and before orphanhood mattered for the short-term impact of the maternal orphanhood.

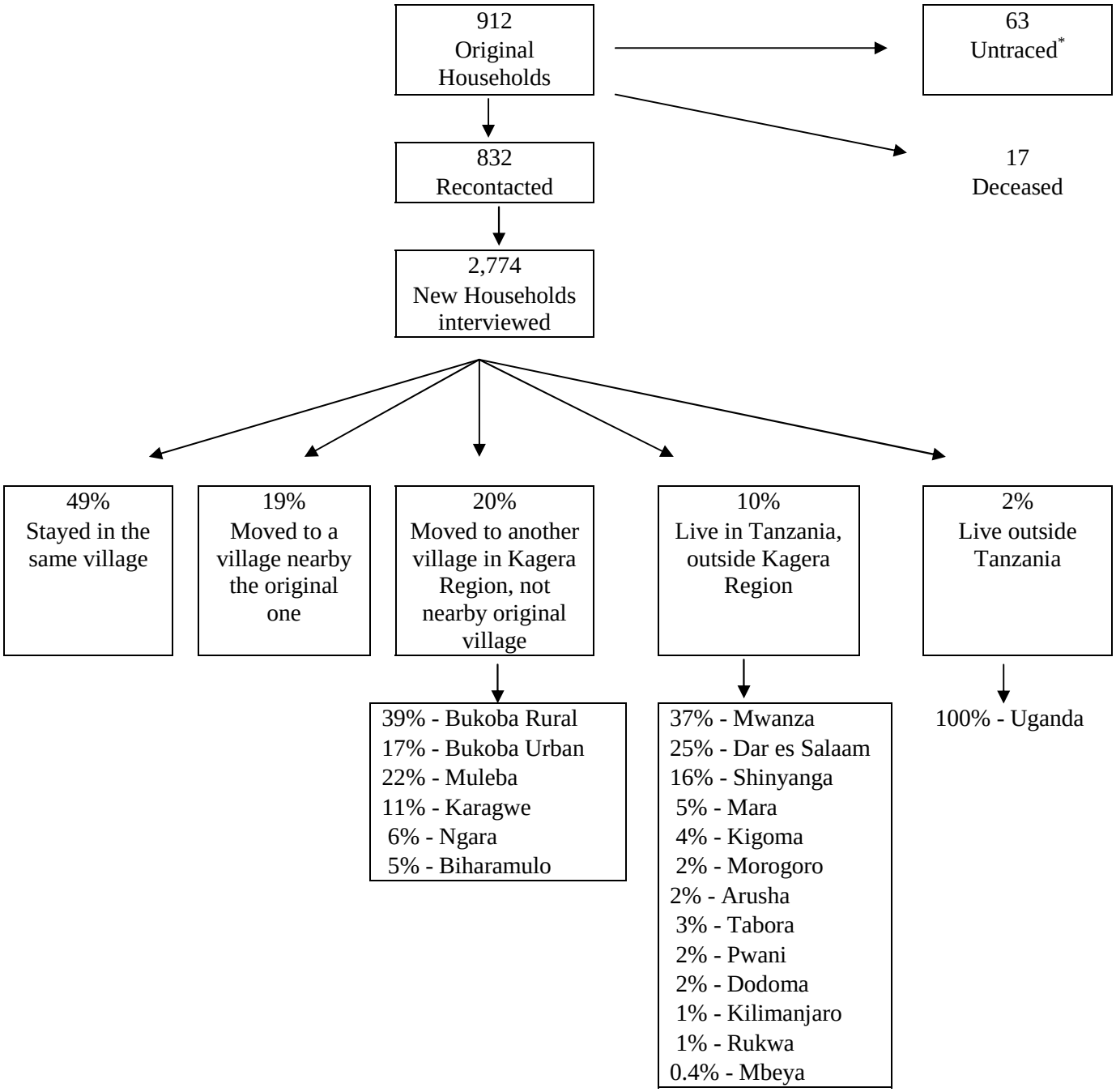
This has crucial policy implications, not least in an area where substantial interventions have taken place. It is not possible to state from this study that the impact is at all affected in positive or negative sense by existing interventions. If anything, the impact exists despite any existing interventions: the effects are still very high, not least for particular orphan groups.

