The Importance of Parents for Access to Education: Findings from South Africa*

Solveig Argeseanu Cunningham, Emory University

Victoria Hosegood, Africa Centre for Health and Population Studies & London School of Hygiene and Tropical Medicine

Christopher R. Cunningham, Federal Reserve Bank of Atlanta

Caterina Hill, Africa Centre for Health and Population Studies

Running head: Is parental involvement associated with access to education?

Word count:

* Analysis was based on data collected through the Africa Centre Demographic Information System. We thank the ACDIS field and data centre staff and the Population Studies Group. This study benefited greatly from the input of Irma Elo, Jere Behrman, Susan Watkins, Elizabeth Vaquera and Sarah Hayford. Funding for this research was provided by NICHD Training Grant number 5-T32-HD07242. The Africa Centre population-based HIV survey and ANC sentinel surveillance was supported by Wellcome Trust grants #65377 and #50535 to the Africa Centre for Health and Population Studies. Please address correspondence to Solveig Argeseanu Cunningham; Hubert Department of Global Health; The Rollins School of Public Health; Emory University; 1518 Clifton Rd N.E.; Atlanta, GA 30322; sargese@sph.emory.edu.

ABSTRACT

This study explores the importance of living with parents for children's school attendance and progress. Using a regional census from KwaZulu-Natal, South Africa, we explore the differences in education between children living with both parents, orphans and children with living parents who do not belong to the same household. The objective of the paper is to examine a number of measures of education access and attainment: number of grades behind, early school entrance, drop out, and attendance of the highest levels. We find that children whose parents are both members of the same household are more likely to be enrolled in school, and have completed more schooling than, children with only one parent household member. Those with only one parent household member have, in turn, more schooling than those with neither parent in the household. Children with deceased parents do not have significantly less schooling than children whose parents are not household members for other reasons. Children are further behind in school if their mothers are absent from the household than if their fathers are absent. They are much further behind if their mothers are deceased than if their fathers are deceased. However, there are no significant differences in education between children whose fathers are deceased and those whose fathers are absent from the household. Boys are on average behind girls, and the difference appears to be driven by boys' greater vulnerability if they do not have a father in the household.

INTRODUCTION

This study explores the importance of living with parents for children's school attendance and progress. Because the education of children necessitates expenses, forgone income and services, and redistribution of scarce resources among household members, it depends upon support within the household. Parents are likely to be most concerned about their children's well-being. However, the support they provide is constrained by the amount of decision-making power they have within the household and by the support and resources they can contribute.

Studies of the effect of household composition on children's education have focused on parental mortality in less developed countries and on parental divorce and non-marital partnerships in the West. While several have examined the effects of parental mortality on children's educational outcomes in Sub-Saharan Africa, they have not taken into account the large number of children living without their parents for others reasons, such as abandonment and parental separation. For example, how do orphans compare with children living with only one parent for other reasons? Is it worse to lose a parent to death or to mobility? These comparisons are relevant to understanding how parents benefit their children's education: through monetary support, supervision, or advocacy. Few study populations and datasets make it possible to adequately address these questions in the developing world. Most surveys conflate household membership with residency, which risks omitting migrant workers who make important contributions to the household, while assuming a familial relationship between individuals who simply share the same quarters.

This paper uses a rich South African dataset that allows households to decide who is considered a member and that makes it possible to track both resident and non-resident household members. The households in the study population experience high extra-marital

fertility, labor migration and mobility, adult mortality, and inter-generational cohabitation and, consequently, diverse and dynamic living arrangements. Parental death is more common than in most parts of the world because of high AIDS mortality. Even children whose parents are not deceased do not necessarily live with both parents, and in fact only about a third of school-aged children do so.[1] Such rich variation of household types is unusual among both developed and developing countries and offers the possibility to include in our study children who are often ignored, such as those who have never lived with more than one parent, those who do not even know whether their parents are alive or deceased, and those who have partnered parents who are nonetheless separated most of the time.

We use several measures of educational progress, to determine whether children with different levels of access to their parents have different education outcomes in terms of school enrollment, progress, and completion.

