ORPHANS AND SCHOOLING IN AFRICA: A LONGITUDINAL ANALYSIS*

DAVID K. EVANS AND EDWARD MIGUEL

AIDS deaths could have a major impact on economic development by affecting the human capital accumulation of the next generation. We estimate the impact of parent death on primary school participation using an unusual five-year panel data set of over 20,000 Kenyan children. There is a substantial decrease in school participation following a parent death and a smaller drop before the death (presumably due to pre-death morbidity). Estimated impacts are smaller in specifications without individual fixed effects, suggesting that estimates based on cross-sectional data are biased toward zero. Effects are largest for children whose mothers died and, in a novel finding, for those with low baseline academic performance.

More than one in nine children under age 18 in sub-Saharan Africa have lost a parent, and the HIV/AIDS pandemic is the leading cause (UNAIDS, UNICEF, and USAID 2004).[1] HIV/AIDS deaths today could have major long-term effects on economic development by affecting the human capital accumulation of the next generation. While some have argued that HIV/AIDS is the key development issue facing Africa (UNAIDS 2000), and children orphaned by AIDS have received considerable international media coverage,[2] surprisingly little systematic empirical research has estimated the impact of parent death on children's education. In the absence of conclusive evidence, a range of views persists regarding the likely impacts.

All sensible observers agree that parent death has an adverse effect on surviving children, but a more complete understanding of these effects—including effects for households and communities with particular characteristics—is critical for the design of programs to successfully assist orphans. Although there are many possible explanations for negative effects on schooling, including lower household income after the parent death, these effects could be mitigated by strong traditional child-fostering norms in Africa. For instance, in the 51 available Demographic and Health Surveys administered in

*David K. Evans, RAND Corporation. Edward Miguel, University of California, Berkeley, and National Bureau of Economic Research. Address correspondence to Edward Miguel, 549 Evans Hall #3880, Department of Economics, University of California, Berkeley, CA 94720-3880; E-mail: emiguel@econ.berkeley.edu. We are grateful for financial support from the World Bank, the Partnership for Child Development, the Harvard Center for International Development, the Harvard University Inequality Program, and the National Institutes of Health Fogarty International Center (R01 TW05612-02). We thank ICS Africa and the Kenya Ministry of Health Division of Vector Borne Diseases for their cooperation in all stages of the project, and we would especially like to acknowledge the contributions of Elizabeth Beasley, Laban Benaya, Pascaline Dupas, Simon Brooker, Alfred Luoba, Sylvie Moulin, Robert Namunyu, Polycarp Waswa, and the PSDP field staff and data group, without whom the project would not have been possible. Harold Alderman, Josh Angrist, Gustavo Bobonis, Anne Case, Caroline Hoxby, Larry Katz, Michael Kremer, participants in seminars at Harvard, the World Bank, 2004 NBER Summer Institute, USC, Oxford, 2005 IUSSP seminar on "Interactions Between Poverty and HIV/AIDS," and University of California, Berkeley, have provided valuable comments, as have the editors of this journal and two anonymous referees. Melissa Gonzalez-Brenes and Tina Green have provided excellent research assistance. This is a substantially revised version of BREAD Working Paper No. 56 (see Evans and Miguel 2004). All errors are our own.

1. In most of the literature on orphanhood in sub-Saharan Africa, a child is referred to as an orphan if her mother has died, if her father has died, or both. In this article, we use the female pronoun when referring to orphans of both genders.

2. Recent popular media articles claim that "as the HIV epidemic deepens in Africa, it is leaving an economically devastated continent in its wake" (Wehrwein 2000). For another media example among many, see Robinson (1999). Young (2005) presents a theoretical case for why the HIV/AIDS epidemic could actually lead to faster African economic growth.

sub-Saharan Africa between 1990 and 2004, the average proportion of households with at least one foster child was greater than one-fifth (ORC Macro 2005).

Several recent studies have examined the issue of parent death and child schooling using a variety of methods and data sources, yielding quite different results. Case, Paxson, and Ableidinger (2004) employed 19 Demographic and Health Surveys (DHS) collected across 10 sub-Saharan African countries between 1992 and 2000 to estimate the impact of parent death on school enrollment. They used a household fixed-effects estimation strategy, which compares orphans and non-orphans in the households that take in orphans. Their main finding was that orphans are significantly less likely to be enrolled in school than non-orphans, even within the same household. The study may suffer from omitted variable bias: only cross-sections were available in their data set, and thus they could not account for fixed characteristics of the orphan child or her original household (which were not observed in the data).[3]

In contrast, earlier studies did not find a substantial negative impact of parent death on child education. For instance, Ainsworth, Beegle, and Koda (2005) analyzed a panel of 1,213 Tanzanian children and found minimal impacts of parent death on schooling. Although these researchers controlled for baseline household characteristics, they did not use child fixed effects and did not fully control for child age in the analysis. Several studies have echoed Ainsworth et al. (2005) in finding little or no difference between orphans' and non-orphans' school enrollment (Kamali et al. 1996; Lloyd and Blanc 1996; Ryder et al. 1994), although these relied on less-conclusive cross-sectional methods. A number of international organization reports have claimed, however, that there are substantial gender differences in parent death impacts on schooling, with girls suffering more than boys (UNAIDS 2002; World Bank 2002).

The absence of consistent negative impacts of parent death on African children in existing work has sometimes been attributed to the strength of extended family and community networks that care for orphans (Foster et al. 1995; Foster and Williamson 2000; Ntozi 1997). An alternative explanation for the small estimated orphan effects in cross-sectional studies is the possibility that African HIV/AIDS victims are often of higher socioeconomic status than nonvictims. This will be the case if individuals in occupations particularly vulnerable to early infection—including truckers, soldiers, and teachers—tend to be relatively affluent. This positive correlation between socioeconomic measures and HIV prevalence has been found in several African studies (e.g., Ainsworth and Semali 1998) and also holds in the 2003 Kenya DHS data (Central Bureau of Statistics and Ministry of Health [Kenya] 2004:223). To the extent that socioeconomic variation is at least partially unobserved by the econometrician, this leads to a bias toward zero in the estimated "impact" of being an orphan on subsequent life outcomes in cross-sectional studies, obscuring negative impacts of parent death.

This issue is less of a concern in longitudinal studies in which fixed differences across households can be controlled for in the analysis. Yamano and Jayne (2005) used a difference-in-differences identification strategy with a panel data set of Kenyan households and found significant negative impacts of adult death on school enrollment, but only among poor children. They estimated the impact of the death of any adult in the child's household because they lacked parent death information. Recent research using longitudinal data from South Africa has found strong negative effects on schooling of the death of a child's mother but not of a child's father (Case and Ardington 2006).

3. The DHS household asset information is collected contemporaneously with the measurement of orphan status and thus is potentially endogenous: households fostering orphans may choose to sell assets, becoming poorer. It is preferable to measure characteristics prior to the parent death.

PRIMARY SCHOOLING AND ORPHANS IN KENYA

The primary school finance context in rural Kenya during 1998–2002 is important in understanding households' decisions about school participation.[4] The national Ministry of Education paid teacher salaries, while local school committees raised funds locally from parents for books, classrooms, and desks. These annual school fees were set by the local school committee and collected by the headmaster. School fees ranged roughly from $4 USD to $10 USD per family during 1998–2002, a nontrivial amount in this area. Children whose parents failed to pay school fees could be temporarily suspended from school (Miguel and Gugerty 2005). However, presidential decrees during the study period prevented schools from permanently expelling students for failure to pay.

Few primary schools in this area made special allowances for orphans in terms of school fee reductions, according to interviews with 48 primary school headmasters conducted in our study area in Western Kenya in 2002. Forty-two of the 48 (88%) headmasters stated that orphans were subject to exactly the same fees as other children. Of the 38 headmasters admitting they had sent some students away temporarily for nonpayment of school fees in the previous year, 32 (84%) claimed that orphans had been sent away just as often as non-orphans. Thus, the inability to pay school fees is a plausible cause for at least part of the drop in school participation after a parent death, to the extent that the death reduces household income (as Yamano and Jayne 2004 found) and in the presence of well-known credit constraints.

Yet there remain many possible channels linking parent death to schooling other than income, including changes in the quality of emotional support from a parent (or other caregiver), psychological trauma resulting from the death, and disruptions caused by fostering. We provide suggestive evidence below that factors other than income play an important role.