Finally, the sample underlying this study is unique in that the project spent considerable resources tracing all individuals initially interviewed, even the very large number of children that had moved outside their baseline village in the 10-13 years since the previous round (about half of the sample). The paper showed some of the problematic inference that could be obtained if this effort had not been expended. For example, the impact of maternal orphanhood on health outcomes would have been much overstated, and the impact of maternal death on schooling would not have been picked up, if this tracing had not taken place.

In further work, the impact of specific programs and interventions will be explored further, as well as other long-term impacts of orphanhood and other shocks. Indeed, one question cannot be addressed by this study: is orphanhood the main source of deprivation among children or just one of many factors

relevant for targeting the poor and vulnerable? We only assessed that orphanhood mattered for these children, but further work on other issues is required to settle this other, larger question.

Figure 1: KHDS 2004, Recontacting Respondents after 10+ years




* The locations of the sample of untraced individuals were reported by informants as: Kagera (48%), Dar es Salaam (8%), Mwanza (12%), other region (10%), other country (6%) and unknown (16%).

**Table 1: KHDS Re-interview of Children**

| | Number of Children at baseline | | |
|---|---|---|---|
| Baseline sample | 3,272 | | |
| Re-interviewed | 2,499 | | |
| | (76.4%) | *1.1.1.1 Location* | |
| | | Same community | 58.2% |
| | | Nearby village | 15.8% |
| | | Elsewhere in Kagera | 16.1% |
| | | Other region | 8.2% |
| | | Other country | 1.6% |
| Untraced | 569 | | |
| | (17.4%) | *1.1.1.2 Location* | |
| | | Kagera | 54.3% |
| | | Dar es Salaam | 11.8% |
| | | Mwanza | 12.0% |
| | | Other region | 8.0% |
| | | Other country | 4.9% |
| | | Don't know | 8.8% |
| Deceased | 204 | | |
| | (6.2%) | Deceased by last round of 1991-1994 surveys | 20.0% |
| | | Deceased after last round and before KHDS 2004 | 80.0% |

*Notes:* Sample of individuals ages 0-15 in their first KHDS 1991-1994 interview. Location for children untraced is reported by other household members from the baseline survey who were successfully located and interviewed, and able to provide location information on the child.

**Table 2: Correlates of Survival and Re-interview**

| | Survival to 2004 | Re-interview in 2004 |
|---|---|---|
| Mean of dependent variables | 0.94 | 0.81 |
| *1991-1994 characteristics of child 0-15* | | |
| Male | -0.06 | -0.02 |
| | (0.80) | (0.30) |
| Age 0-4 | -0.42 | 0.33 |
| | (4.61)*** | (4.72)*** |
| Age 5-9 | -0.16 | 0.03 |
| | 1.63 | 0.43 |
| Living with mother | -0.16 | 0.14 |
| | 1.50 | (1.83)* |
| Living with father | 0.20 | 0.27 |
| | (2.04)** | (3.62)*** |
| Mother deceased | -0.25 | -0.01 |
| | (1.57) | (0.13) |
| Father deceased | 0.00 | 0.09 |
| | (0.02) | (0.97) |
| Both deceased | 0.21 | -0.04 |
| | (0.97) | (0.23) |
| Ln Per capita household expenditure/1000 | 0.04 | -0.13 |
| | 0.54 | (2.73)*** |
| Good flooring material in dwelling | 0.05 | -0.13 |
| | 0.48 | (1.79)* |
| Years of schooling of household head | 0.00 | -0.03 |
| | (0.35) | (3.20)*** |
| Bukoba Urban district | 0.13 | -0.15 |
| | (1.15) | (2.05)** |
| Ngara district | -0.34 | 0.02 |
| | (3.08)*** | (0.20) |
| Karagwe district | 0.34 | 0.29 |
| | (2.45)** | (2.92)*** |
| Biharamulo district | 0.03 | -0.27 |
| | (0.22) | (2.62)*** |
| Muleba district | -0.04 | 0.00 |
| | (0.35) | (0.01) |
| Number of observations | 3,272 | 3,068 |

Notes: Probit estimates with robust standard errors. The sample for re-interview is conditional on survival to 2004. Absolute value of z statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. All variables are binary (0/1) except real per capita expenditure (measured in Tanzania shillings) and years of schooling. Additional indicator for missing education of household head is included. Bukoba rural is the omitted district.