PARENTS AND EDUCATION

With AIDS disrupting both household composition and education systems in sub-Saharan Africa, there has been considerable research on orphans and the consequences of losing a parent for educational attainment. Ainsworth and Filmer (2002) looked at the effects of poverty and AIDS on child enrollment in school at ages 7 to 14 using 39 nationally representative household surveys from developing countries. They found that orphans were not always concentrated in the poorest households. However, within households, orphans were less likely to attend school than other children. They also pointed out that households can redistribute orphans among themselves in order to maximize state aid.

[1] Authors' calculations.

Bicego et al. (2002) compared DHS data from several sub-Saharan African countries and found that orphans, while not disadvantaged economically compared with non-orphans, were less likely to be at age-appropriate grade levels than other children. Their results highlight that cultural factors that govern how families respond distribute resources among relatives.

Case, Paxson and Ableidinger (2004), using DHS data from African countries, find that orphans are less likely to be enrolled in school than other children and that their economic disadvantages do not entirely explain their lower enrollment rates. In fact, like Ainsworth and Filmer (2002), they find that orphans are disadvantaged in education compared with other children with whom they live. This disadvantage is explained by the tendency of orphans to live with more distant relatives, and there is an inverse relationship between relatedness and investment.

Nyamukapa and Gregson (2004) used quantitative and qualitative data from rural Zimbabwe and found that paternal orphans were as likely as other children to complete primary school, while maternal orphans were less likely than other children to do so. In qualitative responses, mothers expressed more willingness to sacrifice for children's education than did fathers.

Evans and Miguel (2005) observed the effect of parental deaths on primary school participation in Kenya over a period of five years. They compared the educational progress of children whose parents died during the period with those whose parents survived. School participation dropped by 9% after the death of a mother and 4% after the death of a father. The decrease was particularly large for girls under the age of 12. Children whose parents died started to show lower participation two years prior to the death and their participation remained low

after the parental death for as long as the study followed them (maximum 3 years). In addition, drop-out rates doubled after a parental death.

Beegle, De Weerdt and Dercon (2005) traced the effects of a parental death on child well-being in Northwestern Tanzania. They re-interviewed after ten years a group of children who were not orphans at baseline and found that, by the time they were adults, children who had lost their mothers had almost a year less education than non-orphans. The largest gap in education levels was among young children, who were not yet enrolled in school when their parents died: they lost approximately 54% of their expected schooling if they experienced a maternal death and 33% if they experienced a paternal death.

Case and Ardington (2006) used the same South African data source used in this paper to examine the effects of parental death on school enrollment and achievement.[2] They found that paternal orphans had lower educational outcomes, but much of the disadvantage could be explained by socioeconomic characteristics. However, maternal deaths caused poorer education outcomes: maternal orphans were 2-3% less likely to be enrolled in school than non-orphans, and, if they were enrolled, they were on average 0.2 years behind. Maternal orphans also had less household money spent on their education than non-orphans. The authors found no differences between maternal orphans and double orphans and argued that maternal orphans generally have little or no access to the father and so are *de facto* double orphans.

These studies have shown that children who have lost a parent are disadvantaged in their access to and achievement in school. In this essay, we extend the existing literature by considering children who are not members of the same households as their parents. Some of the disadvantages faced by orphans result from the experience of having a parent die; some of the

[2] Case and Ardington use a different sample of children from the one used in the current paper.

disadvantage is also likely the result of simply not having the protection, care, and resources of parents on a regular basis. If the latter effect exists, then other children who are not in the same households as their parents are disadvantaged. In this case, it is important to understand the effects not only of parental mortality but also of the child's access to his or her parents. Secondly, children who are not co-members with their parents cannot be ignored or clumped with other children from whom they are different because they represent a sizeable and growing portion of all children in South Africa. In our study population, 41% of school-aged children are not members of the same household as one parent and over 14% as both parents. Thirdly, by examining children who are not in the same household as their parents we separate mortality from abandonment and migration, and we may gain a better understanding of what pillars of parental support are the most important in ensuring child educational performance. If the parent is a member of the household, he or she is better positioned to influence the distribution of resources towards the child, while parents who are not members of the household have no say in allocation decisions.