DATA AND MEASUREMENT

The data were collected in Busia district, Kenya, a densely settled farming region adjacent to Lake Victoria, in the context of a primary school health program that provided medical treatment for intestinal worm infections (Miguel and Kremer 2004). The Kenyan nongovernmental organization (NGO) ICS Africa began carrying out that program in late 1997, and the 75 schools taking part consist of nearly all rural primary schools in Budalangi division and Funyula division in Busia.[5]

The first data set, the 1998 Pupil Questionnaire, was administered from January to March 1998 and collected information from children on a variety of health measures and household socioeconomic characteristics, providing valuable baseline (pre–parent death) controls for a subset of children initially in Grades 3 through 7. We also have baseline 1998 academic test scores for a slightly smaller subset of children in those same grades.

We also use data on school participation over five school years, from early 1998 to mid-2002. Schools were visited by enumerators four to five times per year to record student school attendance and enrollment, and these visits were not announced to the school in advance. For children in preschool through Grade 8, school participants are defined here as those children present in school on the day of an unannounced check, while absent children and dropouts are considered nonparticipants. Attempts were made to track children who

4. This section describes Kenyan primary school finance before Mwai Kibaki was elected president in December 2002. In early 2003, the Ministry of Education abolished local school fees nationwide and provided some additional resources to compensate for lost local funds.

5. The 75 schools included 84% of the local primary school pupil population according to Busia District Education Office records. Several of the remaining schools were either expensive private schools for the local elite or geographically isolated schools located on islands in Lake Victoria, schools to which students in our sample were unlikely to transfer.

transferred to other schools within Busia and neighboring Teso district, but data were not collected for those who left these two districts.

The third data set we use is the 2002 Tracking Survey. The original 75 schools were visited by enumerators between February and August 2002 in order to track each child from the baseline sample. If the child was present at school, she was asked directly about the mortality of her parents and the year of the death if a parent had died. If the child was not present that day, teachers and other students were asked to provide this information. In practice, it was common for siblings, cousins, and neighbors of absent children to volunteer parent death information, which appears to be quite widely known in rural Kenyan communities.

As a check on data reliability, the parent mortality data collected at school were compared with mortality data collected at children's homes in 2001 for a representative subsample of 72 children (among those who had experienced a parent death and for whom we already had home contact information). These home surveys were typically collected from intimate relatives of the dead parent. There is a moderate correlation in the reported year of father death between the two surveys (0.87), but the correlation for mother deaths is considerably lower (0.61). There are many fewer mother deaths (24 deaths) than father deaths (48 deaths) in this sample, so this latter figure is based on relatively few observations. The reported year is identical in the two surveys (at school and at home) for 71% of father deaths and differs by more than one year in only one case for either the father or mother deaths.[6]

Sample Size and Attrition

The scarcity of African panel data sets is due in part to difficulties in tracking respondents through time, and we are not immune to this problem. Migration, fostering, and imperfect recall all complicate our task and lead to nontrivial rates of missing data, especially on the year of parent death. We conduct simulations (described later) to place bounds on the extent of bias due to missing data.

The baseline sample includes all 24,111 children who were not orphans at baseline in early 1998, were enrolled in the 75 NGO program schools in Grades 1 through 7, and were between 5 and 18 years old. Preschool and Grade 8 children were excluded due to the low tracking rates for those two groups. We use two samples of children in the main analysis, the *full sample* and the *restricted sample.* The full sample of 18,133 children includes all baseline students for whom there is mortality data for both parents.[7] Most cases of unknown orphan status were among children initially in the upper grades in 1998. They had been out of primary school longer than younger pupils and were often not as well-known to the other schoolchildren during the tracking survey.

The restricted sample contains 7,815 children from the full sample for whom 1998 Pupil Questionnaire data are available. The restricted sample first drops all 6,718 students (of the 18,133 students in the full sample) initially enrolled in Grades 1 and 2 in 1998 because the 1998 Pupil Questionnaire was administered only in Grades 3 through 7. Of the remaining 11,415 students, there is survey information for only a subset of students—those present on the day of that survey—leaving 7,815 children in the restricted

6. We find similarly high reliability in the reporting of orphan status, rather than simply the year of parent death. Among the 161 home-based survey reports of father death (for all years, including those before 1998), 128, or 80%, were also reported in the school-based tracking exercise. Similarly, among the 74 home-based survey reports of mother death, 53, or 72%, were also reported in the school survey.

7. Age data are also missing for 3,163 children in the full sample. It is not uncommon for individuals in rural Kenya not to know their birth year, and formal birth certificates are rare. However, we include these observations in the analysis using indicator variable controls for observations with missing values. Individuals are excluded from the sample when they reach age 18 due to the difficulty in collecting reliable schooling information for them past that age.

sample. For 2,194 of the restricted sample students, school participation data are available for only a subset of the five years we study because in some cases, students moved away from the area and there is no information regarding subsequent schooling. For both the full and restricted sample, such children are included in the analysis only in the years in which we observe them.[8]

Child characteristics are similar for both the full and restricted samples (Appendix Table A1). A relatively large proportion (8%) of children who were non-orphans in 1998 became orphans during 1999–2002. Fully 15% of schoolchildren were orphans at baseline in 1998, and this proportion varies widely across the 75 schools, from near zero in some areas up to 41% in others. Children in this region are quite poor even by Kenyan standards. For instance, only 14% of children wore shoes to school. This translates into poor health and nutrition (Appendix Table A1, Panel B): nearly 20% of households lacked a latrine (or toilet) at home, and almost two-fifths reported experiencing a fever in the month preceding the survey.[9] The average weight-for-age *z* score is –1.44, which is similar to the overall average for Kenyan children in this age group (UNDP 2002).

EMPIRICAL STRATEGY

Estimation Approach

Most existing studies on the impact of parent death have estimated differences between orphans and non-orphans at a single point in time, controlling for a limited set of current observable child characteristics. The results of such studies may be misleading due to both omitted variables and endogeneity: in the absence of longitudinal data, it is impossible to know whether these orphans and non-orphans were comparable before the parent death, and more important, the current child and household characteristics used as controls may have themselves been affected by the death. Moreover, because parent death is relatively rare in most populations, few studies have sufficient statistical precision to reliably estimate moderate impacts.

In an attempt to address these concerns, we compare changes in the school participation of children whose parents died during the period 1999–2002 to changes for children whose parents did not die. Annual school participation, or the fraction of unannounced enumerator visits for which the child was present at school, takes on a range of values between 0 and 1. The main estimation approach in this article is linear regression with child fixed effects, where the "events" of interest are parent deaths. The fixed effect captures time-invariant child characteristics that affect school participation. In some specifications, we examine effects on school enrollment—an indicator variable for students present at school during at least one enumerator visit over the course of the year—as an alternative outcome.

To the extent that the unobserved differences between children who become orphans and those who do not are time-invariant, Eq. (1) yields unbiased estimates of the effect of parent death on child schooling. (We discuss this assumption further later.) In some specifications, baseline characteristics, rather than the child fixed effects, are included as explanatory variables. Disturbance terms are allowed to be correlated within schools (the "1" subscript here refers to the equation number):

$$Y_{ijt} = \alpha_{1ij} + \rho_{1jt} + \sum_{\tau} \beta_1^S \cdot 1(\tau = S)_{ijt} + \sum_{c} \gamma_1^C \cdot 1(c = C)_{ijt} + \delta_1 T_{jt} + u_{1j} + e_{1ijt}. \tag{1}$$

8. Later, we also extend the analysis to those who were already orphans at baseline ("always orphans"). The 2,676 "always orphans" are selected using the same criteria as the full sample.

9. This includes disease episodes classified by children as either "fever" or "malaria." Although children were asked to report recent cases of malaria (as opposed to fever), fevers of any cause are often reported as "malaria" in areas like Busia, where testing is costly (Watson 1992).

Y_{ijt} is the school participation rate for student i in school j during year t, α_{ij} is the student fixed effect, and ρ_{ij} is a region-year indicator variable (at the level of the administrative division). The school participation of children could simply be compared before and after a parent death to arrive at an estimated parent death effect, but such a specification imposes a constant effect of parent death on subsequent child outcomes regardless of when the parent died. It is theoretically possible that the effects of parent death might either compound over time (if an initial adverse shock increases the probability of negative outcomes in successive periods) or diminish (if coping mechanisms emerge over time). To allow for such effects, we include indicator variables $\sum_{\tau}\beta^{S}\cdot 1(\tau = S)_{ijt}$ in some specifications, where τ is the number of years since the parent death, and S is a value that τ can take on (i.e., the number of years since the death); τ also takes on negative values before the death, for instance, due to AIDS-related morbidity.

Medical researchers estimate that AIDS deaths in nearby Uganda are typically preceded by 4 to 17 months of AIDS-related illness (Morgan et al. 2000; Morgan and Whitworth 2001), and thus we might observe negative effects up to two calendar years before the parent death. In practice, we include indicators for each year from three years before the parent death to three years after the death (where the omitted category is four or more years before a death). We do not observe children four or more years after a parent death because the main analysis is restricted to children who were non-orphans at baseline in 1998, and we observe them only through 2002.

