Title

Orphans and Schooling in Africa: A Longitudinal Analysis

Authors
Evans, David
Miguel, Edward A.

Publication Date
2005-03-01

UNIVERSITY OF CALIFORNIA, BERKELEY

Department of Economics

Berkeley, California 94720-3880

CENTER FOR INTERNATIONAL AND DEVELOPMENT ECONOMICS RESEARCH
Working Paper No. **C05-143**

Orphans and Schooling in Africa: A Longitudinal Analysis

David Evans

Department of Economics,

Harvard University

Edward Miguel

Department of Economics,

University of California, Berkeley

March 2005

Key words: Parent death, education, HIV/AIDS, Africa

JEL Classification: I12, I20, J13, O12

* 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. 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 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, Caroline Hoxby, Larry Katz, Michael Kremer, and participants in seminars at Harvard, the World Bank, 2004 NBER Summer Institute, USC, and University of California Berkeley have provided valuable comments. Melissa Gonzalez-Brenes and Tina Green have provided excellent research assistance. This is a substantially revised version of BREAD Working Paper #56 (March 2004). All errors are our own. This paper can be found online at the UC eScholarship Digital Repository site: http://repositories.cdlib.org/iber/cider.

C05-143 Orphans and Schooling in Africa: A Longitudinal Analysis

Abstract

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 those with low baseline academic performance.

David Evans
Department of Economics
Harvard University
Cambridge, MA
dkevans@fas.harvard.edu

Edward Miguel
Department of Economics
University of California
Berkeley, CA
emiguel@econ.berkeley.edu

Orphans and Schooling in Africa: A Longitudinal Analysis

David Evans
Harvard University

Edward Miguel
University of California, Berkeley and NBER

March 2005

Abstract: 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 those with low baseline academic performance.

*** Acknowledgements:** 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. 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 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, Caroline Hoxby, Larry Katz, Michael Kremer, and participants in seminars at Harvard, the World Bank, 2004 NBER Summer Institute, USC, and University of California Berkeley have provided valuable comments. Melissa Gonzalez-Brenes and Tina Green have provided excellent research assistance. This is a substantially revised version of BREAD Working Paper #56 (March 2004). All errors are our own.
D. Evans, Department of Economics, Harvard University, dkevans@fas.harvard.edu.
E. Miguel (Corresponding author), 549 Evans Hall #3880, Department of Economics, University of California, Berkeley, CA 94720-3880, USA, emiguel@econ.berkeley.edu, phone 1-510-642-7162, fax 1-510-642-6615.

1. Introduction

More than one in nine children in Sub-Saharan Africa have lost a parent, and the HIV/AIDS pandemic is the leading cause (UNAIDS 2004).[1] HIV/AIDS deaths today could potentially have major long-run effects on economic development by affecting the human capital accumulation of the next generation. Yet 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, and 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 partially

[1] In this paper, as 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.

[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), and that African "economies are collapsing and famines are growing … Africa has seen poverty, but this will be worse than anything we have ever known" (Wax 2003). 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.

mitigated by strong traditional child fostering norms in Africa: for instance, Demographic and Health Survey data indicate that one fifth of households in Africa had taken in a foster child (whose parents were still alive, authors' calculation) during the 1990s, and this extensive informal fostering network might also benefit orphans.

This paper uses an unusual panel data set gathered over nearly five years to measure the impact of parent mortality on school participation in a high HIV prevalence district in rural western Kenya. This data allows us to estimate longer-term impacts than existing studies, as well as effects in the years preceding the parent death, due to AIDS-related morbidity, for instance. We focus on children who began the study period as non-orphans, and compare school participation patterns of those who subsequently lost a parent to those who did not. Baseline survey data on child and household characteristics allows us to test if parent death has differential impacts for children from particular types of households. Finally, the dataset contains information on local orphan rates, allowing us to estimate the extent to which community safety nets break down in areas with more orphans.

In the main result, we find a substantial and highly statistically significant negative impact of a parent death on primary school participation: school participation rates declined on average by 5 percentage points after a parent death. Thus informal social networks appear unable to fully insure households that suffer parent deaths in rural Kenya. Estimates in an alternative specification without individual fixed effects are considerably smaller, suggesting that estimates in existing studies that use cross-sectional data – simply comparing orphans to non-orphans at a single point in time – are biased and understate parent death impacts.

Impacts are more than twice as large for maternal deaths (9 percentage points) than paternal deaths (4 percentage points). Young girls under age 12 are significantly more likely to experience falling school participation following a parent death than other children, implying that parent deaths are likely to exacerbate large existing education gender gaps in Kenya. Children with lower baseline (pre-death) academic test scores experience significantly larger drops 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 death on more promising students.

Although these estimates indicate that informal social networks are inadequate to fully insure households against a shock as severe as parent death, there is no evidence that social networks have completely disintegrated due to HIV/AIDS, either. The proportion of children who have been orphaned in a community does not affect school participation, either for orphans or non-orphans.

In a further result, households with fewer assets (e.g., livestock, latrines) at baseline do not experience significantly different parent death effects than other households. Unfortunately, the data set does not contain detailed information on household income (due to research resource constraints, which prevented the collection of baseline consumption expenditures). Thus we cannot directly test whether the estimated schooling impacts are caused by households' inability to pay primary school fees or through several other channels, such as psychological trauma, the loss of parental emotional support, or due to lower perceived life expectancy for the surviving parent (or for the orphan herself). More generally, we lack sufficient data to convincingly establish which theoretical explanation is the main cause of the negative parent death impacts we estimate, and this is a key limitation of the paper. Still the finding that parent death effects are not systematically larger for children from asset-poor households, together with the result that maternal deaths lead to much larger school participation drops than paternal deaths, suggest that lower income is not the only factor driving school participation decisions after parent death, a result consistent with recent findings in Gertler, Levine, and Martinez (2004).[3]

We attempt to rule out the possibility that the estimated schooling impacts are spurious in several ways. First, the results are robust to numerous baseline household and community characteristics, and to individual fixed effects. Second, a range of baseline pre-death characteristics are nearly identical for the children who experience a parent death during the period (whom we call the "became orphan" group) and for those who do not (the "never orphan" group), and pre-death school participation levels and trends are also similar for the two groups, suggesting they are broadly comparable. Third, the results are the same when the baseline characteristics are interacted with time controls and these terms are included as additional explanatory variables, in an attempt to partially capture unobserved time-varying factors. Finally, the main findings are nearly unchanged when the "became orphan" group is compared to an alternative comparison group, those who began the sample period as orphans (the "always orphan" group).

It is worth stressing up front that we neither have individual biomedical information on HIV infection status, nor data on whether the cause of a parent death was AIDS: while HIV prevalence in the study region is estimated to lie between 20 and 30% (NASCOP 2001), and thus it is likely that many adult deaths are HIV/AIDS-related, we cannot determine the exact proportion in our sample. Since we lack data on the cause of parent death, we are unable to test whether AIDS orphans fare differently than other orphans, due to stigma attached to AIDS for instance. Nonetheless, UNAIDS et al. (2002) estimate that in Kenya as a whole, where the HIV

[3] In fact, Gertler, Levine, and Martinez (2004) suggest that the negative impact of parent death on schooling in Indonesia is almost entirely explained by causes other than the loss of income.

prevalence rate was estimated at 15% in 2001, 54% of orphans up to 14 years old had lost at least one parent to AIDS – and this proportion is likely to be even greater in our study region (Busia district), where prevalence is considerably higher.[4] Thus we feel confident in attributing the bulk of the parent deaths in our sample to HIV/AIDS.

A second important limitation of the analysis is the limited range of child outcomes: while we have exceptionally detailed schooling data, we lack information on child nutrition, health, and child labor. Nor do we have information on living arrangements and so cannot test whether the biological relationship between an orphan and her caregiver matters for outcomes (as argued by Case et al. 2004), or whether informal insurance against parent death is stronger when orphans have close adult relatives living nearby.

Yamano and Jayne (2005) is the most closely related work to the current study, but the current study represents several improvements. They use a panel data set of Kenyan household surveys, and a difference-in-differences identification strategy related to the one we employ in this paper. Yamano and Jayne find significant negative impacts of adult death on school enrollment, but only among poor children, which differs from our results. Yamano and Jayne also find effects prior to the adult death, as we do in this study, but our larger sample size and continuous data collection throughout the study period (the Yamano and Jayne data set has three observations, for 1997, 2000 and 2002) allows us to more precisely measure parent death and morbidity impacts. Another advantage of our approach is the use of school attendance and enrollment data collected at school by enumerators during unannounced visits, rather than

---

[4] Recent Kenyan government reports indicate that HIV prevalence may in fact be considerably lower – perhaps by half – in the country as a whole and in the study region, see Central Bureau of Statistics [Kenya] et al. (2004).

relying on parent (or caregiver) reports of school enrollment (Yamano and Jayne 2005 and other studies rely on the latter).[5] Finally, Yamano and Jayne estimate the impact of the death of any adults in the child's household on schooling since they lack precise parent death information, while we estimate parent death effects directly. In Kenya, where many households contain adults other than children's parents (Yamano and Jayne 2005, Table A1), they are thus estimating a different parameter. Gertler, Levine, and Ames (2004) estimate similar parent death impacts in Indonesia during the mid-1990s: a parent death during the previous twelve months leads to a doubling in the probability a child drops out of school that year. However, an important limitation is their inability to estimate impacts of parent death over periods longer than one year.