Using detailed qualitative data from a neighboring area of South Africa, Townsend, Madhavan and Garey (2006) find that the framework that is usually used in the study of children's access to their parents is not applicable to Sub-Saharan Africa. They find that even though fewer than half of children were members of the same household as their fathers, only about 20% of children had no contact at all with their father. A substantial portion of those who did not share a household with their father nonetheless received some financial support from them. This is an interesting insight that poses a challenge to studies of parents' involvement in their children's wellbeing. Hosegood et al (2006), using data from the study population discussed in the present essay, also find that children who do not live with their fathers many do receive

money for school fees from their fathers. In this essay, we start from the assumption that, intentions aside, a parent's household membership allows the parent to advocate for the child within the household and direct resources towards the child. Therefore, if a child lives with a relative, even if the child has contact with his father, the father might not have much authority in the relative's household to influence the child's access to school.

An additional question addressed in this paper is whether one parent is a better advocate or agent for the child than the other. Several previous studies from the literature on orphans report that losing one's mother is worse for children than losing one's father (for example Bicego et al. 2002, Nyamukapa and Gregson 2004). Given that men tend to have higher earnings than women, the absence of the father, simply from a resources perspective, should be a greater shock to the child's well-being. We re-visit this question and extend it beyond membership, to determine whether the observed relationship holds when we look at each parent's co-membership with the child, as well as survival.

Though using the same dataset, the present study addresses essentially different research questions than Case and Ardington (2006). Case and Ardington investigated the effect of orphanhood on access to education, comparing those children who had lost one or both parents with those whose parents were alive. In this paper, we explore how children's education is affected by the composition of the household in which they are members. Certainly, parental mortality is an important component. However, in this analysis, we do not only compare orphans with non-orphans. Our comparison group consists of those children generally believed to be in the most supportive household arrangement: those children who are members of the same household as both their mothers and fathers. We also consider an important segment of the

population: children who are members of different households from their parents.[3] Though their parents are alive, they likely do not have access to the socio-economic resources and social and emotional support enjoyed by children living with their parents. By distinguishing non-orphans into those who are members of the same household as their parents and those who are not, distinguishes between two very different groups, allowing for clearer estimates of orphanhood on one hand and parental co-habitation on the other.

DATA AND METHODS

The study is set in a population of about 100,000 individuals located in the province of KwaZulu-Natal. It includes a township (a town designated for blacks during the apartheid regime), an informal settlement that is rapidly growing around it, and a rural population living on Zulu tribal land (Solarsh et al, n.d., Hosegood 2002). The average adult has 7.3 years of education (7.6 among men and 7.1 among women). Access to education has been changing over time, with older people, and especially older women, having limited education while younger cohorts having on average at least some high school education. Of people aged 25 and over, 22% have no schooling while another 22% have completed 12 years of education.[4]

There are over 38,000 school age children in the study area during the period of observation. About 75% of them are enrolled in school. In the age range for which school is mandatory, between ages 7 and 15, 95% of children are attending school. There are 52 primary

[3] Case and Ardington (2006) identify children whose parents are members but non-resident but they do not include them as a separate category when examining educational outcomes.

[4] Authors' calculations

schools and 19 high schools in the study area. School-aged children live on average 1.3 kilometers from a primary school and 2 kilometers from a high school.[5]

The data are from the Africa Centre Demographic Information System (ACDIS), collected by the Africa Centre for Health and Population Studies. These data offer a unique opportunity to study the implications of household composition and bargaining within the household on children's education. Unlike other comparable surveys, the ACDIS asks respondents who the members of the household are. This approach allows the cultural and economic considerations of the household to determine membership. At the same time, it allows us to distinguish between individuals' household of membership and place of residence, which might not always be the same. It also captures migrant workers who may be the principle earners and are likely to be part of household decisions, but who are usually absent from the homestead. We can identify and distinguish between individuals who are members in residence and individuals who are members of the household but reside elsewhere (discussion based on Hosegood and Timæus 2005; Hosegood, Benzler, and Solarsh 2005).

This paper focuses on all individuals between 6 and 19 years of age between January 2000 and December 2004 (see Appendix 1 for information on how household members were linked to households). The final analytic dataset consists of a panel of 37,731 children who were observed at least once between 2000 and 2004, for a total of 58,478 observations. More than half of the children (19,865) were observed at least twice.[6]

---

[5] Calculations by authors.