**Table 3: Transitions in Orphan Status**

| | Full sample | Reached adulthood |
|---|---|---|
| Both parents alive at age 15 | 1,143 | 549 |
| | (77%) | (81%) |
| Lost mother before age 15 | 143 | 50 |
| | (10%) | (7%) |
| Lost father before age 15 | 262 | 95 |
| | (18%) | (14%) |
| Double orphan by age 15 | 63 | 13 |
| | (4%) | (2%) |
| Number of observations | 1,485 | 681 |

Notes: Lost mother and lost father categories include double orphans, hence columns do not add up to 100%. "Full sample" is the sample of panel respondents not orphaned and between the ages 0-15 in the baseline. "Reached adulthood" restricts the full sample to those age 19 and older in 2004.

**Table 4: Difference in Baseline Characteristics by future orphan status**

| | Full sample | | | | | Reached adulthood | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Remain Non-Orphan to age 15 | | Loses One or Both Parents before age 15 | | sig | Remain Non-Orphan to age 15 | | Loses One or Both Parents before age 15 | | sig |
| *Baseline characteristic* | | | | | | | | | | |
| Age | 6.42 | (5.01) | 5.13 | (4.21) | *** | 10.93 | (2.82) | 9.58 | (2.48) | *** |
| Male♈ | 0.50 | (0.50) | 0.49 | (0.50) | ns | 0.49 | (0.50) | 0.43 | (0.50) | ns |
| Living with mother♈ | 0.82 | (0.38) | 0.74 | (0.44) | *** | 0.79 | (0.41) | 0.64 | (0.48) | *** |
| Living with father♈ | 0.74 | (0.44) | 0.63 | (0.48) | *** | 0.78 | (0.42) | 0.66 | (0.48) | *** |
| Years of schooling of household head | 4.60 | (2.91) | 3.94 | (2.82) | *** | 4.51 | (2.82) | 3.55 | (2.69) | *** |
| Male household head♈ | 0.87 | (0.33) | 0.78 | (0.42) | *** | 0.90 | (0.30) | 0.80 | (0.40) | *** |
| Age household head | 47.36 | (14.91) | 52.07 | (15.21) | *** | 50.27 | (13.61) | 55.37 | (14.02) | *** |
| Ln per capita consumption | 5.20 | (0.59) | 5.20 | (0.57) | ns | 5.23 | (0.63) | 5.22 | (0.59) | ns |
| Good flooring material in dwelling♈ | 0.16 | (0.36) | 0.16 | (0.37) | ns | 0.14 | (0.35) | 0.16 | (0.37) | ns |
| Height of mother (cm) | 157.60 | (5.77) | 157.90 | (5.95) | ns | 157.60 | (5.54) | 157.89 | (6.51) | ns |
| Migrated out of baseline community by 2004 | 0.38 | (0.49) | 0.44 | (0.50) | ** | 0.49 | (0.50) | 0.58 | (0.50) | ** |
| Number of observations | 1,143 | | 342 | | | 549 | | 132 | | |

Notes: "Full sample" is the sample of panel respondents not orphaned and between the ages 0-15 in the baseline. "Reached adulthood" restricts the full sample to those age 19 and older in 2004. Standard deviations are in parentheses. Column "sig" indicates the level of significance of a (one-sided) *t*-test for the difference between the two columns: * significant at 1%, ** significant at 5%, *** significant at 10%, "ns" not significant at 10%.♈ indicates binary (0/1) variables.