Finally, our results provide insight into the debate 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 independent of orphan status. The latter position – that orphans should not in general be targeted for assistance – has been advocated by many in the field (see Lundberg and Over 2000, Ainsworth and Filmer 2002) often drawing on the results of earlier studies that found relatively small parent death impacts. However, our results indicate that orphans are a particularly

---

[5] In a related paper, Yamano and Jayne (2004) find large negative impacts of the death of male household heads on the net value of household crop production – a massive drop of 68% – as well as in off-farm income, and speculate that an income shock of this magnitude would negatively impact school enrollment by making it difficult to pay school fees. For the death of a female household head, or of the female spouse of the male household head, the measured impact on crop production is smaller (at 46%) than for males.

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 died and young girl orphans experience particularly adverse schooling impacts and these easily observable characteristics could be used to improve targeting.

The remainder of the paper is structured as follows. Section 2 describes the data set and discusses sample attrition, section 3 outlines the estimation strategy, section 4 presents the main empirical results, and the final section concludes.

2. Data and Measurement

The data set was collected in Busia district, Kenya, a densely settled farming region adjacent to Lake Victoria, in the context of a primary school health program which provided medical treatment for intestinal worm infections (Miguel and Kremer 2004). The Kenyan non-governmental 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.

The foremost strength of the primary school data set we use, and what sets it apart from most other African data sets, is its length over five school years, from early 1998 to mid-2002. A second strength is that schools were visited by enumerators multiple times each year to record student school attendance and enrollment, and these visits were not announced to the school in advance. This results in a more compelling measure of school participation than either data collected from primary school attendance registers, which are thought to be unreliable in less developed countries, or parent responses about the school enrollment status of their children, which are used in virtually every other study. School participation often varies with the agricultural season and with immediate household economic needs, and thus parent reports on current enrollment status may be unreliable measures of actual school attendance patterns. 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 non-participants. Attempts were made to track the school participation of children who transferred to other schools within Busia district and neighboring Teso district, but data was not collected for those who left these two districts.

The second dataset 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 1998 baseline sample. If the child was present at school, she was asked directly about the mortality of her parents, and about the exact year of the death if a parent had died. If the child was not present at school that day, teachers and other students in the school were asked to provide this information. In practice, it was common for siblings, cousins, and neighbors of absent children to volunteer the information on parent death, as this information appears to be quite widely known in rural Kenyan communities. As a check on data reliability, the parent mortality data collected at school was compared to mortality data collected at children's homes in 2001 for a representative sub-sample of 69 of the children (among those children who had experienced a parent death and for whom we already had home contact information). These home surveys were then typically collected from intimate relatives of the dead parent. There is a high correlation in the reported year of father death between the two surveys (correlation coefficient 0.87), but the correlation for mother deaths is considerably lower (at 0.61), although since there are many fewer mother deaths than father deaths, this latter figure is based on relatively few observations. The reported year is identical in the two surveys (at school and at home) for 70% of father deaths in this small sub-sample, and in the remaining cases it never differs by more than two years.

The third 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 information on baseline 1998 academic test score performance for a slightly smaller subset of children initially in those same grades. Finally, we employ the 2002 Headmaster Questionnaire, which collected information from primary school headmasters during May to July 2002 regarding policies toward orphans in a subsample of 48 of the 75 program schools.

2.1 Sample Size and Attrition

The scarcity of panel data sets in Sub-Saharan Africa is, in part, due to difficulties in tracking survey respondents through time, and we are not immune to this problem. Migration, child fostering, and imperfect recall all complicate our task and lead to non-trivial rates of missing data, especially on the precise year of parent death. We conduct simulations to bound the extent of bias due to missing data below (section 4.2).

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. 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 reliable mortality data for both parents.[6] Most cases of unknown or unreliable orphan status were among children initially in the upper grades in 1998. They had been out of primary school longer than younger pupils and thus were often not as well-known to other children in the school during the tracking survey. Individuals with missing or unreliable parent death information tend to have worse school participation outcomes than other students (as discussed below).

The restricted sample contains 7,815 children from the full sample for whom baseline individual and household measures from the 1998 Pupil Questionnaire data set 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, since the 1998 pupil questionnaire was only administered in grades 3 and higher. Of the remaining 11,415 students, there is baseline survey information for only a subset of students in grades 3 to 7, those who were present on the day of that survey, leaving 7,815 children in the restricted sample. For 2,194 of the restricted sample students there is school participation data for only a subset of the five years we study, since in some cases students moved away from the area and there is no information regarding subsequent schooling.

[6] Age data are also missing for 3,163 children in the full sample. It is common for individuals not to know their birth year in rural Kenya, where 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.

For both the full and restricted sample, such children are included in the analysis only in the years that we observe them.[7]

Child characteristics are similar for both the full and restricted samples (Table 1). A relatively large proportion (8%) of children who were non-orphans in 1998 became orphans during 1999-2002, and most of those experienced a paternal death. Fully 15% of schoolchildren in this area were orphans at baseline in 1998, and this proportion varies widely across the 75 communities, from near zero in some areas up to 41% in others. Children in this region are quite poor even by Kenyan standards, and this translates into poor health and nutrition (Table 1, Panel B): nearly 20% of households lack a latrine (or toilet) at home, only 14% of children wore shoes to school in 1998, and almost two-fifths reported experiencing a fever in the month previous to the survey.[8] 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).

We next test whether students with missing data are significantly different from other students along observable dimensions. There are two sources of missing observations that reduce our "baseline sample" of 24,111 children, missing parent death information and missing school participation data. In the first case, the dependent variable is an indicator that takes on a value of

---

[7] Below 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, namely, those with parent mortality data.

[8] In practice, this includes disease episodes classified by children as either "fever" or "malaria" in the survey. Although poor children were asked to report recent cases of malaria (as opposed to fevers), fevers of any cause are often reported as "malaria" in areas like Busia where testing is costly (Watson 1992).

one if the child is missing parent death information. The dependent variable has a mean of 0.25, i.e., 25% of the sample lacks reliable parent death information for one or both parents. In the baseline sample of 24,111 children, the older children and, unsurprisingly, those with missing age data are more likely to have missing orphan status information (Table 2, regression 1). Among those with detailed baseline characteristics from the 1998 Pupil questionnaire, most indicators of household asset ownership are negatively related to missing orphan status information, including ownership of cattle, goats, and poultry, suggesting that poorer households are more likely to be lost from the sample (regression 2). Children wearing shoes in the baseline survey were significantly more likely to have missing data, while the opposite holds for children wearing a uniform, although the explanation for this pattern is unclear.

Considering attrition – namely, missing school participation data – as the dependent variable, children who we know become orphans during the period are significantly less likely to have missing schooling data, among the 18,133 children with parent mortality data (Table 2, regression 3).[9] This result appears counter-intuitive 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 switches signs and becomes positive; for instance, if a randomly chosen subset of just 12% of the children whose orphan status is unknown (far less than the actual proportion of orphans in this population – see Table 1) are assigned to be orphans in a simulation, the coefficient estimate on the orphan term

[9] We find that 28% of children attrit from the sample at some point during 1998-2002, although often temporarily. All full sample and restricted sample children have at least one school participation observation in 1998.

in regressions 3 and 4 becomes positive (results not shown). This result is similar when examining attrition in both the full sample (regression 3) and the restricted sample (regression 4). The attrition effects are significantly larger for maternal deaths than for paternal deaths (regressions not shown). This is consistent with Evans' (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 after the death, exiting the sample. In the restricted sample (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.

2.2 Primary School Fees and Orphans

The primary school finance context in rural Kenya during 1998-2002 is also important in understanding households' school participation decisions.[10] Both the central government and local school committees played important roles in Kenyan primary school finance. The national Ministry of Education paid teacher salaries, while school committees raised funds locally for books, chalk, classrooms, and desks. Parents raised the bulk of local school funds through two mechanisms: school fees and village fundraising events (called *harambees* in Swahili). Annual school fees were set by the school committee – each primary school is managed by its own committee – and collected by the headmaster. School fees ranged roughly from 4-10 U.S. dollars per family during 1998-2002, a non-trivial amount in this area. A variety of informal sanctions

[10] This section describes Kenyan primary school finance before Mwai Kibaki was elected president in December 2002. Starting in early 2003, the Kenya Ministry of Education abolished local primary school fees nationwide and agreed to provide some additional resources to primary schools to compensate for lost local funds.

could be employed against parents if they failed to make school fee and *harambee* payments, including publicly naming them at community meetings, and temporarily suspending their children from school (Miguel and Gugerty 2004).