[6] Almost 75% of the children who only have one observation were age-censored by our study age brackets: they were under 9 or over 17 at the time they were interviewed and so were too young or too old to fit the age brackets at other household interviews. The others were only observed once because they died, moved into or out of the study

Ensuring an Inclusive Sample

Of children aged 6 to 19, 20% are not linked to their mothers and 55% are not linked to their fathers.[7] There are three reasons why the identity of a parent would be unknown in the dataset. 1) The parent died before the beginning of data collection and so does not have an individual record in the data. 2) The parent lives outside the study area and is not a member of a household in the study area and therefore does not have an individual record. This would happen primarily if the child moved into the study area without that parent, for example, to live with the other parent, with another relative, or independently to receive care, to work, or to attend school in the area. 3) The parent is a member of a household in the study area but has not been a resident or a member of the same household as the child since the beginning of data collection. In this case, the child and parent are not likely to be linked to each other in the data, even though each has individual records.[8]

One of the primary objectives of this paper is to extend the existing research to children with absent parents. This necessitates certain accommodations in order to include the children who cannot be linked with their parents. Because including in the analyses variables pertaining

---

area, or were otherwise lost to follow-up. About 3% of the children were observed more than once in the same year, most of them being members of multiple households.

[7] Survival status is known for 99% of mother and 84% of fathers, but detailed information about education, partnerships and presence is known for only about 40% of the fathers and 70% of mothers. We have also estimated models using only those children for whom at least the mother is known and including mother's education level. The results are similar to those presented in the paper.

[8] Children and parents are linked only if they are in the same household or if the interviewer has contact with the parent and the child on the same day, for example if the parent is a member of another household in the same compound.

to the parents would necessitate excluding these children, we use a slightly different specification from that which has been used in studies where most of the children live in two-parent households. In place of birth order, we use age order, the rank of the child in terms of age among all household members younger than 20. Another set of variables that is often used in such analyses is the parents' educational attainment. We use instead a measure of the education level of the most educated individual in the household.

Even when a parent and child are not linked, the survey asks about the survival status of the parent. Therefore, we know whether the parent is alive or deceased. There may be cases where the child does not maintain contact with the parent.[9] Since the data can only tell us the information that the respondent was able to provide, some of the parents who are coded as not members of the same household as the child may actually be deceased, for example if the child lost touch with the parent and does not find out that he died.

Measuring Educational Progress

Our primary measure of educational attainment is the education gap. It is the difference between the number of grades the child has completed and the number of grades the child would have completed had he or she enrolled at the appropriate age and subsequently passed all grades. The greater the gap, the further behind the child is. The education gap is well suited to measuring education in a study setting such as this, where most children attend school but many do not succeed in progressing from grade to grade and so fall behind and might never pass the higher grades. According to UNICEF (2006), in South Africa between 2000 and 2004, the net primary

[9] The respondent may not be knowledgeable and/or have a particular motivation to say that the father (often) has no contact with the child.

school enrollment ratio was 89 for girls and boys, and the net primary school attendance ratio was 94 for girls and 93 for boys.[10]

One weakness of the education gap measure is that controlling for child's age may not fully control for the effect of age on the education gap, because, by construction, a younger child can only have a small gap. Another weakness is that the gap requires time to accumulate, so it reflects what has happened to the child throughout his or her life. That is, if a child's mother died in 2001, and the child was on schedule in school before but he dropped out of school after her death, when the child is interviewed in 2003, he would only have a small gap in his education. For these reasons, it is important to also consider measures. We therefore also introduce four binomial outcome measures. One measure that is similar to the gap but overcomes the problem of the age bias, is whether the child is at least one standard deviation behind the average education actually completed by other children of the same age in the study. This measure also allows us to compare the child to his or her peers, who are exposed to many of the same community and school conditions. The second measure is whether the child dropped out. It is estimated for children younger than 18 at the time of interview.[11] This is a discrete outcome that does not have a direct connection to past education investment and is thus potentially more responsive to inter-temporal variation in household composition.

---

[10] The net primary school enrollment ratio is the number of children enrolled in primary school who belong to the age group that officially corresponds to primary schooling, divided by the total population of the same age group. The net primary school attendance ratio is the percentage of children who are in an age group officially corresponding to primary schooling and who actually attend primary school.

[11] Individuals over age 15 are not required to attend school, so those over age 16 who stopped attending school might not be thought of as actually dropping out. Therefore, we also ran this model restricting it to children under age 16.