**Table 5: Baseline and 2004 Outcomes by Current Orphan Status**

| | Full sample (N=1,485) | | Reached adulthood (N=681) | |
|---|---|---|---|---|
| | Remain Non-Orphan till at least age 15 | Loses One or Both Parents before the age 15 | Remain Non-Orphan till at least age 15 | Loses One or Both Parents before the age 15 |
| Years of schooling | | | | |
| Baseline | 0.82 | 0.34*** | 1.70 | 0.88*** |
| | (1.66) | (1.03) | (2.06) | (1.50) |
| 2004 | 4.96 | 4.36*** | 6.16 | 5.40*** |
| | (2.93) | (2.81) | (2.98) | (3.11) |
| Ln Height (cm) | | | | |
| Baseline | 4.62 | 4.56*** | 4.89 | 4.84*** |
| | (0.32) | (0.28) | (0.12) | (0.10) |
| 2004 | 5.00 | 4.93*** | 5.09 | 5.08*** |
| | (0.33) | (0.48) | (0.05) | (0.05) |
| Number of observations | 1,143 | 342 | 549 | 142 |

Notes: Standard deviations are in parentheses. *** indicates 1% significance of a (one-sided) *t*-test for the difference between the two columns. "Full sample" is the sample of panel respondents not orphaned and between the ages 0-15 in the baseline. "Reached adulthood" restricts the full sample to those age 19 and older in 2004.

**Table 6: Determinants of Height and Years of Schooling in 2004**

| | (1) full sample Ln height | (2) reached adulthood Ln height | (3) full sample Years of schooling | (4) reached adulthood Years of schooling |
|---|---|---|---|---|
| Mother died between ages 0-15 | -0.109 | -0.011 | -0.380 | -0.757 |
| | [2.33]** | [1.87]* | [1.86]* | [2.07]** |
| Father died between ages 0-15 | -0.006 | 0.001 | -0.337 | -0.366 |
| | [0.25] | [0.12] | [2.01]** | [1.16] |
| Number of observations | 1,485 | 681 | 1,485 | 681 |

Notes: OLS estimates with community fixed effects and robust standard errors. "Full sample" (n=1706) is the sample of panel respondents not orphaned and between the ages 0-15 in the baseline. "Reached adulthood" (n=902) restricts the full sample to those age 19 and older in 2004. T-statistics in brackets. * significant at 10%, ** significant at 5%, *** significant at 1 %. Includes controls for child characteristics (sex and age dummies), baseline characteristics (residing with mother and residing with father; household consumption, flooring material, age, years of education and sex of the household head). The height regressions include the height of the child at baseline and, when available, mother's height from the baseline data. The schooling regressions include the years of education of the child and whether the child was at school in the baseline.

**Table 7: Predicting Baseline Outcomes by Future Orphan Transitions**

| | (1) full sample Ln height | (2) reached adulthood Ln height | (3) full sample Years of schooling | (4) reached adulthood Years of schooling |
|---|---|---|---|---|
| Mother died between ages 0-15 | -0.004 | -0.011 | 0.097 | 0.230 |
| | [0.49] | [1.02] | [1.19] | [1.18] |
| Father died between ages 0-15 | -0.008 | -0.003 | -0.065 | -0.094 |
| | [1.40] | [0.35] | [1.01] | [0.61] |
| Number of observations | 1,485 | 681 | 1,485 | 681 |

Notes: OLS estimates with community fixed effects and robust standard errors. T-statistics in brackets. See notes to Table 6. The height regressions include, when available, mother's height from the baseline data.