While the threat of such sanctions was an important tool used by headmasters and school committees to enforce payment, the children of non-contributing parents could not permanently be removed from school: former President Daniel Arap Moi of Kenya repeatedly stated that no child could be refused a primary education because of nonpayment of fees, and while official Ministry of Education policy was unclear during 1998-2002, in practice these decrees limited the discretion of schools to expel students. The content of threatened sanctions therefore did not include complete exclusion from public education, but only temporary suspension. This likely dampened the effect of household income differences on child primary school enrollment.

Few primary schools in this area made special allowances for orphans in terms of school fee reductions, according to the 2002 Headmaster Questionnaire. Forty-two of the 48 (88%) surveyed headmasters stated that orphans were subject to exactly the same school fees as other children. Of the 38 headmasters admitting they had sent some students temporarily away for nonpayment of school fees in the previous year, 32 of 38 (84%) claimed that orphans had been sent away just as often as non-orphans. More than two-thirds of the headmasters stated they believed that a main cause of dropping out for orphans was nonpayment of school fees or not owning a school uniform (another large financial cost of schooling in Kenya). Thus the inability to pay school fees appears to be a plausible cause for at least part of the drop in school participation after a parent death, to the extent that parent death reduces household disposable income (as Yamano and Jayne 2004 find) and in the presence of credit constraints.

Yet there remain many possible channels linking parent death to schooling other than income, including changes in the quality of parent (or other caregiver) emotional support, psychological trauma resulting from the death, and disruptions caused by fostering, as discussed further below. Although we are unable to definitively determine the key causal mechanism underlying the reduced form parent death impacts we estimate, there is suggestive evidence that factors other than income play an important role.

3. Empirical strategy

3.1 Identification Issues in Existing Research

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. Most existing studies estimate 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 importantly, the current child and household characteristics used as controls may have themselves been affected by the death. Moreover, since parent death is relatively rare in most populations, few studies have sufficient statistical precision to reliably estimate moderate impacts.

Case et al. (2004) employ nineteen Demographic and Health Surveys (DHS) collected across ten Sub-Saharan African countries between 1992 and 2000 to estimate the impact of parent death on school enrollment. The relatively high incidence of parent death in their sample allow them to estimate the impact of parent death with high statistical precision, and the use of data from many different countries means their results are likely to generalize within Africa. They use a clever household fixed effects estimation strategy which compares orphans and non-orphans in the households that take in orphans, thus accounting for any fixed household level characteristics. Their main finding is that orphans are significantly less likely to be enrolled in school than non-orphans, even among children within the same household. Nevertheless, despite their impressive data effort and innovative estimation strategy, the study may suffer from the omitted variable bias problem mentioned above, since only use cross-sections are available in their dataset, rather than a panel, and thus they cannot account for any fixed characteristics of the orphan child herself or of her original household (which is not observed in the data).[11]

Earlier studies do not find substantial negative parent death impacts on child education. For instance, Ainsworth et al. (2002) analyze a panel of 1,213 children in northwestern Tanzania and find minimal impacts of parent death on schooling. In particular, child school enrollment is unaffected by parent death among non-poor households, while for poor households, they find that enrollment is merely delayed for the youngest children but basically unaffected for older children. There are larger adverse effects for maternal orphans than paternal orphans in their data. Note that, although Ainsworth et al. (2002) control for baseline household characteristics, they do not use child fixed effects. Several studies echo Ainsworth et al. (2002) in finding little or no difference between orphans and non-orphans in terms of school enrollment (see Kamali et

[11] 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 off assets, becoming poorer. It is methodologically preferable to measure characteristics prior to the parent death. The data in Case et al. (2004) also give no indication of how long a child has been an orphan and thus cannot shed light on how they fare over time.

al. 1996, Ryder et al. 1994, Lloyd and Blanc 1996), although these all rely 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 girl orphans suffering more than boys (World Bank 2002, UNAIDS 2002).

The absence of consistently negative impacts on African children in existing work has 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). One explanation for differences across geographic settings is the possibility that these insurance networks weaken or break down when local orphan rates surpass a certain critical level. However, note that the large estimated effects in Indonesia (Gertler, Levine, and Ames 2004) and smaller effects in Tanzania (Ainsworth et al. 2002) do not seem to fit this interpretation, given the much higher rate of orphanhood in Tanzania.

An alternative explanation for the small estimated orphan effects in cross-sectional studies is the possibility that African HIV/AIDS victims are often of somewhat higher socioeconomic status than non-victims, at least early in the epidemic; 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 strong positive correlation between socioeconomic measures and HIV prevalence has been found in several existing African studies (e.g., Ainsworth and Semali 1998) and also holds in the 2003 Kenya DHS data (Central Bureau of Statistics [Kenya] et al. 2004: 223). To the extent that socio-economic variation is at least partially unobserved by the econometrician, this would lead to an upward bias in the estimated "impact" of being an orphan on subsequent life outcomes in cross-sectional studies, and could obscure negative parent death impacts. This is less of a concern in longitudinal studies, like the current study, where fixed levels differences across households can be controlled for in the analysis.

Finally, Breierova (2002) uses an instrumental variable method to estimate the impact of AIDS on school enrollment in western Kenya, relying on differences in circumcision practices between members of the Luo ethnic community and other groups, and estimates drops of 3.5-5.6% in enrollment due to AIDS. This approach constitutes an improvement over most existing cross-sectional studies, although note that the exclusion restriction may fail with this method to the extent that underlying trends in Luo school enrollment would have differed from other groups during her study period even in the absence of HIV/AIDS.

3.2 Estimation Approach

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. Average annual school participation takes on a range of values between zero and one, and this is the fraction of unannounced enumerator visits for which the child was present at school. The main estimation approach in the current paper is linear regression with child fixed effects, where the "events" of interest are parent deaths. The child fixed effect captures time-invariant child (and household) characteristics that affect school participation. In some specifications we examine effects on school enrollment – an indicator variable that takes a value of one for students who were present in school during at least one enumerator visit during the course of the school 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, then equation 1 yields unbiased estimates of the effect of parent death on child schooling. (We discuss this assumption in greater detail below.) In some specifications, baseline household characteristics are included as explanatory variables, rather than the child fixed effects, and these generate somewhat weaker results. Disturbance terms are allowed to be correlated within the same school. This leads to equation 1 (where the "1" subscript refers to the equation number):

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

Y_{ijt} is the school participation rate for student *i* in school *j* during year *t*, α_{ij} is a student fixed effect, and ρ_{jt} 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 there are a number of limitations to this approach. Such a specification imposes a constant effect of parent death on subsequent child outcomes regardless of when the parent died. It is possible that the effects of parent death might either compound over time (i.e., children whose parents died long ago experience increasingly adverse outcomes) or perhaps diminish if coping mechanisms emerge over time. To allow for such effects, we include indicator variables $\sum_{\tau} \beta^S 1(\tau = S)_{ijt}$ in some specifications, where τ is the number of years since the parent death; τ also takes on negative values in years before the parent death, for instance due to AIDS-related morbidity.

Medical researchers estimate that AIDS deaths in nearby rural 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 expect negative effects up to two 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 parent death).

We do not observe children four or more years after a parent death, since the main analysis is restricted to children who were non-orphans at baseline in 1998 and we only observe the children through 2002.

In order to account for cohort and year-specific trends and 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, e.g., 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 was found to be related to school participation (Miguel and Kremer 2004).

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

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

Parent death may theoretically have differential effects based on the gender of the parent; for instance, to the extent that mothers' income and care-giving are more important than fathers' income and care-giving, or maternal deaths have different implications for subsequent fostering patterns, maternal deaths will have a greater impact on child outcomes (or vice versa). To estimate differential effects based on the gender of the parent, we include separate indicators for maternal death and paternal death, and also explore the impact of the first parent death versus the second parent death. The parent death indicators are also interacted with household and community characteristics to test whether individuals from particular types of households or communities are differentially affected by parent death. For example, the magnitude of the parent death effect may depend on child age since older children are better substitutes for parents in the labor market, perhaps making them more likely to drop out of school after an adverse household income shock (although we find below that this is not the case in our setting).

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 or 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 this rural area where most adults engage in subsistence agriculture few have formal sector jobs to lose. Note that child morbidity due to HIV infections contracted from parents is unlikely to affect the estimation, since the overwhelming majority of children born HIV-positive in rural Africa die before reaching school age (Adetunji 2000).

To begin addressing this issue, we restrict attention to individuals whose parents were both alive at baseline in January 1998, and compare individuals whose parents subsequently died during the period 1999-2002 to those whose parents did not die. We make the case that that these two groups of individuals – the "became orphan" and "never orphan" groups – are comparable along a range of observable characteristics at baseline. There are no significant differences in terms of baseline school participation or demographic characteristics in the full sample (Table 3, Panel A). In the restricted sample, the two groups look remarkably similar along fourteen baseline 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 academic test scores (with an average difference of only 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 the sample schools – there is no significant difference between the 1997-98 trends in school participation for children who later become orphans and those who do not (Panel C), evidence that they were similar in terms of both school participation levels and trends before the parent deaths.