The two other measures ask whether the household is making extra investments into the child's education. One is whether the child entered school early. Some families try to enroll their children in school before the compulsory age of 7, which requires some effort, either to send the child to a private school or to convince the public school that the child is able to start school early. The second measure is whether the child completed more than the required grades. Individuals over age 15 and those who have completed the ninth grade are not required to continue school, so those who complete more than nine grades are making a greater investment in education. This specification is restricted to children over age 15.

Other Explanatory Variables[12]

The child-specific variables are the child's gender[13], age, age order, and whether the child was ever a member of multiple households simultaneously[14]. Our controls for household

[12] Discussion of the other explanatory variables is in Appendix 2.

[13] In other settings, access to education has often been documented to be more difficult for girls than for boys because they are expected to help around the house or because the education of a girl is not seen as a good investment. However, there is evidence that discrimination against girls is less common in South Africa, at least at young ages than in other settings (for example, see Duflo 2003; Case, Paxson and Ablerdinger 2002). In fact, girls may have an advantage. For example, Duflo (2003) found that South African girls grew significantly more in height and weight if they lived with a grandmother of pension-eligible age, but there were no such benefits for grandsons. Girls and boys appear to be almost equally enrolled in school in the study area: 84% of girls and 86% of boys aged 6 to 19 were full time students the first time they were visited.

[14] Some children in the area are members of several households, for example splitting their time between maternal and paternal households (for details see Hosegood and Timaeus 2006). For the analyses, the child is assigned to one of the households in which he or she is a member so that each child appears only once in the dataset at any given time period. Information on other members and other household variables pertain to the household to which the child is assigned.

composition indicate whether any of the following are household members: the child's parents, as outlined above, the child's grandparents or other older men or women over 50, pension-eligible elderly, other school-aged boys and girls, and boys and girls under the age of 6. Several variables indicating socio-economic status are included in the models. We use information on the resources owned by the household as a proxy for wealth by creating an index of household wealth using principal component analysis (Detail on the methods and the variables used to create the index are shown in Appendix 3. See also Filmer and Pritchett 1999 and Dunteman 1989). The index is based on the factor loadings of variables indicating house construction materials; household amenities, ownership of commodities; and farm production and is calculated separately for 2000-2002 and 2003-2004. The wealth measure is created by segmenting this index into quartiles. A variable denotes the education level of the most educated individual in the household. Children who live far away from a school may find it more difficult or costly to attend, so we include the number of kilometers from the household to the nearest primary and the nearest secondary school, calculated using MapInfo software as the linear, or Euclidean, distance based on GIS-coded locations of households and schools.

Finally, we control for some aspects of time with a series of dummy variables for interview year of observation (2000-2004) to remove purely inter-temporal variation in educational attainment rates and a term indicating whether the child was eligible to be enrolled in school before a series of changes in age at school enrollment began.

Analytic Strategy

The final dataset contains 58,478 observations on 37,731 individuals.[15] GLS (generalized least squares) random effects models are used to examine the determinants of the education gap:

$$ED_{it} = \beta_0 + \beta_1 I_{it} + \beta_3 H_{kt} + \alpha S_{kt} + \beta_5 D_{kt} + \beta_4 Y + E_{it} \quad (1)$$

(1)

where ED_{it} is the education gap, I_{it} is a vector of child-specific characteristics as discussed above, H_{kt} is a vector of household endowments as discussed above, D_{kt} *is a* vector of distances to schools and Y is a vector of year dummies. The household composition vector S_{kt} consists of the following:

$$S_{kt} = \alpha_1 P_{it} + \alpha_2 w_{kt} + \alpha_3 D_{kt} + \alpha_4 C_{kt} + \alpha_5 YC_{kt} \quad (2)$$

(2)

P_{it} is a vector of dummy variables indicating the eight possible combinations of parental survival and membership in the household in which the child is a member: mother is a member of the household but father is not, father is a member but mother not, mother is a member but

[15] Observations with missing values on some variables were dropped. The final dataset contains 58,478 observations on 37,731 individuals. Some observations had to be dropped because of incomplete information: 6 observations were dropped because the gender was not known; 4,296 observations were dropped because no information was available about the household in which the children were members; 1,967 observations were dropped because they could not be matched to a particular residence and thus their distance to nearest school could not be calculated; 4,518 observations were dropped because there was no information on the number of years of education completed by the child. Among these, 215 observations had information on years of education completed that were most likely wrong and were therefore coded as missing. That is, the children were recorded as having completed substantially more years of school than possible for their age.

father is deceased, father is a member but mother is deceased, father is not a member and mother is deceased, mother is not a member and father is deceased, neither parent is a member of the household, both parents are deceased. The omitted category value is both parents are members of the same household as the child.