**Table 8: Determinants of Years of Schooling: interactions**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | full sample | reached adulthood | full sample | reached adulthood | full sample | reached adulthood | full sample | reached adulthood | full sample | reached adulthood |
| Mother died between ages 0-15 | -0.989 [2.90]*** | -3.269 [3.32]*** | -1.183 [2.19]** | -2.589 [2.34]** | 0.151 [0.43] | 0.453 [0.91] | -0.147 [0.49] | -0.69 [0.94] | -0.974 [1.44] | -2.559 [2.41]** |
| Father died between ages 0-15 | -1.079 [4.77]*** | -2.008 [3.61]*** | -0.592 [2.01]** | -0.567 [1.09] | 0.144 [0.53] | 0.537 [1.09] | -0.58 [2.90]*** | -0.56 [1.35] | -0.758 [2.00]** | -1.32 [1.76]* |
| *Interaction terms*: | | | | | | | | | | |
| Mother died and in school when she died | 1.256 [3.24]*** | 3.149 [3.05]*** | | | | | | | 1.091 [2.73]*** | 2.658 [2.51]** |
| Father died and in school when he died | 1.618 [5.80]*** | 2.46 [4.00]*** | | | | | | | 1.647 [5.90]*** | 2.668 [4.49]*** |
| Mother died and in top 75% of cons pc | | | 0.971 [1.67]* | 2.08 [1.78]* | | | | | 0.446 [0.82] | 0.722 [0.73] |
| Father died and richer than median | | | 0.402 [1.17] | 0.307 [0.49] | | | | | 0.336 [1.06] | 0.501 [0.90] |
| Mother died and in top 75% of cons pc | | | | | -0.723 [1.70]* | -1.982 [3.01]*** | | | -0.391 [0.96] | -1.626 [2.77]*** |
| Father died and living with father | | | | | -0.681 [2.01]** | -1.277 [2.00]** | | | -0.807 [2.50]** | -1.622 [2.68]*** |
| Moved away | | | | | | | -0.037 [0.26] | 0.355 [1.54] | | |
| Mother died and moved away | | | | | | | -0.442 [1.09] | -0.22 [0.25] | | |
| Father died and moved away | | | | | | | 0.604 [1.78]* | 0.438 [0.71] | | |
| Number of observations | 1,485 | 681 | 1,485 | 681 | 1,485 | 681 | 1,485 | 681 | 1,485 | 681 |

Notes: OLS estimates with community fixed effects and robust standard errors. T-statistics in brackets. * significant at 10%, ** significant at 5%, *** significant at 1 %. Specifications as described in Table 6, supplemented with specific interaction terms. Wealth (1=if top 75% in consumption per capita) and living with father or mother are baseline characteristics. 'Moved away' refers to 2004 residence and is one if the child is not living in the same community as in the baseline. Whether enrolled at time of death are measured relative to the year of death of the parent.

**Table 9: Determinants of Height: interactions**

| | (1) full sample | (2) reached adulthood | (3) full sample | (4) reached adulthood | (5) full sample | (7) reached adulthood | (8) full sample | (9) reached adulthood |
|---|---|---|---|---|---|---|---|---|
| Mother died between ages 0-15 | -0.205 [1.45] | -0.003 [0.17] | -0.299 [2.56]** | -0.02 [1.72]* | -0.159 [1.99]** | -0.032 [4.20]*** | -0.455 [2.34]** | -0.014 [0.73] |
| Father died between ages 0-15 | 0.025 [0.75] | -0.003 [0.38] | -0.019 [0.40] | 0.003 [0.45] | -0.027 [0.65] | -0.004 [0.72] | 0.022 [0.39] | 0 [0.02] |
| *Interaction terms*: | | | | | | | | |
| Mother died and in top 75% of cons pc | 0.12 [0.79] | -0.01 [0.58] | | | | | 0.171 [1.13] | -0.007 [0.42] |
| Father died and in top 75% of cons pc | -0.045 [1.04] | 0.005 [0.63] | | | | | -0.054 [1.23] | 0.005 [0.59] |
| Mother died and living with mother | | | 0.257 [1.97]** | 0.014 [1.06] | | | 0.282 [2.12]** | 0.013 [1.00] |
| Father died and living with father | | | 0.008 [0.14] | -0.005 [0.52] | | | 0.003 [0.05] | -0.004 [0.47] |
| Moved away | | | | | 0.025 [1.24] | -0.003 [0.89] | | |
| Mother died and moved away | | | | | 0.086 [0.88] | 0.028 [2.65]*** | | |
| Father died and moved away | | | | | 0.064 [1.25] | 0.01 [1.27] | | |
| Number of observations | 1,485 | 681 | 1,485 | 681 | 1,485 | 681 | 1,485 | 681 |

Notes: OLS estimates with community fixed effects and robust standard errors. T-statistics in brackets. * significant at 10%, ** significant at 5%, *** significant at 1 %. Specifications as described in Table 6, supplemented with specific interaction terms. Wealth (1=if top 75% in consumption per capita) and living with father or mother are baseline characteristics. 'Moved away' refers to 2004 residence and is one if the child is not living in the same community as in the baseline. Whether enrolled at time of death are measured relative to the year of death of the parent.