These arguments do not completely eliminate concerns about the suitability of the comparison group, and in the absence of a natural experiment, it may be impossible to do so. Yet we feel that this allays most reasonable concerns about the comparability of the two groups: if "became orphan" and "never orphan" households indeed differed sharply along unobserved dimensions – for instance, parents' commitment to their children's education, or their intertemporal discount rate – it is likely that these differences would also be reflected along observable dimensions given the rich set of baseline characteristics we employ, but we do not find systematic differences.

In a further attempt to address unobserved time-varying factors, in some specifications we include all of our standard baseline socioeconomic, educational, and demographic characteristics – including those that differed significantly across the became orphan and never orphan groups, e.g., child cleanliness, age, and baseline 1998 test score – interacted with a full set of year indicator variables, and find that the main empirical results are unchanged, as discussed below. As an additional robustness check, below we compare the "became orphan" children to the "always orphans", the 2,676 children who began the study period as orphans. Although they are not as compelling a comparison group as the "never orphans", at least in terms of matching up along baseline characteristics (see Appendix Table A1), this specification yields nearly identical parent death estimates.

Several possible sources of bias remain, but most tend to bias the parent death estimates toward zero, leading our estimates to serve as bounds on actual impacts. In other words, parent death impacts are likely to be even larger than our estimates. First, note that children with unknown orphan status have lower baseline school participation and less household asset ownership than other children (Table 3, right column). To the extent orphans are more likely to drop out of school and leave the area, it is likely that these children are also disproportionately orphans, in which case excluding them from the analysis would likely lead us to underestimate actual parent death impacts.

Second, we are likely to further underestimate parent death impacts if the survey method captures information on parent death years with error (Aigner 1973).[12] 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 due to the parent's AIDS-related morbidity, and this would reduce the average school participation rate of children classified as "never orphans". Yet this bias is likely to be very small: only 2% of baseline non-orphans become an orphan each year during 1999-2002 (Table 1), and the pre-death effects we estimate are moderate, so the resulting product of these two quantities is likely to be negligible.

---

[12] The authors conducted a simulation in which the year of parent death is replaced with a "noisy" proxy, where the distribution of noise is the difference between the year of parent death recorded in the home tracking survey versus in the school tracking survey. In regression specifications like those in Table 4, the main parent death impacts are attenuated only slightly in the simulation (results not shown).

We also test whether children's school participation fares worse in areas where there have already been more parent deaths, and in particular whether orphans fare disproportionately worse in such areas. Parent deaths limit child fostering options for others in their social network and community, leading to adverse spillovers. In practice, we estimate externality impacts using the share of children in the school community that have experienced a parent death. Local orphan rates evolve slowly over time, and thus the panel dataset we use contains limited intertemporal variation in community orphan rates once child fixed effects and geographic variables are included as controls, unfortunately limiting statistical precision despite the wide variation in baseline orphan rates.

This study has a number of other limitations worth noting. The estimation approach does not permit us to estimate broad regional or national effects of the HIV/AIDS epidemic on primary school participation due to, for example, reduced national school funding, teacher shortages, or falling demand for education (which is theoretically possible in a society where 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, as these children may never enroll in primary school in the first place, and thus are not in our dataset. As with any microeconomic empirical study, questions of generalizeability remain important since the impact of parent death could differ across settings – in rural versus urban areas for instance, or as a function of local school fees – an issue we cannot address in this study's entirely rural Kenyan sample.

4. Empirical Results

4.1 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 afterwards (Table 4, 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 parental morbidity, and the timing coincides with the period of AIDS-related morbidity described in the existing literature (Morgan et al. 2000, Morgan and Whitmore 2001). There is no evidence of orphan recovery after parent death, in contrast to some earlier work (Ainsworth et al. 2002), suggesting that long-run parent death effects 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-value=0.56), nor is the equality of effects in the two years immediately pre-death (p-value=0.86). Figure 1 graphically presents the school participation time pattern for the full sample using point estimates from regression 1.

We combine two years before parent death to estimate pre-death morbidity effects and combine the post death years in most subsequent specifications, as in equation 2 above. In this specification, parent death has a moderate negative impact on child school participation: on average, school participation falls by 0.055 (standard error 0.017, statistically significant at 99% confidence) after parent death in a specification with individual fixed effects[13], and the average effect in the two years before the parent death is again negative although smaller and not statistically significant (-0.021, standard error 0.015 – Table 4, regression 2). In a specification

[13] The analogous estimate in Yamano and Jayne (2005) for a specification including child and household controls is similar, at -0.060 (refer to their Table 4, regression A).

without child fixed effects, similar to many cross-sectional specifications found in existing literature, the analogous point estimates are smaller in magnitude at -0.040 (standard error 0.007) for the post-death effect and -0.018 (standard error 0.007 – regression not shown) for the pre-death effect. This implies that omitted variable bias is positive and would lead researchers to understate parent death impacts if child fixed effects were not included.

The time pattern of effects is similar with the smaller restricted sample (Table 4, regression 3), and the parent death effect is nearly identical to the effect in the full sample, with an estimate of -0.054 (standard error 0.022, significant at 95% confidence – regression 4) while the pre-death effect is somewhat larger at -0.032 (standard error 0.019, significant at 90% confidence). All baseline covariates (shown in Table 4) interacted with the year indicator variables are next included to partially address concerns related to time-varying omitted variables, and results are nearly identical (parent death effect -0.053, standard error 0.022, and pre-death effect -0.031, standard error 0.019 – regression 5). When the baseline 1998 test score is interacted with the year controls, as well, the sample falls slightly to 28,665 observations but the coefficient estimates remain almost unchanged at -0.053 (standard error 0.022, post-death effect) and -0.029 (standard error 0.019, pre-death effect – regression not shown).

Both parent death and pre-death effects remain statistically significant when a range of baseline controls are included instead of fixed effects, but point estimates are somewhat smaller in magnitude (Table 4, regression 6), again suggesting that fixed effects address omitted variable bias. The parent death impact is larger than the estimated effects of various socioeconomic proxies in this specification, including livestock ownership and wearing a uniform to school.

Effects are robust to an alternative schooling measure, the school enrollment indicator variable, yielding estimated magnitudes similar to Gertler, Levine, and Ames (2004) in Indonesia

– more than a doubling of the drop-out rate after parent death – with time patterns similar to school participation (Appendix Table A2).

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 (standard error 0.013, statistically significant at 99% confidence – Table 5, regression 2), and the pre-death effect is again smaller at -0.008 (standard error 0.012). When the "always orphan" and "never orphan" groups are both included in the regression as the comparison group for the "became orphan" group, the parent death effect remains stable and significant at 99% confidence at -0.045 (standard error 0.014, regression 3).

A second outcome measure is child mortality, but we do not find a statistically significant impact of parent death on child mortality during the sample period: in a specification analogous to Table 4, regression 2, the point estimate on parent death is just -0.001 (regression not shown).

<u>4.2 Bounding Bias due to Missing Data</u>

We place bounds on the impact of parent death on school participation in the full sample accounting for both missing parent death information and schooling data, and this 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 on schooling, so we feel confident in asserting that parent death has a negative impact on school participation in rural Kenya.

This exercise requires assumptions on three groups: (A) children who are missing school participation data only (i.e., attrition), (B) children who are missing orphan status data only, and (C) children who are missing both. To establish an upper bound on the magnitude of impacts, for group A children we assume school participation is equal to zero if they become orphans (with zero participation starting two years before the parent death), and one if they are not orphans. We assume group B children become orphans if their school participation is lower than the mean school participation rate among the full sample; otherwise we assume they do not become orphans. Finally, we assume that all children in group C become orphans and that their school participation is equal to zero (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 parent mortality who are assigned as "became orphans".

In a fixed effects specification (analogous to Table 4, regression 2), and with 100 runs of the simulation, the average upper bound is a -0.18 decrease in school participation after parent death, and a -0.08 decline in the two years pre-death. Since 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 parent death effect 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: in group A, school participation is assumed to be one if children become orphans (with the perfect school participation starting two years before the parent death) and zero for non-orphans. Group B children are assumed to become orphans if their school participation is above average, and zero 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 zero school participation. This yields a mean lower bound of a 0.00 effect post parent death (and 0.01 increase pre-death), again in 100 runs of the simulation.

4.3 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 (standard error 0.022, statistically significant at 99% confidence), and the post-death effect is -0.093 (standard error 0.025 –Table 6, regression 1). The analogous effects for fathers are less than half as large, with the father death effect at -0.036 (standard error 0.022, not statistically significant at traditional confidence levels) and a pre-death effect of only -0.005 (standard error 0.018). The difference between pre-death maternal and paternal effects is statistically significant at 95% confidence (p-value=0.03) and the post-death effects at 90% confidence (p-value=0.09).[14]

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 disruption of fostering may also account for part of the large estimated maternal death effect: Evans (2004) finds that child fostering patterns

[14] Note that this result differs from an earlier version (BREAD Working Paper #56), where we found no statistically significant difference between maternal and paternal death impacts. The explanation for this discrepancy, and for most other differences between the two versions, is the different econometric specification we employed in the earlier version, namely one that failed to include pre-death terms. Failing to account for the substantial drops in the two years before a parent death led us to estimate considerably smaller parent death impacts.

differ significantly depending on the gender of the parent that dies, with orphans significantly more likely to be sent to live in other households following maternal deaths than paternal deaths. These different fostering patterns by parent gender allow the possibility that sample attrition is driving some of the estimated difference between maternal and paternal death impacts, but we feel that differential sample attrition – if anything – is likely to lead us to understate maternal death impacts, as low performing orphans are more likely to leave the sample.