The variables D_{kt} indicate that the household has member men and women over 50 and w_{kt} denotes that the household has pension-eligible elderly (men over age 64 and women over age 59). The variables C_{kt} count the number of other members who are school aged boys and girls (ages 6-19) and YC_{kt} indicates whether there are boy and girl members aged 5 and younger.

The error term, E_{it} is decomposed as follows:

$$E_{it} = v_i + e_{it} \tag{3}$$

where v_i is a time-invariant error term that accounts for unobserved, but normally distributed, endowments of scholastic ability and e_{it} is the standard error term. This method controls for unobserved heterogeneity that is uncorrelated with any of the right hand side variables.

A separate specification estimates the differences between girls and boys. The equation takes the same form and variables, but includes an additional interaction term:

$$ED_{it} = \beta_0 + \beta_1 I_{it} + \beta_2 H_{kt} + \alpha S_{kt} * \beta_5 d^{male} + \beta_5 D_{kt} + \beta_4 Y + E_{it} \tag{4}$$

where the gender variable, d^{male}, is interacted with the household composition variables, S_{kt}.

Finally, we present four alternative measures of progress in school: 1) entered school before required age; 2) completed required grades; 3) is one standard deviation behind average schooling for age; and 4) dropped out. Each of these binomial measures is estimated with logistic regression with random effects:

$$\ln\left(\frac{ED}{1-ED}\right) = \beta_0 + \beta_1 I_{it} + \beta_3 H_{kt} + \alpha S_{kt} + \beta_5 D_{kt} + \beta_4 Y + E_{it} \tag{5}$$

RESULTS

INSERT TABLE 1

Descriptive statistics of the children and their households are shown in Table 1. The average gap between age-appropriate and actual educational attainment for the study population is 1.2 years. Figure 1 illustrates the size of the gap between age-appropriate and actual education for each age between 6 and 19. After the age of 15, the gap between actual and expected years of education widens as school is no longer compulsory after this age. The figure also shows that boys start to show a larger gap than girls as early as age 9 and the difference between the genders is highest between ages 13 and 16, reaching about half a year of school.

INSERT FIGURE 1

As shown in Table 1, at the first observation, about 30% of children were members of the same household as both parents. Almost 45% were living with only one parent, over 80% of these with the mother. For the majority of these children, the other parent is alive but is not a member of the same household, while almost 12% are orphans. Twenty-four percent of children live with neither parent. About 20% of children are single orphans and almost 3% are double.

The Education Gap

Table 2 shows results comparing the education gap of children who are in the same household as both parents with those of children in other household arrangements.[16] Children who are in the

[16] In the text, we convert the coefficients of the education gap into school time for easier interpretation. In KwaZulu-Natal, children attend school 41 weeks (between 200 and 204 days (South African Government Information 2006). To convert the coefficient estimates into number of school weeks, we multiply the coefficient by 41 and round to the nearest week.

same household as both parents are far better off than children in other household arrangements. Those living with only one parent because the other parent is deceased are between 3 and 7 weeks behind, while those living with only one parent because the other parent is not a member of the same household are 6 to 7 weeks behind those living with both parents. This suggests that orphans are not significantly further behind other children living with just one parent. Children living only with their fathers are about 7 weeks behind, whether their mothers are deceased or living elsewhere.

Furthest behind are those children who have no parent in the household because the parents are in other households, or significantly worse, deceased ($p<0.001$[17]). They have 3 to 4 months less education than children living with both parents.

INSERT TABLE 2

A striking finding in this table is that children are at least as far behind in school if their fathers or mothers are not members of the same household as if they are deceased. If the child's mother is not a member of the household, he or she tends to be even further behind in school if the father is also absent than if he is deceased ($p=0.0268$). This is surprising, since an absent parent might still provide some small amount of help and encouragement to the child. Furthermore, children are likely to be emotionally scarred and possibly physically ill after the death of a parent.[18]

[17] P-values shown in this section are from Wald tests of linear hypotheses about the parameters of the model.