In order to account for cohort and year-specific trends as well as gender differences in school participation, which are independent of parent death, a full set of age cohort–year–gender indicator variables (where c denotes a particular age cohort–gender group in a particular year, such as girls born in 1986, observed in 1999) are always included, $\sum_{c}\gamma^{C}\cdot 1(c = C)_{ijt}$. We also include an indicator variable for medical treatment through the school-based deworming program in school j in year t, T_{jt}, which boosted school participation (Miguel and Kremer 2004).

The preferred specification is more parsimonious, including the two mutually exclusive terms: (1) $ORPHAN_{ijt}$, which takes a value of 1 if the individual is an orphan in period t (in other words, for all years during and following the parent death) and 0 otherwise, and (2) PRE_{jt}, which takes a value of 1 during the two years before an individual becomes an orphan:

$$Y_{ijt} = \alpha_{2ij} + \rho_{2jt} + \beta_2^{PRE}PRE + \beta_2^{POST}ORPHAN + \sum_{c}\gamma_2^{C}\cdot 1(c = C)_{ijt} + \delta_2 T_{jt} + u_{2j} + e_{2ijt}. \quad (2)$$

Parent death may have differential effects as a function of parent gender—for instance, to the extent that mothers' income and caregiving are more (or less) important than fathers' income and caregiving, or if maternal deaths have different implications for subsequent fostering patterns. To estimate differential effects, we include separate indicators for maternal and paternal deaths and also estimate the impact of the first parent death versus the second parent death. The parent death indicators are also interacted with individual, household, and community characteristics to test for other differential effects of parent death. For example, the magnitude of the parent death effect may depend on child age because older children are better labor market substitutes for parents, perhaps making them more likely to drop out of school after an adverse household income shock (although, as we discuss later, we do not find this in our setting).

Potential Bias Due to Unobserved Time-Varying Factors

The key concern for this econometric identification strategy is the possibility of unobserved time-varying factors that affect both parent health and child schooling. The most plausible sources are local weather and crop price shocks, but these are captured in the region-year indicator variables (ρ_{jt}) included in all specifications. Another such shock could be parent job loss. However, in the study area, most adults engage in subsistence agriculture, and few have formal sector jobs to lose. Note that child morbidity due to HIV infections contracted from parents is unlikely to affect the estimation because the overwhelming majority of children born HIV-positive in rural Africa die before reaching school age (Adetunji 2000).

To begin addressing the issue, we restrict attention to children whose parents were both alive at baseline in 1998 and compare those whose parents subsequently died during the period 1999–2002 with those whose parents did not die. We make the case that these two groups—the "became orphan" and "never orphan" groups, respectively—are comparable along a range of observable baseline characteristics. There are no significant differences in terms of baseline school participation or demographic characteristics in the full sample (Appendix Table A2, Panel A). In the restricted sample, the two groups are remarkably similar along 14 characteristics, including measures of child nutrition and health and household socioeconomic status (Panel B). There are statistically significant but minor differences in child cleanliness and age, as well as in baseline 1998 test scores (with an average difference of 0.13 standard deviations). For the 2,923 students for whom we have 1997 school participation data—gathered for the evaluation of an education intervention in a subset of sample schools—there is no significant difference between the 1997–1998 school participation trends for children who later became orphans and those who did not (Panel C), evidence that they were similar in terms of both schooling levels and trends.

These arguments do not completely eliminate concerns about the suitability of the comparison group, yet we feel that these patterns allay most reasonable concerns about bias. If the "became orphan" and "never orphan" groups indeed differed sharply along unobserved dimensions, such as parents' commitment to education or their discount rate, it is likely that these differences would also be reflected along some observable dimensions, but we do not find systematic differences.

In a further attempt to address unobserved time-varying factors, in some specifications, we include the baseline controls—including those that differed significantly across the "became orphan" and "never orphan" groups, such as child cleanliness, age, and baseline 1998 test score—interacted with a full set of year indicator variables, and we find that the main empirical results are unchanged, as discussed later. As an additional robustness check, we compare the children who became orphans with the 2,676 children who began the study period as orphans. This specification yields nearly identical parent death estimates.

Potential Bias Due to Attrition and Measurement Error

We next test whether students with missing data are significantly different from other students along observable dimensions. Our baseline sample of 24,111 children is reduced because of missing information on parent deaths. In our first analysis, the dependent variable is an indicator that takes on a value of 1 if the child is missing information on parent death. The dependent variable has a mean of 0.25; that is, 25% of the sample lacks death information for one or both parents. In the baseline sample of 24,111 children, the older children and, not surprisingly, those with missing data on age are more likely to have missing orphan status information (Appendix Table A3, regression 1). Among those with 1998 Pupil Questionnaire data, most indicators of household asset ownership are negatively related to missing orphan information, including ownership of cattle, goats, and poultry, suggesting that poorer households are more likely to be lost (regression 2). This could bias us against finding strong parent death effects. Children wearing shoes were significantly more likely to have missing data, while the opposite holds for those wearing a uniform, although the explanation for this pattern is unclear.

We next consider attrition—namely, missing data on school participation—as the dependent variable. School participation was recorded as missing when the child was absent from school and her former peers and teachers did not know her current school participation status, perhaps because the child had moved and was out of tracking range. While children with missing parent death information are uniformly dropped from the sample, children missing school participation information are dropped from the sample only in those years for which data are missing; some children exit and reenter the sample after a temporary residential move. In Appendix Table A3, the dependent variable takes on a value of 1 if the child attrited at any point during 1998–2002.

Children who we know became orphans are significantly less likely to have missing schooling data (Appendix Table A3, regression 3). This result appears counterintuitive at first but is consistent with the notion that reliable orphan status information is more likely to be missing altogether for children who have already attrited. However, if even a small fraction of the children with unknown orphan status are in fact orphans, then the coefficient estimate on the orphan indicator becomes positive; for instance, if a randomly chosen subset of just 10% of the children whose orphan status is unknown (only two-thirds of the actual proportion of orphans in this population; see Appendix Table A1) are assigned to be orphans in a simulation, the coefficient estimate on the orphan term in regression 3 becomes positive (results not shown).

The attrition effects are significantly larger for maternal deaths than for paternal deaths (regressions not shown). This is consistent with Evans's (2004) finding that maternal orphans are 50% more likely not to live with a surviving parent than paternal orphans and hence are more likely to move away from the area, exiting the sample. In the restricted sample (Appendix Table A3, regression 4), household asset ownership measures are not consistently related to attrition, but older girls are more likely to have missing schooling data than younger girls.

Several possible sources of bias remain, but most tend to bias the parent death estimates toward zero, leading our estimates to serve as bounds. In other words, actual impacts of parent death are likely to be even larger than our estimates, although it is difficult to specify exactly how much larger they are. First, children with unknown orphan status have lower baseline school participation and less household asset ownership than other children (Appendix Table A2, right column). To the extent that orphans are more likely to drop out of school and leave the area, it is likely that these children with unknown orphan status are also disproportionately orphans, in which case, excluding them would likely lead us to underestimate actual impacts of parent death.

Second, we are likely to further underestimate parent death impacts if the surveys captured information on parent death years with error (Aigner 1973), as we discuss further later. Another reason that true effects may be underestimated is that this data set does not include information on future parent deaths unknown at the time of data collection. In other words, in the 2002 data, children who (unbeknownst to the econometrician) will experience a parent death in 2003 or 2004 could already be experiencing adverse impacts because of a parent's AIDS-related morbidity, and this would reduce average school participation of some classified as "never orphans." Yet this latter bias is likely to be very small: only 2% of baseline non-orphans became an orphan each year during 1999–2002 (Appendix Table A1), and the estimated pre-death effects are moderate, so the product of these two quantities is negligible.