The additional impact of becoming a double orphan, on top of the summed effects of losing both mother and father, on school participation is near zero and not statistically significant (Table 6, regression 2), although there is limited statistical precision on this coefficient estimate as a result of the relatively small number of double orphans in the sample (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 7, regression 1).[15] The explanation may lie in the underlying academic ability of enrolled primary school students of different ages. Given high dropout rates during primary school in Kenya, those students still in school during their teenage years are positively selected on academic

[15] Regressions 1 and 2 in Table 7 also include an indicator variable for "Missing age data", and interactions between the "Missing age data" indicator and the parent death indicators (coefficients not reported).

ability, thus this result that older children are less likely to drop out following a parent death – despite the higher opportunity costs for older children, since labor market prospects are better for them – 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 below. The gender of the parent who dies does not differentially affect young versus older students (regression 2).

Girls are no more likely than boys overall to experience falling school participation following a parent death (Table 7, regression 3), and this holds independent of the gender of the parent who dies (regression 4). However, the double interaction specification suggests that the group most likely to experience falling school participation following a parent death are young girls (statistically significant at 90% confidence, regression 5). Both young and older boys experience average drops in school participation of approximately 5 percentage points following a parent death (summing the relevant coefficient estimates in regression 5), older girls experience a drop of 2 percentage points, while for young girls under age 12 the coefficient estimate is a massive 12 percentage points. Mother versus father 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 examining triple interactions of this sort (regression not shown).

One possible interpretation is that the perceived returns to primary schooling are lower for young girls than for young boys in these households, but there is no evidence that actual returns to education differ by gender in Sub-Saharan Africa (Schultz 2003). Another explanation is simple discrimination against girls, and in particular the possibility that young girls are systematically called upon to work after a parent death by custom (despite their low labor productivity relative to older children) but this remains speculative.

The likelihood that an orphan is removed from school should in part depend on the child's expected returns to continued primary schooling, which is likely to be an increasing function of her academic ability in this context since only the best students are typically able to continue on to secondary school. We employ 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 in the right tail of the baseline exam distribution are largely unaffected. For a child with a test score of zero at baseline (the mean score by construction), the pre-death effect is -0.029 (standard error 0.019 – Table 8, regression 1) and the parent death effect is -0.053 (standard error 0.021, statistically significant at 95% confidence), while the analogous pre- and post-death impacts for a child with a baseline test score of +1 standard deviation are essentially zero (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. Although impacts are roughly linear in the initial score, negative parent death impacts appear especially large for those in the lowest quintile of the baseline test distribution (regression not shown).

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 goes in the expected negative direction, but is not statistically significant (Table 8, regression 2), and the interaction terms with a poverty index similar to that used in existing studies[16] are also not

[16] 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, following Filmer and Pritchett (2001). Unfortunately, we lack detailed information on parent occupation.

statistically significant (regression 3). These findings suggest that moderate amounts of household wealth do not significantly buffer children from the shock of parent death, at least at the low asset levels found in our sample (although note that these are rough socioeconomic proxies, and so results should be interpreted with caution). The baseline test score interaction terms remain nearly unchanged when the socioeconomic controls and interactions are included (regression 4), and when gender interactions are included (regression not shown), suggesting that the test score is not simply proxying for household socioeconomic status or gender.

4.4 Community Impacts

Orphans do not fare significantly worse in primary school communities with higher current orphan rates, although limited statistical precision means we cannot rule out moderate negative impacts. The point estimate is reasonably large and in the expected direction but is not statistically significant (point estimate on the interaction term -0.232, standard error 0.230 – Table 9, regression 1). The result that the local orphan rate is not strongly associated with orphan schooling is robust to an alternative definition of overall local orphan burden (regression 2), and to examining local maternal and paternal orphanhood separately (regression 3).[17] Results are similar using initial 1998 orphan rates rather than contemporaneous rates by year (regression not shown).

Any negative spillovers of higher orphan populations on other community members are probably small. Non-orphans do fare somewhat worse on average in communities with a higher

[17] In contrast, Yamano and Jayne (2005) find that average school attendance is strongly negatively correlated with lagged provincial HIV-prevalence in their sample (note that their measure is at a much higher level of aggregation).

proportion of orphans (coefficient estimate -0.310, standard error 0.335 – Table 9, regression 1). This point estimate suggests that an increase in the proportion of orphans in the community from 10% to 20% reduces non-orphan school participation by 3.1 percentage points, a moderate effect, but this is also not statistically significant.

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 and should be re-examined. Further research is needed to understand how general these findings are beyond rural western Kenya, but the fact that there is little evidence networks are breaking down in this region, with its high orphan rates and high HIV prevalence, suggests that this issue is even less of a concern where prevalence is lower. Note that these patterns resonate with Evans' (2004) finding that the adverse effects on schooling experienced by non-orphan children when an orphan moves into their household are small at worst.

Finally, there is no evidence parent death impacts are related to the local ethnic composition of orphans, the match between orphan ethnic affiliation and local ethnic composition, or to the average socioeconomic status of the community as a whole (regressions not shown).

5. Conclusion

To summarize the main finding, parent death is associated with a sharp drop in primary school participation in rural western Kenya, with particularly adverse impacts among maternal orphans, young girls, and children who are weak students. Our empirical approach addresses a number of methodological shortcomings of recent studies, in particular the omitted variable bias toward zero in existing work that relies on cross-sectional data, and highlights the usefulness of collecting panel data sets on children's health and education in less developed countries.

The results also shed light on the optimal targeting of assistance programs to mitigate the impact of HIV/AIDS on education in rural Africa, at least in the sense of identifying groups that experience particularly adverse outcomes following parent deaths. Still, we feel 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 confidently develop further policy recommendations in this area.

In future work, we plan to extend the analysis in this paper for several additional years using the Kenya Life Panel Survey, which is currently collecting labor market, education, health and demographic outcomes on a subsample of these children, in order to more accurately estimate long-term impacts of parent death on a broader range of life outcomes. Further empirical research on the impact of parent HIV/AIDS deaths on children in other African settings, as well as on the design of programs to assist orphans is urgently needed given the rapidly growing numbers of orphans across Sub-Saharan Africa.

References

Adetunji, Jacob. (2000). "Trends in under-5 mortality rates and the HIV/AIDS epidemic", *Bulletin of the World Health Organization*, 78, 1200-1206.

Aigner, Dennis J., "Regression with a Binary Independent Variable Subject to Errors of Observation," *Journal of Econometrics*, 1973, 1(1), 49-50.

Ainsworth, Martha, Kathleen Beegle and Godlike Koda. "The impact of adult mortality on primary school enrolment in Northwestern Tanzania." Africa Region Human Development Working Paper, World Bank, 2002.

Ainsworth, Martha, and Deon Filmer. "Poverty, AIDS, and Children's Schooling: A Targeting Dilemma" World Bank Policy Research Working Paper #2885.

Ainsworth, Martha, and I. 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.

Breirova, Lucia. "AIDS and Schooling in Kenya", unpublished working paper, MIT, 2002.

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

Central Bureau of Statistics [Kenya], et al. (2004). *Kenya Demographic and Health Survey 2003*. Calverton, MD.

Evans, David. "The Distribution and Impact of Africa's Orphan Crisis", unpublished working paper, Harvard University, 2004.

Filmer, Deon, and Lant Pritchett. "Estimating wealth effects without expenditure data – or tears:

an application to educational enrollment in states of India." *Demography,* 2001, 38, 115-32.

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

Foster, Geoff, and John Williamson. "A review of current literature on the impact of HIV/AIDS on children in Sub-Saharan Africa." *AIDS*, 2000, 14 (suppl 3), S275-S284.

Gertler, Paul, David Levine and Minnie Ames. "Schooling and Parental Death", *Review of Economics and Statistics*, 86(1), February 2004.

Gertler, Paul, David Levine and Sebastian Martinez. "The presence and presents of parents: do parents matter for more than their money?" Unpublished manuscript, U.C. Berkeley, 2004.

Kamali, A., J. A. Seeley, A. J. Nunn, J. F. Kengeya-Kayondo, A. Ruberantwari, and D. W. Mulder. "The orphans problem: experience of a Sub-Saharan Africa rural population in the AIDS epidemic." *AIDS Care*, 1996, 8(5), 509-515.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. (2004). "Incentives to Learn", unpublished working paper.