[18] In this paper, the cause of parental death is not considered, in part because many parents will not have been registered in the data at the time of their death and therefore the cause of death was not known. From the cause of death profiles of the population, however, we know that high male mortality of non-AIDS causes meant that many

Children whose mothers are deceased or absent are significantly further behind than those missing their fathers (p=0.0451 and p=0.000, respectively).

Different gap for boys and girls

As shown in Table 3, boys have an education gap 4 months greater than girls (mean girl gap is 1.03, mean boy gap is 1.43; .4*41 weeks=16 weeks). Thus, girls actually have an advantage, as has been found in other parts of Sub-Saharan Africa. But let us consider whether the educational progress of girls and boys is associated in similar ways with the presence of parents. Results are found in Table 5. Column 2 shows the estimated effect of the household composition characteristics for girls' education gap. Column 3 shows the additional effect for boys. To calculate the estimated education gap of boys living in the same household conditions, we add the main effects and the interaction effects, as shown in the table notes.

Both boys and girls tend to get more education if they are members of the same households as their parents. Girls are significantly further behind in school if they are not living with both parents. However, boys are even further behind. The only exception is when fathers belong to the household and mothers do not. Girls are not significantly worse off if either of their parents is deceased, as long as the other parent is a household member. However, boys who have lost their fathers and live with their mothers are about 9 weeks behind girls living in the same situation. Boys living with neither parent are 5 to 11 weeks behind girls in the same situation.

INSERT TABLE 3

Other measures of education

---

parents, particularly fathers, will have died of other causes such as violent and accidental deaths (Hosegood, Vanneste and Timaeus (2004).

Above, we argued that the education gap is a good tool for measuring the extent to which children are keeping to the school schedule. We also pointed out that the gap measure has weaknesses and shortcomings, and that it is therefore useful to consider several other measures for a more nuanced assessment of progress in education. In this section, we present results from four alternative measures: whether the child is at least one standard deviation behind the average education actually completed by other children of the same age in the study; whether a child younger than 16 dropped out; whether the child entered school before the required age; whether, among children who are 16 years or older (the age at which children are no longer required to go to school) the child completed more than the required number of grades. Taken together, these measures capture the components of the education gap. Table 4 shows the results for these alternative outcome variables.

INSERT TABLE 4

The results are consistent across measures. As we found by focusing on the education gap, children living with both parents have significantly better educational outcomes than children living in all other arrangements: they are more likely to have started school before the required age and to have completed more than the required grades. They are also more likely to be on schedule in school and less likely to have dropped out. Compared with children who are in a household with both parents, those who are living with just one parent because the other parent is absent or deceased are between 24% and 36% less likely to have entered school early, are between 36% and 57% less likely to stay in school beyond the required grades, are between 62% and 2.6 times more likely to be at least a standard deviation behind their peers, and are between 35% and 2.2 times more likely to have dropped out. Those who have neither parent in the household are even worse off in terms of all measures. Compared with children living with both

parents, those living with neither parent because of death and parental absence are between 40% and 55% less likely to have entered school early, between 73% and 77% less likely to go beyond the required grades, are 4 to 5 times more likely to be at least a standard deviation behind their peers, and are between 2.3 and 3.5 times more likely to have dropped out.

It is interesting to note that, among those whose parents are not household members, orphans fare better than other children living with neither parent on all measures except age of school entry. Thus, while orphans are by far worse off on all measures of education compared with children living with both parents or with one parent, compared with other children living with neither parent, they are more likely to stay in school and to complete more than the required grades and are less likely to be a standard deviation or more behind their peers (though we know from the gap measure that they are more likely to be behind at all than all other children).

DISCUSSION

This paper shows that household composition is associated with children's educational attainment. Not surprisingly, parents are especially important for educational attainment: children are much more likely to progress well in school if they are members of the same household as one, or, even more so, both parents.

Children not living with both parents have completed less education than other children. They are less likely to have started school before the required age and less likely to have completed more than the required grade. They are also more likely to be far behind their peers in school or have dropped out altogether.