EMPIRICAL RESULTS

Parent Death Impacts

School participation is similar for "became orphans" and "never orphans" three years before parent death, but it begins to drop two years before parent death, drops sharply again in the year of the death, and remains at a lower level for at least three years afterward (Table 1, regression 1). The small but growing gaps between orphans and non-orphans during the two years prior to parent death are consistent with the duration of AIDS-related

Table 1. Ordinary Least Squares Coefficients Predicting the Impact of Parent Death on School Participation

Variable	Full Sample (1)	Full Sample (2)	Restricted Sample (3)	Restricted Sample (4)	Restricted Sample (5)	Restricted Sample (6)
3 Years Before Death	–0.013 (0.026)		0.007 (0.036)			
2 Years Before Death	–0.037 (0.029)		–0.037 (0.038)			
1 Year Before Death	–0.039 (0.030)		–0.037 (0.041)			
Year of Parent Death	–0.074* (0.031)		–0.054 (0.041)			
1 Year After Death	–0.060* (0.030)		–0.054 (0.043)			
2 Years After Death	–0.065* (0.033)		–0.071 (0.049)			
3 Years After Death	–0.089* (0.040)		–0.070 (0.059)			
Before Parent Death (1–2 years)		–0.021 (0.015)		–0.032 (0.019)	–0.031 (0.019)	–0.025** (0.009)
After Parent Death		–0.055** (0.017)		–0.054* (0.022)	–0.053* (0.022)	–0.036** (0.012)

(continued)

parental morbidity described in the existing literature (Morgan et al. 2000; Morgan and Whitworth 2001). There is no evidence of orphan recovery after parent death, in contrast to some other work (Ainsworth et al. 2005), suggesting that long-run effects of parent death are possibly quite large. The equality of parent death impacts three years before and three years after the death is rejected at over 95% confidence (in regression 1 using an F test), but the equality of parent death effects during the year of parent death and the following three years is not rejected (p = .56), nor is the equality of effects in the two years immediately before death (p = .86).

We combine the two years before parent death to estimate pre-death morbidity effects and combine the post-death years in most subsequent specifications, as in Eq. (2) above. In this specification, parent death has a moderate negative impact on child school participation: on average, school participation falls by 5.5 percentage points (or 0.055, SE = 0.017, statistically significant at 99% confidence) after parent death in a specification with individual fixed effects,[10] and the average effect in the two years before the parent death is again negative, though not significant (–0.021, SE = 0.015; Table 1, regression 2). In a specification without child fixed effects, similar to cross-sectional specifications found in existing literature, the analogous point estimates are smaller in magnitude, at –0.040 (SE = 0.007) for the post-death effect and –0.018 (SE = 0.007; regression not shown) for the pre-death effect. This implies that omitted variable bias in this sample is positive and would lead us to understate parent death impacts if child fixed effects were not included.

The time pattern of effects is similar with the smaller restricted sample, those children from the full sample for whom 1998 Pupil Questionnaire data are available

10. The analogous estimate in Yamano and Jayne (2005: table 4, regression A) for a specification including child and household controls is similar, at –0.060.

(Table 1, continued)

Variable	Full Sample (1)	Full Sample (2)	Restricted Sample (3)	Restricted Sample (4)	Restricted Sample (5)	Restricted Sample (6)
Baseline Household Controls						
Child's weight-for-age (z score), 1998						0.010** (0.004)
Child had malaria/fever in past month, 1998						–0.010* (0.005)
Child wears shoes, 1998						0.017† (0.009)
Child wears school uniform, 1998						0.035** (0.009)
Child appears "clean," 1998						0.016** (0.005)
Latrine at home, 1998						0.007 (0.007)
Cows at home, 1998						0.020** (0.006)
Goats at home, 1998						–0.005 (0.006)
Poultry at home, 1998						0.022* (0.009)
Student Fixed Effects	Yes	Yes	Yes	Yes	Yes	No
Baseline Controls × Year Controls	No	No	No	No	Yes	Yes
Observations	73,070	73,070	30,817	30,817	30,817	30,817
Mean (SD) of Dependent Variable	0.75 (0.35)	0.75 (0.35)	0.77 (0.34)	0.77 (0.34)	0.77 (0.34)	0.77 (0.34)
R^2	0.54	0.54	0.58	0.58	0.58	0.11

Notes: Standard errors, shown in parentheses, are clustered at the school level. Unreported controls include deworming program (PSDP) treatment variables, the full set of birth-year cohort-year-gender indicator variables, and region-year indicator variables. Regressions 1 and 2 contain 18,133 unique pupils, and regressions 3–6 contain 7,815 unique pupils. The before and after death indicator variables are mutually exclusive categories. The additional controls in regressions 5 and 6 include all of the baseline household controls interacted with indicator variables for the years 1999, 2000, 2001, and 2002.

†$p < .10$; *$p < .05$; **$p < .01$

(Table 1, regression 3); the estimated parent death effect is nearly identical at –0.054 (SE = 0.022, significant at 95% confidence; regression 4), while the pre-death effect is somewhat larger, at –0.032 (SE = 0.019, significant at 90% confidence). All baseline covariates (shown in Table 1) interacted with the year indicator variables are next included to partially address concerns related to time-varying omitted variables, and the results are unchanged (regression 5). When the baseline test score is interacted with the year controls as well, the sample falls slightly to 28,665 child-year observations, but the coefficient estimates are almost identical (not shown). Both parent death and pre-death effects remain statistically significant when baseline controls are included instead of fixed effects, but point estimates are somewhat smaller in magnitude (Table 1, regression 6), again suggesting that fixed effects address omitted variable bias.

Table 2. Ordinary Least Squares Regression Coefficients Predicting the Impact of Parent Death on School Participation, Alternative Comparison Groups

Variable	Became Orphans Versus Never Orphans (Full Sample) (1)	Became Orphans Versus Always Orphans (2)	Became Orphans Versus Never Orphans Always Orphans (3)
Before Parent Death (1–2 years)	–0.021	–0.008	–0.010
	(0.015)	(0.012)	(0.010)
After Parent Death	–0.055**	–0.042**	–0.045**
	(0.017)	(0.013)	(0.014)
Student Fixed Effects	Yes	Yes	Yes
Observations	73,070	19,176	85,713
Mean (*SD*) of Dependent Variable	0.75	0.73	0.74
	(0.35)	(0.35)	(0.33)
R^2	0.54	0.56	0.54

Notes: Standard errors, shown in parentheses, are clustered at the school level. Unreported controls include deworming program (PSDP) treatment variables, a full set of birth-year cohort-year-gender indicator variables, and region-year indicator variables. Regression 1 contains 18,133 unique pupils, regression 2 contains 4,690 unique pupils, and regression 3 contains 21,348 unique pupils. Regression 1 reproduces the result in Table 1, regression 2.

$^{**}p < .01$

Effects are robust to an alternative schooling measure, the school enrollment indicator variable, yielding estimated magnitudes similar to those Gertler, Levine, and Ames (2004) estimated for Indonesia—more than a doubling of the drop-out rate after parent death—with time patterns similar to school participation (regressions not shown). This is a useful measure to consider because it is closer to the schooling data used in other studies. However, there is a crucial difference between our measure and other ones: existing enrollment measures are based on household reports, while ours are based on unannounced school visits. Thus, our school enrollment variable is not directly comparable to other measures.

The main results are robust to the use of an alternative comparison group, those children who began the study period as orphans. In a fixed effects specification, the estimated parent death effect is –0.042 (Table 2, regression 2), and the pre-death effect is –0.008. When the "always orphan" and "never orphan" groups are both used as the comparison group, the parent death effect remains stable at –0.045 and significant at 99% confidence (regression 3).

Impacts by Parent, Child, and Household Characteristics

Maternal deaths have a much larger impact than paternal deaths, and most of the difference is driven by the sharp drop in school participation among children in the two years before their mother dies: the maternal pre-death effect is –0.065 (*SE* = 0.022, statistically significant at 99% confidence), and the post-death effect is –0.093 (*SE* = 0.025; Table 3, regression 1). The analogous effects for fathers are less than half as large, with the father death effect at –0.036 (*SE* = 0.022, not statistically significant at traditional confidence levels) and a pre-death effect of only –0.005 (*SE* = 0.018). The difference between maternal and paternal effects before death is statistically significant at 95% confidence ($p = .03$) and after death at 90% confidence ($p = .09$).

This finding implies that the encouragement and income provided by (healthy) mothers is more important, on average, in determining child schooling participation than the encouragement and income provided by fathers in rural Kenya. The disruptions caused by fostering may also account for part of the large maternal death effect. Evans (2004) found

Table 3. Ordinary Least-Squares Regression Coefficients Predicting the Impact of Maternal and Paternal Deaths on School Participation: Full Sample

Variable	(1)	(2)	(3)
Before Maternal Death (1–2 years)	–0.065**	–0.065**	–0.067**
	(0.022)	(0.022)	(0.025)
After Maternal Death	–0.093**	–0.096**	–0.091**
	(0.025)	(0.026)	(0.029)
Before Paternal Death (1–2 years)	–0.005	–0.005	–0.009
	(0.018)	(0.018)	(0.029)
After Paternal Death	–0.036	–0.037	–0.032
	(0.022)	(0.023)	(0.030)
After Maternal Death × After Paternal Death		0.014	
		(0.037)	
Before First Parent Death			0.004
			(0.023)
After First Parent Death			–0.004
			(0.026)
Student Fixed Effects	Yes	Yes	Yes
Observations	73,070	73,070	73,070
Mean (*SD*) of Dependent Variable	0.75	0.75	0.75
	(0.33)	(0.33)	(0.33)
R^2	0.54	0.54	0.54

Notes: Standard errors, shown in parentheses, are clustered at the school level. Unreported controls include deworming program (PSDP) treatment variables, birth-year cohort-year-gender indicator variables, and region-year indicator variables. All regressions contain 18,133 unique pupils.