**Table 10: Implications of Tracking on Inference**

| | Within Baseline Community in 2004 | | | | Within + Migrated Out of Baseline Community *(Table 6)* | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) reached adulthood | (3) | (4) reached adulthood | (5) | (6) reached adulthood | (7) | (8) reached adulthood |
| | full sample Ln height | Ln height | full sample Years of schooling | Years of schooling | full sample Ln height | Ln height | full sample Years of schooling | Years of schooling |
| Mother died between ages 0-15 | -0.147 | -0.028 | -0.194 | -0.663 | -0.107 | -0.010 | -0.418 | -0.784 |
| | [1.88]* | [3.68]*** | [0.66] | [0.84] | [2.32]** | [1.75]* | [2.08]** | [2.25]** |
| Father died between ages 0-15 | -0.019 | -0.004 | -0.557 | -0.598 | -0.007 | -0.001 | -0.278 | -0.285 |
| | [0.45] | [0.50] | [2.76]*** | [1.28] | [0.26] | [0.26] | [1.66]* | [0.92] |
| Number of observations | 899 | 337 | 899 | 337 | 1,485 | 681 | 1,485 | 681 |

Notes: OLS estimates with community fixed effects and robust standard errors. "Within Baseline Community" is the subset of the Table 6 samples of panel respondents who still resided in the baseline community in 2004. See Table 6 notes. T-statistics in brackets. * significant at 10%, ** significant at 5%, *** significant at 1 %.

**Appendix Table 1: KHDS Individuals, by Age**

| Age at baseline 1991-1994 | Re-contacted | Deceased | Untraced | Reinterview among survivors |
|---|---|---|---|---|
| <10 years | 1,605 (77.1%) | 160 (7.7%) | 317 (15.2%) | 83.5% |
| 10-19 years | 1,406 (73.0%) | 104 (5.4%) | 415 (21.6%) | 77.2% |
| 20-39 years | 828 (63.4%) | 287 (22.1%) | 190 (14.6%) | 81.4% |
| 40-59 years | 434 (70.6%) | 148 (24.0%) | 34 (5.5%) | 92.3% |
| 60+ years | 163 (37.6%) | 261 (60.4%) | 9 (2.1%) | 94.2% |
| Overall | 4,435 (69.7%) | 961 (15.1%) | 965 (15.2%) | 82.1% |

*Notes:* Sample of individuals interviewed at least once in KHDS 1991-1994. Age categories are based on age at first interview. Some deceased individuals died during the KHDS 1991-1994 panel. Row percentages in parentheses.

**References**


Ainsworth, Martha, Kathleen Beegle, and Godlike Koda. 2005. "The Impact of Adult Mortality and Parental Deaths on Schooling in Northwestern Tanzania." *Journal of Development Studies* 41(3): 412-439.

Ainsworth, Martha, and Deon Filmer. 2002. "Poverty, AIDS, and Children's Schooling: A Targeting Dilemma." World Bank Policy Research Working Paper No. 2885.

Ainsworth, Martha and Innocent Semali. 1998. "Who is Most Likely to Die of AIDS? Socioeconomic Correlates of Adult Deaths in Kagera Region, Tanzania." in M. Ainsworth, L. Fransen, and M. Over, eds., *Confronting AIDS: Evidence from the Developing World*. Brussels: European Union.

Ainsworth, Martha and Innocent Semali. 2000. "The Impact of Adult Deaths on Children's Health in Northwestern Tanzania" Policy Research Working Paper 2266, Development Research Group. The World Bank: Washington DC.

Alderman, Harold, Jere R. Berman, Hans-Peter Kohler, John Maluccio, and Susan Cotts Watkins. 2001. "Attrition in Longitudinal Household Survey Data: Some Tests for Three Developing-Country Samples." *Demographic Research* 5(4).