Lloyd, Cynthia, and Ann Blanc. "Children's schooling in Sub-Saharan Africa: the role of fathers, mothers, and others." *Population and Development Review,* 1996, 22(2), 265-298.

Lundberg, Mattias, and Mead Over. "Transfers and Household Welfare in Kagera", unpublished, 2000.

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

Miguel, Edward, and Mary Kay Gugerty. "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya", forthcoming *Journal of Public Economics*, 2004.

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

Morgan, Dilys, S. S. Malamba, J. Orem, B. Mayanja, M. Okongo, and J.A.G. Whitworth. "Survival by AIDS defining condition in rural Uganda," *Sexually Transmitted Infections*, 2000, 76, 193-97.

Morgan, Dilys, and Jimmy A.G. Whitworth. "The natural history of HIV-1 infection in Africa," *Nature Medicine*, February 2001, 7(2), 143-145.

NASCOP (National AIDS and STD's Control Programme). *Estimating National HIV Prevalence in Kenya from Sentinel Surveillance Data*, Nairobi, Kenya, June 2001.

Ntozi, James P.M. "Effect of AIDS on Children: The Problem of Orphans in Uganda", *Health Transition Review*, 1997, supplement to volume 7, 23-40.

Robinson, Simon. "Orphans of AIDS," *TIME*, December 13, 1999.

Ryder, Robert, Munkolenkole Kamenga, Muniaka Nkusu, Veronique Batter, and William Heyward. "AIDS orphans in Kinshasa, Zaire: incidence and socioeconomic consequences," *AIDS*, December 13, 1999, 8, 673-679.

Schultz, T. Paul. "Evidence of Returns to Schooling in Africa from Household Surveys: Monitoring and Restructuring the Market for Education", *Yale University Economic Growth Center Discussion Paper #875*, 2003.

UNAIDS, UNICEF, and USAID. *Children on the Brink 2002: a joint report of orphan estimates and program strategies*. 2002.

UNAIDS, UNICEF, and USAID. *Children on the Brink 2004: a joint report of new orphan* *estimates and a framework for action*. 2004.

UNAIDS. *Socioeconomic Impact of HIV/AIDS in Africa*. (www.unaids.org), 2000.

UNAIDS. *Report on the Global HIV/AIDS Epidemic*. (www.unaids.org), 2002.

United Nations Development Programme (UNDP). *Human Development Indicators 2002*.

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

Wax, Emily. "A Generation Orphaned by AIDS," *Washington Post*, August 13, 2003, pp. A01.

Wehrwein, Peter. "AIDS Leaves Africa's Economic Future in Doubt" (cnn.com), 2000.

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

Yamano, Takashi, and T.S. Jayne. "Measuring the impact of working-age adult mortality on small-scale farm households in Kenya," *World Development*, 2004, 32(1), 91-119.

Yamano, Takashi, and T.S. Jayne. "Working-age adult mortality and primary school attendance in rural Kenya," *Economic Development and Cultural Change*, 2005, 53(3), 619-54.

Young, Alwyn. (2005). "The Gift of Dying: The Tragedy of AIDS and the Welfare of Future African Generations", forthcoming *Quarterly Journal of Economics.*

Table 1: **Summary statistics**

	Obs	Mean	Std. Dev	Min	Max
Panel A: Full sample					
Female	18133	0.48	0.5	0	1
Age, 1998	14970	11.8	2.5	5	18
Became orphan, during 1999-2002	18133	0.08	0.27	0	1
Became maternal orphan, during 1999-2002	18133	0.06	0.23	0	1
Became paternal orphan, during 1999-2002	18133	0.03	0.17	0	1
Proportion orphans in school, 1998	18133	0.15	0.05	0.01	0.41
Proportion maternal orphans in school, 1998	18133	0.05	0.02	0.01	0.20
Proportion paternal orphans in school, 1998	18133	0.12	0.04	0	0.33
Proportion double orphans in school, 1998	18133	0.02	0.01	0	0.12
School participation, 1998	18133	0.85	0.23	0	1
School enrollment, 1998	18133	0.98	0.14	0	1
Panel B: Restricted sample					
Female	7815	0.48	0.50	0	1
Age, 1998	7769	12.9	2.0	6	18
Became orphan, 1999-2002	7815	0.09	0.28	0	1
Became maternal orphan, during 1999-2002	7815	0.06	0.24	0	1
Became paternal orphan, during 1999-2002	7815	0.03	0.17	0	1
Proportion orphans in school, 1998	7815	0.14	0.05	0.01	0.35
Proportion maternal orphans in school, 1998	7815	0.05	0.02	0.01	0.15
Proportion paternal orphans in school, 1998	7815	0.11	0.04	0	0.31
Proportion double orphans in school, 1998	7815	0.02	0.01	0	0.12
School participation, 1998	7815	0.92	0.17	0	1
School enrollment, 1998	7815	1	0.06	0	1
Child weight-for-age (z-score), 1998	7815	-1.44	0.82	-4.79	2.34
Child had malaria/fever in past month, 1998	7815	0.39	0.49	0	1
Child wears shoes, 1998	7815	0.14	0.35	0	1
Child wears school uniform, 1998	7815	0.86	0.34	0	1
Child appears "clean", 1998	7815	0.62	0.49	0	1
Latrine at home, 1998	7815	0.82	0.38	0	1
Cows at home, 1998	7815	0.49	0.50	0	1
Goats at home, 1998	7815	0.41	0.49	0	1
Poultry at home, 1998	7815	0.93	0.25	0	1

Notes: School participation variables are from regular unannounced checks collected throughout the 1998 to 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" is due to missing data; in the regressions, we include an indicator for observations with missing age data.

Table 2: Attrition and child characteristics

	Dependent variable:			
	Missing parent death information in 2002		Attrited during 1998-2002 (missing school participation)	
	Probit	Probit	Probit	Probit
	(1)	(2)	(3)	(4)
Female	0.029**	0.034	-0.006	-0.281***
	(0.012)	(0.063)	(0.016)	(0.057)
Age, 1998	0.016***	0.020***	0.007***	-0.012
	(0.002)	(0.007)	(0.001)	(0.008)
Missing age data	0.399***		0.083***	
	(0.033)		(0.025)	
Female * Age, 1998	0.000	-0.001	0.001	0.021***
	(0.001)	(0.005)	(0.002)	(0.005)
School in Budalangi Division	0.073***	0.065***	0.008	0.004
	(0.018)	(0.018)	(0.013)	(0.016)
Child weight-for-age (z-score), 1998		-0.074***		0.001
		(0.025)		(0.035)
Child had malaria/fever in past month, 1998		0.014*		-0.010
		(0.009)		(0.009)
Latrine at home, 1998		-0.063		0.126*
		(0.054)		(0.060)
Cows at home, 1998		-0.107*		-0.004
		(0.057)		(0.071)
Goats at home, 1998		-0.018**		-0.022*
		(0.008)		(0.012)
Poultry at home, 1998		-0.037**		-0.004
		(0.016)		(0.022)
Child wears shoes, 1998		0.033***		0.051**
		(0.012)		(0.022)
Child wears school uniform, 1998		-0.026*		0.050***
		(0.014)		(0.014)
Child appears 'clean,' 1998		-0.030		0.340***
		(0.054)		(0.048)
Cows at home, 1998 * Age, 1998		0.007		-0.001
		(0.004)		(0.005)
Latrine at home, 1998 * Age, 1998		0.005		-0.008
		(0.004)		(0.006)
Child appears 'clean', 1998 * Age, 1998		0.002		-0.025***
		(0.004)		(0.005)
Child weight-for-age(z-score), 1998 * Age, 1998		0.007		-0.001
		(0.002)		(0.003)
Became orphan, 1999-2002			-0.056***	-0.053***
			(0.011)	(0.016)
Observations	24,111	9,789	18,133	7,815
Mean (s.d.) of dependent variable	0.25	0.20	0.23	0.28
	(0.43)	(0.40)	(0.42)	(0.45)

Notes: All regressions are probits, with marginal effects reported. Standard errors are clustered at the school level. * significant at 90%; ** 95%; *** 99% confidence. Unreported controls include deworming program (PSDP) treatment group indicator variables, region-year indicator variables, and the constant. Regression 1 includes all 24,111 children from the baseline sample, and regression 2 includes all baseline children for whom baseline covariates are available. Regression 3 includes all 18,133 baseline children for whom parent mortality data is available, and regression 4 includes those among the 18,133 for whom covariates are available. Age data is missing for 5,095 baseline children.

Table 3: Baseline characteristics for children who lost a parent versus those who did not

	Became orphans	Never orphans	Difference Became – Never (s.e.)	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)	12.35
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	1245	13725		4046
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 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 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	7148		1938
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)	-0.08
Observations	250	2673		904

Notes: Standard errors are clustered at the school level. * significant at 90%; ** 95%; *** 99% confidence. 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.