$^{**}p < .01$

that orphans are significantly more likely to be sent to live in other households following maternal deaths than following paternal deaths. These different fostering patterns by parent gender allow the possibility that sample attrition is driving some of the difference between maternal and paternal death impacts, but we believe that differential attrition, if anything, is likely to lead us to understate maternal death impacts because low-performing orphans are more likely to leave the sample. Another key difference between fathers and mothers that could in part account for this pattern is presence at home. In the 1998 Kenya DHS survey in Western Province (ORC Macro 2005), only 67% of children aged 0–14 years with both parents alive actually lived with their fathers, whereas the analogous percentage for mothers is much higher at 89%. The death of an absent parent is unlikely to have as large an impact on children as the death of a present parent.

The additional impact of becoming a double orphan, on top of the summed effects of losing both mother and father, is near zero and not statistically significant (Table 3, regression 2), although there is limited statistical precision as a result of the small number of double orphans (recall that all children in these regressions began the sample period as non-orphans). Note that the different maternal and paternal death effects are not simply the result of the fact that paternal deaths usually precede maternal deaths (regression 3), and among those who have lost both parents, there is no significant difference between having lost one's father first versus one's mother first (regression not shown).

Young children (under age 12 at parent death) are somewhat more likely (at 90% confidence) to drop out of school following a parent death in one specification (Table 4,

Table 4. Ordinary Least-Squares Regression Coefficients Predicting the Impact of Parent Death on School Participation, by Child Age and Gender: Full Sample

Variable	(1)	(2)	(3)	(4)	(5)
Before Parent Death (1–2 years)	–0.011	0.010	–0.024	–0.003	–0.010
	(0.017)	(0.020)	(0.019)	(0.022)	(0.019)
After Parent Death	–0.036	–0.011	–0.051*	–0.030	–0.051†
	(0.022)	(0.025)	(0.022)	(0.025)	(0.027)
Before Maternal Death (1–2 years)		–0.068*		–0.066*	
		(0.032)		(0.030)	
After Maternal Death		–0.076†		–0.066†	
		(0.039)		(0.037)	
Child Below Age 12 × Before Parent Death (1–2 years)	–0.038	–0.040			–0.040
	(0.026)	(0.034)			(0.032)
Child Below Age 12 × After Parent Death	–0.049†	–0.051			0.001
	(0.025)	(0.034)			(0.034)
Child Below Age 12 × Before Maternal Death (1–2 years)		0.007			
		(0.058)			
Child Below Age 12 × After Maternal Death		0.005			
		(0.061)			
Female Child × Before Parent Death (1–2 years)			0.006	0.006	–0.003
			(0.027)	(0.030)	(0.033)
Female Child × After Parent Death			–0.007	–0.004	0.034
			(0.030)	(0.031)	(0.042)
Female Child × Before Maternal Death (1–2 years)				0.004	
				(0.040)	
Female Child × After Maternal Death				–0.005	
				(0.052)	
Child Below Age 12 × Female × Before Parent Death					0.008
					(0.051)
Child Below Age 12 × Female × After Parent Death					–0.104†
					(0.060)
Student Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	73,070	73,070	73,070	73,070	73,070
Mean (*SD*) of Dependent Variable	0.75	0.75	0.75	0.75	0.75
	(0.33)	(0.33)	(0.33)	(0.33)	(0.33)
R^2	0.54	0.54	0.54	0.54	0.54

Notes: Standard errors, shown in parentheses, are clustered at the school level. Unreported controls include deworming program (PSDP) treatment variables, birth-year cohort-year-gender indicator variables, and region-year indicator variables. "Child below age 12" refers to their age in the year of the parent death. Regressions 1, 2, and 5 also include the indicator variable for missing data on age and interactions between the indicator for missing data on age and parent death terms; coefficient estimates not reported. All regressions contain 18,133 unique pupils.

†$p < .10$; *$p < .05$

regression 1).[11] The explanation may lie in the underlying academic ability of enrolled primary school students of different ages. Given high drop-out rates during primary school in Kenya, those students still in school at baseline during their teenage years are positively selected on academic ability; thus, this result that older children are less likely to drop out

11. Regressions 1 and 2 in Table 4 also include an indicator for "missing age data" and interactions between "missing age data" and the parent death terms (coefficients not reported).

following a parent death—despite the higher opportunity costs for older children, since their labor market prospects are better—is a first hint that academically stronger students are less likely to be removed from school following a parent death, a finding we confirm more conclusively later. The gender of the parent who dies does not differentially affect young versus older children (regression 2).

Girls are no more likely than boys overall to experience decreased school participation following a parent death (Table 4, regression 3), and this holds independent of the gender of the parent who dies (regression 4). However, the double interaction specification suggests that young girls are the most likely to experience decreased school participation following a parent death (statistically significant at 90% confidence, regression 5). Both young and older boys experience average school participation drops of approximately five percentage points following a parent death (summing the relevant coefficient estimates in regression 5); older girls experience a drop of two percentage points, while for young girls under age 12, the coefficient estimate is 12 percentage points (although this estimate is quite imprecise). Maternal versus paternal deaths do not have significantly different effects on young girls' school participation relative to other groups, although small cell sizes and limited statistical power are a concern when triple interactions of this sort are examined (regression not shown).

The likelihood that an orphan is removed from school should in part depend on the child's expected returns to continued schooling, which is likely to be an increasing function of her academic ability in this context because only the best students are typically able to continue to secondary school. We use the child's normalized baseline 1998 academic test score as a measure of ability and find that parent death has the most adverse negative impacts on children with low baseline scores, while children with high baseline scores are largely unaffected. For a child with a test score of 0 at baseline (the mean score by construction), the pre-death effect is –0.029 (*SE* = 0.019; Table 5, regression 1) and the parent death effect is –0.053 (*SE* = 0.021, statistically significant at 95% confidence); the analogous pre- and post-death impacts for a child with a baseline test score of +1 standard deviation are essentially 0 (at 0.009 and 0.004, respectively, neither of which is statistically significant). In contrast, the post-death effect for a child with an initial baseline test score of –1 standard deviation is extremely large, at nearly –0.11, twice the magnitude for a child with an average baseline score.

Household asset ownership is not robustly associated with parent death impacts. The coefficient estimate on the interaction term of parent death and not having a latrine at home is not statistically significant (Table 5, regression 2), nor are the interaction terms with a poverty index similar to that used in existing studies[12] (regression 3). These findings suggest that increasing levels of household wealth do not buffer children from parent death, at least at the low asset levels found in our sample (although these are rough socioeconomic proxies, and so results should be interpreted with caution). The baseline test score interaction terms remain nearly unchanged when the poverty index and interaction are included (regression 4), indicating that the test score is not simply proxying for household socioeconomic status, an important robustness check.

Community Impacts

Orphans do not fare significantly worse in primary school communities with higher orphan rates, although limited statistical precision means that we cannot rule out moderate negative impacts. The point estimate is not statistically significant (the point estimate on the interaction term is –0.232, *SE* = 0.230; regressions not shown). The result that the local orphan

12. Following Filmer and Pritchett (2001), we use principal components to construct an index of household assets, including latrines, cows, goats, poultry, shoes, and school uniforms, as well as child cleanliness. Unfortunately, we lack detailed information on parent occupation.