Beegle, Kathleen. "Economic Mobility in Indonesia and Vietnam: What Missing Data Can't Tell Us." mimeo, The World Bank.

Bell, Clive, Shanta Devarajan and Hans Gersbach. 2003. "The Long-run Economic Costs of AIDS: With an Application to South Africa." World Bank Policy Research Working Paper No. 3152.

Bennell, Paul. 2005. "The Impact of the AIDS Epidemic on the Schooling of Orphans and Other Directly Affected Children in Sub-Saharan Africa." *Journal of Development Studies*, 41(3): 467-488.

Bhargava, Alok. 2005. "AIDS epidemic and the psychological well-being and school participation of Ethiopian orphans." Psychology, Health and Medicin 10(3): 263-275.

Bicego, George, Shea Rutstein, and Kiersten Johnson. 2003. "Dimensions of the Emerging Orphan Crisis in Sub-Saharan Africa." *Social Science and Medicine* 56: 1235-1247.

Case, Anne and Cally Ardington. 2004. "The Impact of Parental Death on School Enrollment and Achievement: Longitudinal Evidence from South Africa." mimeo.

Case, Anne, Christina Paxson and Joseph Ableidinger. 2004. "Orphans in Africa: Parental Death, Poverty and School Enrollment," *Demography* 41(3): 483-508.

Chatterji, Minki, Leanne Dougherty, Tom Ventimiglia, Yvonne Mulenga, Andrew Jones, Antoinette Mukaneza, Nancy Murray, Kathy Buek, William Winfrey, and Joseph Amon. 2005. "The Well-Being of Children Affected by HIV/AIDS in Lusaka, Zambia, and Gitarama Province, Rwanda: Findings from a Study". Community REACH Working Paper No. 2. Washington, DC: Community REACH Program, Pact.

Corrigan, Paul, Gerhard Glomm, and Fabio Mendez. 2005. "AIDS Crisis and Growth." *Journal of Development Economics* 77(1): 107-124

Crampin Amelia, Sian Floyd, Judith Glynn, Nyovani Madise, Andrew Nyondo, Masiya Khondowe, Chance Njoka, Huxley Kanyongoloka, Bagrey Ngwira, Basia Zaba, and Paul Fine. 2003. "The Long-term Impact of HIV and Orphanhood on the Mortality and Physical Well-being of Children in Rural Malawi." *AIDS* 17(3):389-97.

Evans, David and Edward Miguel. 2005. "Orphans and Schooling in Africa: A Longitudinal Analysis." mimeo, University of California.

Haddad, Lawrence. and Howard . Bouis. 1991. "The Impact of Nutritional Status on Agricultural Productivity: Wage Evidence from the Philippines." *Oxford Bulletin of Economics and Statistics* 53:45-68.

Hargreaves, James R., and Judith R. Glynn. 2002. "Educational Attainment and HIV-1 Infection in Developing Countries: A Systematic Review." *Tropical Medicine and International Health* 7(6): 489-498.

Lindblade, Kim, Frank Odhiambo, Daniel Rosen, Kevin DeCock. 2003. "Health and Nutritional Status of Orphans <6 Years Old Cared for by Relatives in Western Kenya." *Tropical Medicine and International Health* 8(1):67-72.

Rosenzweig, Mark. 2003. "Payoffs from Panels in Low-Income Countries: Economic Development and Economic Mobility." BREAD Working Paper No. 011.

Tanzania Commission for AIDS (TACAIDS), National Bureau of Statistics (NBS) and ORC Macro. 2005. "Tanzania HIV/AIDS Indicator Survey 2003-04. Calverton, Maryland, USA: TACAIDS, NBS and ORC Macro.

Thomas, Duncan and John Strauss. 1997. "Health and Wages: Evidence on Men and Women in Urban Brazil." *Journal of Econometrics* 77-159-185.

World Bank. 2004. "User's Guide to the Kagera Health and Development Survey Datasets." mimeo.

World Bank. 2005. "Kagera Health and Development Survey 2004 Basic Information Document." mimeo.