Table 4: Impact of parent death on school participation

	Dependent variable: Total school participation					
	Full sample (1)	Full sample (2)	Restricted sample (3)	Restricted sample (4)	Restricted sample (5)	Restricted sample (6)
3 years pre-death	-0.013		0.007			
	(0.026)		(0.036)			
2 years pre-death	-0.037		-0.037			
	(0.029)		(0.038)			
1 year pre-death	-0.039		-0.037			
	(0.030)		(0.041)			
Year of parent death	-0.074**		-0.054			
	(0.031)		(0.041)			
1 year post-death	-0.060**		-0.054			
	(0.030)		(0.043)			
2 years post-death	-0.065**		-0.071			
	(0.033)		(0.049)			
3 years post-death	-0.089**		-0.070			
	(0.040)		(0.059)			
Pre parent death (1-2 years)		-0.021		-0.032*	-0.031*	-0.025***
		(0.015)		(0.019)	(0.019)	(0.009)
Post parent death		-0.055***		-0.054**	-0.053**	-0.036***
		(0.017)		(0.022)	(0.022)	(0.012)
Child 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	73070	73070	30817	30817	30817	30817
Mean (s.d.) of dependent variable	0.75	0.75	0.77	0.77	0.77	0.77
	(0.35)	(0.35)	(0.34)	(0.34)	(0.34)	(0.34)
R^2	0.54	0.54	0.58	0.58	0.58	0.11

Notes: All regressions are ordinary least squares (OLS). Standard errors are clustered at the school level. * significant at 90%; ** 95%; *** 99% confidence. Unreported controls include deworming program (PSDP) treatment group indicator variables, the full set of birth-year cohort-year-gender indicator variables, and region-year indicator variables, and the constant term. Regressions 1 and 2 contain 18,133 unique pupils, and regressions 3-6 contain 7,815 unique pupils. The pre/post death indicator variables are mutually exclusive categories. The additional controls in regressions 5 and 6 include all of the baseline household controls ("Child height-for-age (z-score), 1998" through "Poultry at home, 1998") interacted with indicator variables for the years 1999, 2000, 2001, and 2002.

Table 5: Impact of parent death on school participation, alternative comparison groups

	Dependent variable: Total school participation		
	Became orphans vs. Never orphans (Full sample)	Became orphans vs. Always orphans	Became orphans vs. Never orphans, Always orphans
	(1)	(2)	(3)
Pre parent death (1-2 years)	-0.021	-0.008	-0.010
	(0.015)	(0.012)	(0.010)
Post parent death	-0.055***	-0.042***	-0.045***
	(0.017)	(0.013)	(0.014)
Student fixed effects	Yes	Yes	Yes
Observations	73070	19176	85713
Mean (s.d.) of dependent variable	0.75	0.73	0.74
	(0.35)	(0.35)	(0.33)
R^2	0.54	0.56	0.54

Notes: All regressions are ordinary least squares (OLS). Standard errors are clustered at the school level. * significant at 90%; ** 95%; *** 99%. Unreported controls include deworming program (PSDP) treatment group indicator variables, a full set of birth-year cohort-year-gender indicator variables, region-year indicator variables, and the constant term. 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 4, regression 2.

Table 6: Impact of maternal and paternal deaths

	Dependent variable: Total school participation		
	Full sample	Full sample	Full sample
	(1)	(2)	(3)
Pre maternal death (1-2 years)	-0.065***	-0.065***	-0.067***
	(0.022)	(0.022)	(0.025)
Post maternal death	-0.093***	-0.096***	-0.091***
	(0.025)	(0.026)	(0.029)
Pre paternal death (1-2 years)	-0.005	-0.005	-0.009
	(0.018)	(0.018)	(0.029)
Post paternal death	-0.036	-0.037	-0.032
	(0.022)	(0.023)	(0.030)
Post-maternal death * Post-paternal death		0.014	
		(0.037)	
Pre first parent death			0.004
			(0.023)
Post first parent death			-0.004
			(0.026)
Student fixed effects	Yes	Yes	Yes
Observations	73070	73070	73070
Mean (s.d.) of dependent variable	0.75	0.75	0.75
	(0.33)	(0.33)	(0.33)
R^2	0.54	0.54	0.54

Notes: All regressions are ordinary least squares (OLS). Standard errors are clustered at the school level. * significant at 90%; ** 95%; *** 99% confidence. Unreported controls include deworming program (PSDP) treatment group indicator variables, birth-year cohort-year-gender indicator variables, region-year indicator variables, and the constant term. All regressions contain 18,133 unique pupils.

Table 7: Impact of parent death by child age and gender

	Dependent variable: Total school participation				
	Full sample (1)	Full sample (2)	Full sample (3)	Full sample (4)	Full sample (5)
Pre parent death (1-2 years)	-0.010	0.010	-0.024	-0.003	-0.010
	(0.017)	(0.020)	(0.019)	(0.022)	(0.019)
Post parent death	-0.036	-0.011	-0.051**	-0.030	-0.051*
	(0.022)	(0.025)	(0.022)	(0.025)	(0.027)
Pre maternal death (1-2 years)		-0.068**		-0.066**	
		(0.032)		(0.030)	
Post maternal death		-0.076**		-0.066*	
		(0.039)		(0.037)	
Child below age 12 * Pre parent death (1-2 years)	-0.038	-0.040			-0.040
	(0.026)	(0.034)			(0.032)
Child below age 12 * Post parent death	-0.049*	-0.051			0.001
	(0.025)	(0.034)			(0.034)
Child below age 12 * Pre maternal death (1-2 years)		0.007			
		(0.058)			
Child below age 12 * Post maternal death		0.005			
		(0.061)			
Female child * Pre parent death (1-2 years)			0.006	0.006	-0.003
			(0.027)	(0.030)	(0.033)
Female child * Post parent death			-0.007	-0.004	0.034
			(0.030)	(0.031)	(0.042)
Female child * Pre maternal death (1-2 years)				0.004	
				(0.040)	
Female child * Post maternal death				-0.005	
				(0.052)	
Child below age 12 * Female * Pre parent death					0.008
					(0.051)
Child below age 12 * Female * Post parent death					-0.104*
					(0.060)
Student fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	73070	73070	73070	73070	73070
Mean (s.d.) 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: All regressions are ordinary least squares (OLS). Standard errors are clustered at the school level. * significant at 90%; ** 95%; *** 99% confidence. Unreported controls include deworming program (PSDP) treatment group indicator variables, birth-year cohort-year-gender indicator variables, region-year indicator variables, and the constant term. The "Child below age 12" term refers to their age in the year of the parent death. Regressions 1, 2 and 5 also include the Missing age data indicator variable, and interactions between the Missing age data indicator and parent death terms – coefficient estimates not reported. All regressions contain 18,133 unique pupils.

<u>Table 8:</u> Impact of parent death by child and household characteristics

	Dependent variable: Total school participation			
	Restricted sample (1)	Restricted sample (2)	Restricted sample (3)	Restricted sample (4)
Pre parent death (1-2 years)	-0.029	-0.033	-0.041**	-0.033*
	(0.019)	(0.020)	(0.021)	(0.020)
Post parent death	-0.053**	-0.040	-0.054**	-0.043*
	(0.021)	(0.025)	(0.026)	(0.024)
1998 Test Score * Pre parent death (1-2 years)	0.038**			0.039**
	(0.018)			(0.018)
1998 Test Score * Post parent death	0.057**			0.057**
	(0.023)			(0.023)
No latrine at home * Pre parent death (1-2 years)		0.008		0.026
		(0.043)		(0.042)
No latrine at home * Post parent death		-0.072		-0.050
		(0.050)		(0.053)
Poor household * Pre parent death (1-2 years)			0.052	
			(0.042)	
Poor household * Post parent death			0.005	
			(0.055)	
Student fixed effects	Yes	Yes	Yes	Yes
Observations	28665	30817	30817	28665
Mean (s.d.) 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

<u>Notes:</u> All regressions are ordinary least squares (OLS). Standard errors are clustered at the school level. * significant at 90%; ** 95%; *** 99%. Unreported controls include deworming program (PSDP) treatment group indicator variables, birth-year cohort-year-gender indicator variables, region-year indicator variables, and the constant term. "Poor" is an indicator variable that takes on a value of one 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, child is "clean"). Regressions 1 and 4 contain 7,210 unique pupils, and regressions 2 and 3 contain 7,815 unique pupils.