Table 5. Ordinary Least-Squares Coefficients Predicting the Impact of Parent Death on School Participation, by Child and Household Characteristics: Restricted Sample

Variable	(1)	(2)	(3)	(4)
Before Parent Death (1–2 years)	–0.029	–0.033	–0.041†	–0.039†
	(0.019)	(0.020)	(0.021)	(0.021)
After Parent Death	–0.053*	–0.040	–0.054*	–0.054*
	(0.021)	(0.025)	(0.026)	(0.024)
1998 Test Score × Before Parent Death (1–2 years)	0.038*			0.038*
	(0.018)			(0.018)
1998 Test Score × After Parent Death	0.057*			0.056*
	(0.023)			(0.023)
No Latrine at Home × Before Parent Death (1–2 years)		0.008		
		(0.043)		
No Latrine at Home × After Parent Death		–0.072		
		(0.050)		
Poor Household × Before Parent Death (1–2 years)			0.052	0.055
			(0.042)	(0.042)
Poor Household × After Parent Death			0.005	0.009
			(0.055)	(0.058)
Student Fixed Effects	Yes	Yes	Yes	Yes
Observations	28,665	30,817	30,817	28,665
Mean (*SD*) of Dependent Variable	0.77	0.77	0.77	0.77
	(0.34)	(0.34)	(0.34)	(0.34)
R^2	0.54	0.58	0.58	0.54

Notes: Standard errors, shown in parentheses, are clustered at the school level. Unreported controls include deworming program (PSDP) treatment variables, birth-year cohort-year-gender indicator variables, and region-year indicator variables. “Poor” is an indicator variable that takes on a value of 1 for students whose households are in the bottom quintile of a poverty index; the index is created using a principal components approach, and the inputs are the household socioeconomic measures (latrine ownership, cow ownership, goat ownership, poultry ownership, child wears shoes, child wears school uniform, and child is “clean”). Regressions 1 and 4 contain 7,210 unique pupils, and regressions 2 and 3 contain 7,815 unique pupils.

$^{\dagger}p < .10$; $^{*}p < .05$

rate is not strongly associated with orphan schooling is robust to an alternative definition of overall local orphan burden and to examining local maternal and paternal orphanhood separately.[13] Results are similar using initial 1998 orphan rates rather than contemporaneous rates by year (regressions not shown). Any negative spillovers of higher orphan populations on other community members are also not apparent: non-orphans do not fare significantly worse on average in communities with a higher proportion of orphans (coefficient estimate = –0.310, *SE* = 0.335; regression not shown).

These findings suggest that recent claims in the popular media that social networks in rural Africa are rapidly breaking down under the strain of HIV/AIDS deaths—and that as a result, neither orphans nor other children can be adequately taken care of by surviving relatives—are probably overstated. Further research is needed to understand how general these findings are beyond rural Kenya, of course. But the fact that there is little evidence that networks are breaking down in this region, with its relatively high orphan rates, suggests that this issue is even less of a concern in other rural areas where orphan rates are lower.

13. In contrast, Yamano and Jayne (2005) found that school attendance is negatively correlated with lagged provincial HIV prevalence. (Note that their measure is at a higher level of aggregation.)

DISCUSSION AND LIMITATIONS

Yamano and Jayne (2005) produced the most closely related work to the current study, but our current study represents several improvements. Our larger sample size and continuous data collection throughout the study period allow us to more precisely estimate impacts; Yamano and Jayne's data set included observations for only three years. Another advantage of our approach is the use of schooling data collected at school by enumerators during unannounced visits, rather than relying on parent or caregiver reports; Yamano and Jayne and other studies relied on the latter. Finally, we estimate parent death effects directly, whereas Yamano and Jayne estimated the impact of the death of any adult in the child's household on schooling because they lacked parent death information. In Kenya, where many households contain adults other than children's parents (Yamano and Jayne 2005: table A1), they thus estimated a different parameter.

Our study has other important limitations. The estimation approach does not permit us to estimate broad regional or national effects of the HIV/AIDS epidemic on primary school participation, for example, due to reduced national school funding, teacher shortages, or decreased demand for education (which is theoretically possible in a society in which life expectancy is dropping rapidly). A cross-region or cross-country analysis is needed to capture these broader impacts. The estimates we present in this study also miss effects of parent death on the schooling of children below age five, who may never enroll in primary school and thus are not in our data set. As with any microeconomic empirical study, questions of generalizeability remain important because impacts could differ across settings—in rural versus urban areas, for instance, or as a function of school fees—issues we cannot address in this study's entirely rural Kenyan sample.

It is also worth stressing that we have neither individual biomedical information on HIV infection status nor data on whether the cause of a parent death was AIDS. Although many adult deaths are likely to be HIV/AIDS–related, we cannot determine the exact proportion in our sample. Thus, we are unable to test whether AIDS orphans fare differently than other orphans, because of AIDS stigma, for instance. Nonetheless, UNAIDS et al. (2002) estimated that in Kenya as a whole, where the HIV prevalence rate was estimated at 6.7% in 2003 (Marum et al. 2004), 54% of orphans up to 14 years old had lost at least one parent to AIDS. This proportion is also likely to be high in our study region (Western Province), where the prevalence was 4.9%. Thus, we feel confident that HIV/AIDS–related illnesses are a leading cause of parent death in our sample.

Finally, we conduct two simulation exercises to understand the likely effect of data problems on our estimates. The first simulation places bounds on parent death impacts accounting for both missing information on parent deaths and missing schooling data. The second exercise establishes the likely degree of attenuation bias resulting from mismeasured years of parent death.

Placing bounds on effects due to missing data generates a wide but always negative range of estimated parent death impacts. Thus, even under implausibly conservative assumptions, zero is a bound on parent death impacts, so we feel confident in asserting that parent death has a negative impact on school participation in rural Kenya.

This exercise requires assumptions for three groups of children: (1) those who are missing school participation data only (i.e., those subject to attrition; Group A), (2) those who are missing orphan status data only (Group B), and (3) those who are missing both (Group C). To establish an upper bound on the magnitude of impacts, we assume for Group A children that school participation is equal to 0 if they become orphans (with zero participation starting two years before the parent death) and 1 if they are not orphans. We assume that Group B children become orphans if their school participation is lower than the mean school participation rate among the full sample; otherwise, we assume that they do not become orphans. Finally, we assume that all children in Group C become orphans and that their school participation is equal to 0 (starting two years before the parent death), an extreme bounding procedure related to Manski (1995). The timing of observed parent deaths across years is used to generate simulated parent death years for the children with missing information on parent mortality who are designated "became orphans."[14]

In a fixed-effects specification (analogous to Table 1, regression 2), and with 100 runs of the simulation, the average upper bound is a massive –0.36 decrease in school participation after parent death and a –0.14 decline in the two years before death. Because orphanhood is often disruptive to living arrangements (Evans 2004), making orphans more likely to leave the sample than non-orphans, it is plausible that the actual effect of parent death lies between our estimated effects and these upper bounds, rather than lying closer to the lower bound presented below.

The lower bound procedure makes polar opposite assumptions: for Group A, school participation is assumed to be 1 if children become orphans (with perfect school participation starting two years before the parent death) and 0 for non-orphans. Children in Group B are assumed to become orphans if their school participation is above average, and 0 otherwise. For Group C, we assume that the same proportion of children become orphans as in the full sample, and that these children all have perfect school participation (starting two years before the parent death), while the remaining children are assumed to be non-orphans with 0 school participation. This yields a mean lower bound of a –0.02 decrease after parent death (and a 0.02 increase in school participation before death), again in 100 runs of the simulation.

In the second exercise, the year of parent death is replaced with a "noisy" proxy to assess the likely magnitude of attenuation bias in our estimates. Specifically, we take the difference between the year of parent death recorded in the home tracking survey versus in the school tracking survey to be the distribution of errors in year of parent death reports, assuming perfect accuracy of the home reports. For example, 71% of reports of the year of paternal death were accurate, 6% were one year "early," and 23% were one year "late." In 500 runs of the same specification as in Table 1, regression 2, but in which the year of parent death is replaced with the "noisy" proxy, we find that the average pre-death effect is –0.014 (with an average *SE* of 0.014) and the average post-death effect is –0.043 (average *SE* = 0.016). These compare to pre- and post-death effects of –0.021 and –0.055, respectively, in Table 1, indicating that measurement error in the year of parent death of the magnitude found in our data leads to moderate attenuation in estimates of the impact of parent death, further evidence that our main estimates are lower bounds.

Because maternal deaths are less reliably reported than paternal deaths, we carry out similar simulations allowing for both maternal and paternal effects (as in Table 3, regression 1), taking into account the different measurement error patterns in these two types of reports. This allows us to test whether the estimated differences between maternal and paternal death effects can be attributed to differences in measurement error, but we do not find that this is the case. We find that the pattern of statistical significance is nearly identical to that shown in Table 3: average maternal pre-death effects are –0.043 (average *SE* = 0.021), maternal post-death effects are –0.068 (average *SE* = 0.023), paternal pre-death effects are –0.003 (average *SE* = 0.019), and paternal post-death effects are –0.030 (average *SE* = 0.022). Thus, estimated pre-death effects fall equally in the simulation for both maternal and paternal deaths, and the post-death effect falls more for maternal than for paternal deaths (28% vs. 15%), suggesting that the true gap between maternal and paternal effects could be even larger.

14. For Groups A and C, we focus on students whose school participation data are missing through the final year of the sample (e.g., for 2000–2002), and these cases constitute a large fraction of all missing data. The bounds on parent death effects are somewhat wider if other missing school participation observations (e.g., just one year of missing data in the middle of the sample—say, in 1999) are also replaced with extreme values in the exercise.

CONCLUSION

To summarize our main findings, we find a substantial and highly statistically significant negative impact of parent death on primary school participation, and our estimates are likely to be lower bounds on true effects. Impacts are more than twice as large for maternal deaths (at 9 percentage points) than paternal deaths (at 4 percentage points). A striking result is that children with lower baseline (pre-death) academic test scores experience significantly larger decreases in school participation after a parent death than children with high test scores, suggesting that households decide to focus their increasingly scarce resources after a parent death on more promising students. This result informs a previously unexamined area and sheds new light on the priorities of rural Kenyan households following an economic shock. Our empirical approach addresses a number of methodological shortcomings of recent studies. In particular, we find evidence of omitted variable bias in cross-section estimates, although more research is clearly needed from other settings to establish the generality of this result.

Our results provide insight into debates over how to target assistance programs to mitigate the impact of HIV/AIDS on education in Africa and, in particular, whether orphans should be specifically targeted or whether transfers should instead be directed to all poor children. The latter position has been advocated by many in the field (Ainsworth and Filmer 2002; Lundberg and Over 2000), often drawing on the results of earlier studies that found relatively small parent death impacts. However, our results indicate that orphans are a particularly disadvantaged group in terms of schooling, even relative to other poor children, and suggest that transfers targeted to orphans directly might be beneficial in the rural Kenyan context. Children whose mothers have died experience particularly adverse schooling impacts, and this easily observable characteristic could be used to improve targeting. Still, we think that a better understanding of the underlying theoretical mechanisms—for instance, the role of resource constraints versus psychological factors versus fostering patterns—is necessary to develop effective policy recommendations in this area.

Appendix Table A1. Summary Statistics

Variable	Observed	Mean	*SD*	Min.	Max.
Panel A. Full Sample					
Female	18,133	0.48	0.5	0	1
Age, 1998	14,970	11.8	2.5	5	18
Became an orphan during 1999–2002	18,133	0.08	0.27	0	1
Became a maternal orphan during 1999–2002	18,133	0.03	0.17	0	1
Became a paternal orphan during 1999–2002	18,133	0.06	0.23	0	1
Proportion of orphans in school, 1998	18,133	0.15	0.05	0.01	0.41
Proportion of maternal orphans in school, 1998	18,133	0.05	0.02	0.01	0.20
Proportion of paternal orphans in school, 1998	18,133	0.12	0.04	0	0.33
Proportion of double orphans in school, 1998	18,133	0.02	0.01	0	0.12
School participation, 1998	18,133	0.85	0.23	0	1
School enrollment, 1998	18,133	0.98	0.14	0	1

(continued)

(Appendix Table A1, continued)

Variable	Observed	Mean	*SD*	Min.	Max.
Panel B. Restricted Sample					
Female	7,815	0.48	0.50	0	1
Age, 1998	7,769	12.9	2.0	6	18
Became an orphan, 1999–2002	7,815	0.09	0.28	0	1
Became a maternal orphan during 1999–2002	7,815	0.03	0.17	0	1
Became a paternal orphan during 1999–2002	7,815	0.06	0.24	0	1
Proportion of orphans in school, 1998	7,815	0.14	0.05	0.01	0.35
Proportion of maternal orphans in school, 1998	7,815	0.05	0.02	0.01	0.15
Proportion of paternal orphans in school, 1998	7,815	0.11	0.04	0	0.31
Proportion of double orphans in school, 1998	7,815	0.02	0.01	0	0.12
School participation, 1998	7,815	0.92	0.17	0	1
School enrollment, 1998	7,815	1	0.06	0	1
Child's weight-for-age (*z* score), 1998	7,815	–1.44	0.82	–4.79	2.34
Child had malaria/fever in past month, 1998	7,815	0.39	0.49	0	1
Child wears shoes, 1998	7,815	0.14	0.35	0	1
Child wears school uniform, 1998	7,815	0.86	0.34	0	1
Child appears "clean," 1998	7,815	0.62	0.49	0	1
Latrine at home, 1998	7,815	0.82	0.38	0	1
Cows at home, 1998	7,815	0.49	0.50	0	1
Goats at home, 1998	7,815	0.41	0.49	0	1
Poultry at home, 1998	7,815	0.93	0.25	0	1

Notes: School participation variables are from regular unannounced checks collected throughout the 1998–2002 school years (see Miguel and Kremer 2004). Orphan status variables are from the 2002 Tracking Data. Demographic and socioeconomic characteristics are from the 1998 Pupil Questionnaire. The reduced samples for "age" are due to missing data; in the regressions, we include an indicator for observations with missing data on age.

Appendix Table A2. Baseline Characteristics for Children Who Lost a Parent Versus Others

Variable	Became Orphans	Never Orphans	Difference Became – Never (*SE*)		Orphan Status Unknown
Panel A. Full Sample					
Female	0.46	0.48	–0.02	(0.02)	0.51
Age, 1998	11.8	11.8	0.0	(0.1)	12.4
School participation, 1998	0.87	0.87	0.00	(0.01)	0.76
School enrollment, 1998	0.99	0.99	0.00	(0.00)	0.93
Observations	1,245	13,725			5,978

(continued)

(Appendix Table A2, continued)

Variable	Became Orphans	Never Orphans	Difference Became – Never (*SE*)		Orphan Status Unknown
Panel B. Restricted Sample					
Female	0.48	0.48	0.00	(0.02)	0.53
Age, 1998	12.7	12.9	–0.2**	(0.1)	13.5
School participation, 1998	0.92	0.92	0.00	(0.01)	0.87
School enrollment, 1998	1.00	1.00	0.00	(0.00)	0.99
Academic test score, 1998 (normalized)	–0.08	0.05	–0.13**	(0.05)	–0.01
Child's weight-for-age (*z* score), 1998	–1.40	–1.45	–0.04	(0.03)	–1.34
Child had malaria/fever in past month, 1998	0.40	0.39	0.01	(0.02)	0.42
Child wears shoes, 1998	0.13	0.14	–0.01	(0.01)	0.19
Child wears a school uniform, 1998	0.85	0.86	–0.02	(0.01)	0.85
Child appears "clean," 1998	0.59	0.62	–0.03	(0.02)	0.64
Latrine at home, 1998	0.81	0.82	–0.01	(0.02)	0.81
Cows at home, 1998	0.49	0.49	0.00	(0.03)	0.44
Goats at home, 1998	0.39	0.41	–0.02	(0.02)	0.37
Poultry at home, 1998	0.93	0.93	0.00	(0.01)	0.91
Observations	667	7,148			1,938
Panel C. Subsample of Children With School Participation Data in 1997 and 1998					
School participation, 1997	0.84	0.81	0.03	(0.02)	0.75
School participation, 1998	0.80	0.79	0.01	(0.02)	0.67
School participation, 1998–1997	–0.04	–0.02	–0.02	(0.03)	–0.09
Observations	250	2,673			904

Notes: Standard errors are clustered at the school level. For Panel C, 27 schools are included which were involved in another NGO program, and thus had 1997 attendance data. The reduction in sample size in Panel A (from 18,133 total students in the full sample to 14,970) is due to missing age information. The final column includes children who would be in the full sample (or restricted sample) but for the lack of parent mortality data. The 1998 test scores are available for a somewhat smaller sample of 622 "became orphans" and 6,588 "never orphans."