<u>Table 9:</u> Community characteristics and the impact of parent death

	Dependent variable: Total school participation		
	Full sample (1)	Full sample (2)	Full sample (3)
Pre parent death (1-2 years)	-0.039	-0.035*	-0.036
	(0.042)	(0.020)	(0.040)
Post parent death	-0.000	-0.063**	0.003
	(0.061)	(0.028)	(0.059)
Proportion orphans in school	-0.310		
	(0.335)		
Proportion orphans in school * Pre parent death	0.082		
	(0.178)		
Proportion orphans in school * Post parent death	-0.232		
	(0.230)		
Top quartile, proportion orphans in school		0.005	
		(0.017)	
Top quartile, proportion orphans in school * Pre parent death		0.022	
		(0.020)	
Top quartile, proportion orphans in school * Post parent death		0.005	
		(0.021)	
Proportion maternal orphans in school			0.138
			(0.489)
Proportion paternal orphans in school			-0.341
			(0.332)
Proportion maternal orphans in school * Pre parent death			-0.647
			(0.533)
Proportion maternal orphans in school * Post parent death			-0.308
			(0.436)
Proportion paternal orphans in school * Pre parent death			0.360
			(0.334)
Proportion paternal orphans in school * Post parent death			-0.162
			(0.346)
Student fixed effects	Yes	Yes	Yes
Observations	73070	73070	73070
Mean (s.d.) of dependent variable	0.75	0.75	0.75
	(0.33)	(0.33)	(0.33)
R^2	0.58	0.58	0.58

<u>Notes:</u> All regressions are ordinary least squares (OLS). Standard errors are clustered at the school level. * significant at 90%; ** 95%; *** 99%. Unreported controls include deworming program (PSDP) treatment group indicator variables, birth-year cohort-year-gender indicator variables, region-year indicator variables, and the constant term. All regressions contain 18,133 unique pupils.

Appendix Tables

Table A1: Baseline characteristics for "became orphan" versus "always orphan" children

	Became orphans	Always orphans	Difference Became – Always (s.e.)
Panel A: Full sample			
Female	0.46	0.46	0.00 (0.02)
Age	11.8	12.1	-0.3*** (0.1)
School participation, 1998	0.87	0.87	0.00 (0.01)
School enrollment, 1998	0.99	0.99	0.00 (0.00)
N	1245	2676	
Panel B: Restricted sample			
Female	0.48	0.46	0.02 (0.03)
Age	12.7	13.0	-0.3*** (0.1)
School participation, 1998	0.92	0.91	0.00 (0.01)
School enrollment, 1998	1.00	0.99	0.00 (0.00)
Academic test score, 1998 (normalized)	-0.08	0.06	-0.13*** (0.05)
Child weight-for-age (z-score), 1998	-1.40	-1.47	0.07** (0.03)
Child had malaria/fever in past month, 1998	0.40	0.38	0.01 (0.03)
Child wears shoes, 1998	0.13	0.16	-0.03* (0.02)
Child wears school uniform, 1998	0.85	0.83	0.01 (0.02)
Child appears "clean", 1998	0.59	0.61	-0.02 (0.02)
Latrine at home, 1998	0.81	0.77	0.03 (0.03)
Cows at home, 1998	0.49	0.43	0.06** (0.03)
Goats at home, 1998	0.39	0.35	0.03 (0.02)
Poultry at home, 1998	0.93	0.90	0.03** (0.01)
N	667	1400	
Panel C: Subsample of children with school participation data in 1997 and 1998			
School participation, 1997	0.84	0.79	0.04 (0.03)
School participation, 1998	0.80	0.78	0.01 (0.02)
School participation, 1998 – 1997	-0.04	-0.01	-0.03 (0.03)
N	250	478	

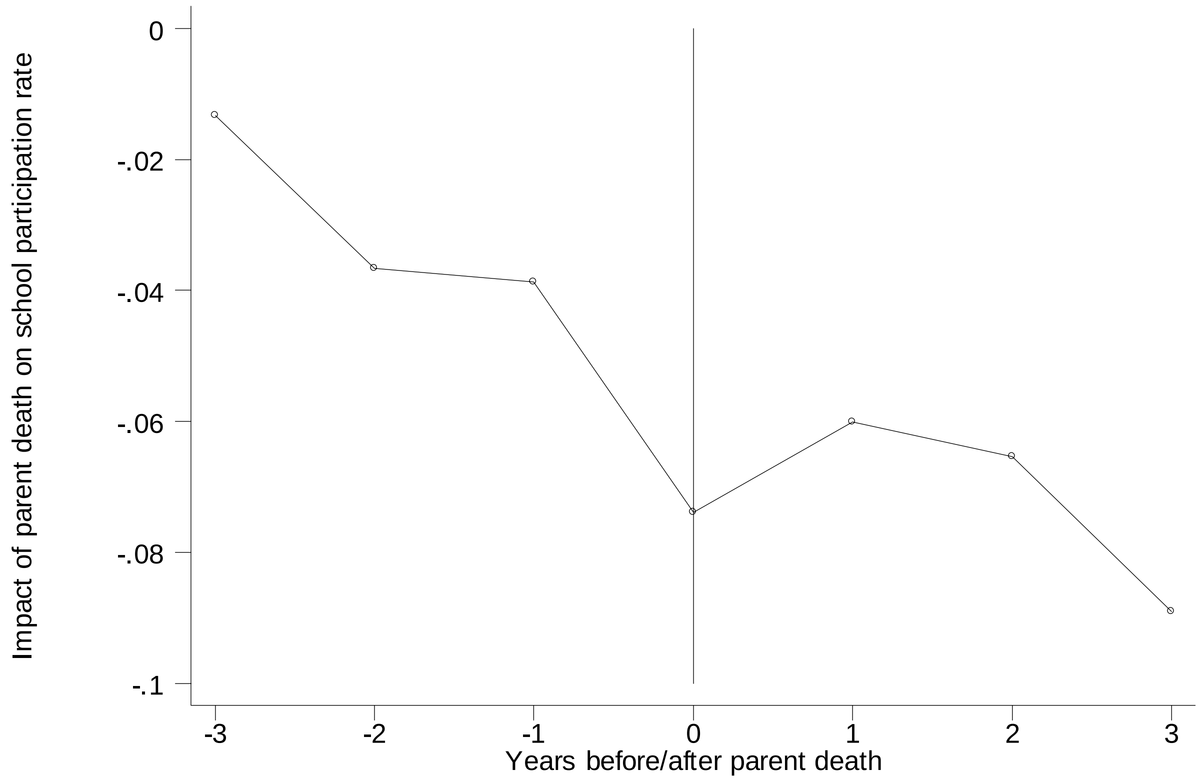
Notes: Standard errors are clustered at the school level. * significant at 90%; ** 95%; *** 99% confidence. For Panel C, 27 schools are included which were involved in another NGO program, and thus had 1997 attendance data. The reduction in sample size in Panel A (from 4,690 students to 3,921 students) is due to missing age information.

Table A2: Impact of parent death on school enrollment

	Dependent variable: Total school enrollment					
	Full sample (1)	Full sample (2)	Restricted sample (3)	Restricted sample (4)	Restricted sample (5)	Restricted sample (6)
3 years pre-death	-0.025		0.002			
	(0.020)		(0.026)			
2 years pre-death	-0.030		-0.012			
	(0.021)		(0.029)			
1 year pre-death	-0.036		-0.012			
	(0.023)		(0.032)			
Year of parent death	-0.055**		-0.027			
	(0.024)		(0.032)			
1 year post-death	-0.056***		-0.041			
	(0.022)		(0.034)			
2 years post-death	-0.064***		-0.055			
	(0.023)		(0.038)			
3 years post-death	-0.090**		-0.065			
	(0.033)		(0.053)			
Pre parent death (1-2 years)		-0.009		0.000	0.001	-0.004
		(0.011)		(0.015)	(0.015)	(0.006)
Post parent death		-0.034**		-0.024	-0.023	-0.026***
		(0.015)		(0.019)	(0.019)	(0.010)
Child weight-for-age (z-score), 1998						-0.013***
						(0.003)
Child had malaria/fever in past month, 1998						0.000
						(0.001)
Child wears shoes, 1998						0.003
						(0.002)
Child wears school uniform, 1998						0.000
						(0.001)
Child appears ‘clean,’ 1998						0.001
						(0.001)
Latrine at home, 1998						0.0014
						(0.0013)
Cows at home, 1998						0.0007
						(0.0007)
Goats at home, 1998						0.000
						(0.001)
Poultry at home, 1998						0.0013
						(0.0023)
Student fixed effects	Yes	Yes	Yes	Yes	Yes	No
Baseline controls * Year controls	No	No	No	No	Yes	Yes
Observations	73070	73070	30817	30817	30817	30817
Mean (s.d.) of dependent variable	0.89	0.89	0.88	0.88	0.88	0.88
	(0.32)	(0.32)	(0.33)	(0.33)	(0.33)	(0.33)
R^2	0.51	0.54	0.53	0.53	0.53	0.12

Notes: All regressions are ordinary least squares (OLS). Standard errors are clustered at the school level. * significant at 90%; ** 95%; *** 99% confidence. Unreported controls include deworming program (PSDP) treatment group indicator variables, the full set of birth-year cohort-year-gender indicator variables, and region-year indicator variables, and the constant term. Regressions 1 and 2 contain 18,133 unique pupils, and regressions 3-6 contain 7,815 unique pupils. The pre/post death indicator variables are mutually exclusive categories. The additional controls in regressions 5 and 6 include all of the baseline household controls (“Child height-for-age (z-score), 1998” through “Poultry at home, 1998”) interacted with indicator variables for the years 1999, 2000, 2001, and 2002.

<u>Figure 1:</u> Parent death and school participation over time (relative to four years prior to parent death)



Notes: Full sample, point estimates from a specification analogous to Table 4, regression 1.