$^{**}p < .01$

Appendix Table A3. Attrition and Child Characteristics

Variable	Missing Information on Parent Death in 2002 (1)	(2)	Attrited During 1998–2002 (Missing School Participation) (3)	(4)
Female	0.030*	0.034	0.018	–0.196**
	(0.012)	(0.063)	(0.013)	(0.057)
Age, 1998	0.016**	0.020**	–0.006**	–0.018*
	(0.002)	(0.007)	(0.001)	(0.008)
Missing Age Data	0.399**		–0.058**	
	(0.033)		(0.021)	
Female × Age, 1998	0.000	–0.001	–0.001	0.015**
	(0.001)	(0.005)	(0.001)	(0.004)
School in Budalangi Division	0.073**	0.065**	0.000	–0.007
	(0.018)	(0.018)	(0.013)	(0.015)

(continued)

(Appendix Table A3, continued)

	Missing Information on Parent Death in 2002		Attrited During 1998–2002 (Missing School Participation)	
Variable	(1)	(2)	(3)	(4)
Child Weight-for-Age (*z* score), 1998		–0.074**		0.030
		(0.025)		(0.028)
Child Had Malaria/Fever in Past Month, 1998		0.014†		–0.008
		(0.009)		(0.009)
Latrine at Home, 1998		–0.063		0.213**
		(0.054)		(0.048)
Cows at Home, 1998		–0.107†		0.007
		(0.057)		(0.052)
Goats at Home, 1998		–0.018*		–0.015
		(0.008)		(0.011)
Poultry at Home, 1998		–0.037*		–0.006
		(0.016)		(0.021)
Child Wears Shoes, 1998		0.033**		0.044*
		(0.012)		(0.021)
Child Wears School Uniform, 1998		–0.026†		0.046**
		(0.014)		(0.014)
Child Appears "Clean," 1998		–0.030		0.275**
		(0.054)		(0.049)
Cows at Home, 1998 × Age, 1998		0.007		–0.002
		(0.004)		(0.004)
Latrine at Home, 1998 × Age, 1998		0.005		–0.017**
		(0.004)		(0.005)
Child Appears "Clean," 1998 × Age, 1998		0.002		–0.020**
		(0.004)		(0.004)
Child Weight-for-Age (*z* score), 1998 × Age, 1998		0.007		–0.003†
		(0.002)		(0.002)
Became Orphan, 1999–2002			–0.059**	–0.054**
			(0.012)	(0.017)
Missing Information on Parent Death in 2002			0.168**	0.082**
			(0.009)	(0.013)
Observations	24,111	9,789	24,111	9,789
Mean (*SD*) of Dependent Variable	0.25	0.20	0.28	0.29
	(0.43)	(0.40)	(0.45)	(0.46)

Note: All regressions are probits, with marginal effects reported. Standard errors are clustered at the school level. Unreported controls include deworming program (PSDP) treatment variables, and region-year indicator variables. Regressions 1 and 3 include all 24,111 children from the baseline sample, and regressions 2 and 4 include all baseline children for whom baseline covariates are available. Age data are missing for 5,095 baseline children.

† $p < .10$; * $p < .05$; ** $p < .01$

REFERENCES

Adetunji, J. 2000. "Trends in Under-5 Mortality Rates and the HIV/AIDS Epidemic." *Bulletin of the World Health Organization* 78:1200–206.

Aigner, D.J. 1973. "Regression With a Binary Independent Variable Subject to Errors of Observation." *Journal of Econometrics* 1(1):49–50.

Ainsworth, M., K. Beegle, and G. Koda. 2005. "The Impact of Adult Mortality and Parental Deaths on Primary Schooling in North-Western Tanzania." *Journal of Development Studies* 41:412–39.

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

Ainsworth, M. and I. Semali. 1998. "Who is Most Likely to Die of AIDS? Socioeconomic Correlates of Adult Deaths in Kagera Region, Tanzania." Pp. 95–109 in *Confronting AIDS: Evidence From the Developing World,* edited by M. Ainsworth, L. Fransen, and M. Over. Brussels: European Union.

Case, A. and C. Ardington. 2006. "The Impact of Parental Death on School Outcomes: Longitudinal Data From South Africa." *Demography* 43:401–20.

Case, A., C. Paxson, and J. Ableidinger. 2004. "Orphans in Africa: Parental Death, Poverty, and School Enrollment." *Demography* 41:483–508.

Central Bureau of Statistics and Ministry of Health [Kenya]. 2004. *Kenya Demographic and Health Survey 2003*. Nairobi, Kenya: Central Bureau of Statistics. Available online at http://www.cbs.go.ke/downloads/pdf/Kenya_Demographic_and_Health_Survey_2003_Preliminary_Report.pdf

Evans, D. 2004. "The Spillover Impacts of Africa's Orphan Crisis." Unpublished manuscript. RAND Corporation, Santa Monica, CA

Evans, D. and E. Miguel. 2004. "Orphans and Schooling in Africa: A Longitudinal Analysis." BREAD Working Paper No. 56. Bureau for Research and Economic Analysis of Development, Cambridge, MA.

Filmer, D. and L. Pritchett. 2001. "Estimating Wealth Effects Without Expenditure Data—or Tears: An Application to Educational Enrollment in States of India." *Demography* 38:115–32.

Foster, G., R. Shakespeare, F. Chinemana, H. Jackson, S. Gregson, C. Marange, and S. Mashumba. 1995. "Orphans Prevalence and Extended Family Care in a Peri-urban Community in Zimbabwe." *AIDS Care* 7(1):3–18.

Foster, G. and J. Williamson. 2000. "A Review of Current Literature on the Impact of HIV/AIDS on Children in Sub-Saharan Africa." *AIDS* 14(Suppl. 3):S275–S284.

Gertler, P., D. Levine, and M. Ames. 2004. "Schooling and Parental Death." *Review of Economics and Statistics* 86(1):211–25.

Kamali, A., J.A. Seeley, A.J. Nunn, J.F. Kengeya-Kayondo, A. Ruberantwari, and D.W. Mulder. 1996. "The Orphans Problem: Experience of a Sub-Saharan Africa Rural Population in the AIDS Epidemic." *AIDS Care* 8:509–15.

Lloyd, C. and A. Blanc. 1996. "Children's Schooling in Sub-Saharan Africa: The Role of Fathers, Mothers, and Others." *Population and Development Review* 22:265–98.

Lundberg, M., and M. Over. 2000. "Transfers and Household Welfare in Kagera." Unpublished manuscript. World Bank, Washington, DC.

Manski, C. 1995. *Identification Problems in the Social Sciences*. Cambridge, MA: Harvard University Press.

Marum, L., J. Muttunga, F. Munene, and B. Cheluget. 2004. "HIV Prevalence and Associated Factors." Pp. 217–32 in *Kenya: Demographic and Health Survey 2003*. Nairobi, Kenya: Central Bureau of Statistics.

Miguel, E. and M.K. Gugerty. 2005. "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya." *Journal of Public Economics* 89:2325–68.

Miguel, E. and M. Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72(1):159–217.

Morgan, D., S.S. Malamba, J. Orem, B. Mayanja, M. Okongo, and J.A.G. Whitworth. 2000. "Survival by AIDS Defining Condition in Rural Uganda." *Sexually Transmitted Infections* 76:193–97.

Morgan, D. and J.A.G. Whitworth. 2001. "The Natural History of HIV-1 Infection in Africa." *Nature Medicine* 7(2):143–45.

Ntozi, J.P.M. 1997. "Effect of AIDS on Children: The Problem of Orphans in Uganda." *Health Transition Review* 7(Suppl.):23–40.

ORC Macro. DHS Statcompiler. 2005. Available online at http://www.measuredhs.com/statcompiler/start.cfm

Robinson, S. 1999. "Orphans of AIDS." *TIME* 154(24):60–61.

Ryder, R., M. Kamenga, M. Nkusu, V. Batter, and W. Heyward. 1994. "AIDS Orphans in Kinshasa, Zaire: Incidence and Socioeconomic Consequences." *AIDS* 8:673–79.

UNAIDS. 2000. *Socioeconomic Impact of HIV/AIDS in Africa*. Available online at http://www.unaids.org

————. 2002. *Report on the Global HIV/AIDS Epidemic*. Available online at http://www.unaids.org

UNAIDS, UNICEF, and USAID. 2002. *Children on the Brink 2002: A Joint Report on Orphan Estimates and Program Strategies*. Available online at http://www.unicef.org/publications/index_4378.html

————. 2004. *Children on the Brink 2004: A Joint Report of New Orphan Estimates and a Framework for Action*. Available online at http://www.unicef.org/publications/index_22212.html

United Nations Development Programme (UNDP). 2002. *Human Development Indicators 2002*. Available online at http://hdr.undp.org

Watson, C. 1992. "RAPing in Chad." Pp. 409–16 in *Rapid Assessment Procedures: Qualitative Methodologies for Planning and Evaluation of Health Related Programmes,* edited by N.S. Scrimshaw and G.R. Gleason. Boston: International Nutrition Foundation for Developing Countries.

Wehrwein, P. 2000. "AIDS Leaves Africa's Economic Future in Doubt." Available online at http://www.cnn.com/SPECIALS/2000/aids/stories/economic.impact/

World Bank. 2002. *Education and HIV/AIDS: A Window of Hope*. Washington, DC: The World Bank.

Yamano, T. and T.S. Jayne. 2004. "Measuring the Impact of Working-Age Adult Mortality on Small-Scale Farm Households in Kenya." *World Development* 32(1):91–119.

————. 2005. "Working-Age Adult Mortality and Primary School Attendance in Rural Kenya." *Economic Development and Cultural Change* 53:619–54.

Young, A. 2005. "The Gift of Dying: The Tragedy of AIDS and the Welfare of Future African Generations." *Quarterly Journal of Economics* 120:423–66.