Those who have no parent in the household are worst off, lagging about 4 months of school behind. There is anecdotal evidence that children sometimes move into households to

take advantage of educational opportunities, either to live close to a good school or to stay with a family that can afford to send them to school. While some children may improve their educational outcomes in this manner, the results suggest that the majority of children living with neither parent are falling behind in school or not attending, perhaps working as servants or herd boys. Ford and Hosegood (2005) emphasize that migration is a "positive choice" for some children and is a situation of "crisis fostering" for others, likely leading to different outcomes for the child.

These relationships differ by gender. The loss of one or both parents is associated with worse outcomes for boys than for girls. Though boys in this population are on average 4 months behind girls, in a two-parent household (the omitted category) boys may do slightly better than girls (as shown in Table 7).

An interesting finding is that children with a deceased parent are not significantly worse off by many measures than children whose parents are not household members. This is surprising, since an absent parent might still provide some small amount of help and encouragement to the child. Also, even if he or she does not have decision-making power in the household, the fact that a parent might be able to monitor the child's education from afar should create the threat, for the household members and for the child, that he or she will return and take action if the child is not making good progress. Furthermore, children are likely to be emotionally scarred and possibly physically ill after the death of a parent. On the other hand, this relationship is consistent with other studies and explanations. It is consistent with Beegle et al.'s (2005), who found that children whose parents are deceased do better if they were not living with their parents at the time of their death. Indeed, some of the parents who are categorized as not members in this analysis may actually be deceased. There are several reasons why orphans

might not be worse off. One reason for this may be that if a parent dies, his or her side of the family is more likely to assist the child than if the parent has abandoned or lost touch with the child. Yet another reason may be that adults caring for orphaned children may receive state aid more easily than non-parental adults caring for other children. It could be that they are more likely to secure the appropriate documentation for state grants, or that they tend to know that they are eligible and therefore to apply. This result may reflect investments in children's education that parents made before they died. It may also suggest that state grants may succeed in ameliorating some of the loss in education resulting from losing a parent.

Finally, this is consistent with Lund and Hosegood's (2005) study of children's residential stability in the same population. They found that children whose parents were members of the same household as them were less likely to move than children whose parents belonged to other households, even if the parents themselves were not actually residing with the household. They suggest that the parents' ties to the household may indicate that the children also have stronger kinship ties there.

We do find that children living without their fathers, especially those whose fathers are not members in the same household, have significantly worse outcomes than children who belong to the same household as their fathers.[19] However, Consistent with other studies, we found that children whose fathers are deceased are better off than those whose mothers are deceased. Access to grants might also be the reason for part of the differences between maternal and paternal orphans. Case, Hosegood, and Lund (2005) documented that mothers are more likely than other relatives to go through the steps necessary to secure orphan grants for their

[19] This finding is also consistent with Ford and Hosegood (2005), who found that children whose fathers were members of the same household, whether or not he was a resident, enjoyed more stable living arrangements.

children. Fathers and other adult caregivers may not even know that they are eligible and so might not apply. However, they found that children whose mothers were absent but were members of the household were less likely to be getting child support grant than those who lived with their mothers, suggesting that there is something about the presence of a mother that champions her children regardless of her financial status.

We find that, under most circumstances, boys have worse education outcomes than girls when they are not members of the same household as both parents and that they are particularly disadvantaged if they are not living with their fathers. It could be that fathers are particularly good advocates for boys within the household, so that boys receive less encouragement and resources if their fathers are not in the household. Alternatively, boys may be in need of more supervision and encouragement from their father, so that they do not attend or participate in school if their father is not there to do this.

As we consider the magnitude of the results, one obvious question is the long-term effect of the educational disadvantages on children. It is lamentable that children who live with neither parent receive less schooling than children who live with both parents. But is this difference large enough to affect the child's health, social, or economic wellbeing beyond this? This is a difficult question to answer without access to other indicators over the lifetime. Some studies that assessed the effects of education have focused on earnings. A standard estimate, based on numbers from the U.S., is that a one year increase in education increases earnings by 10%. Similar estimates based on South African data estimate the earnings increase at 4.1% using 1994 data (data compiled by Psacharopoulos and Patrinos 2004). Children in the study area living in difficult household arrangements are up to half a year behind. On the job market, this may be effectively rounded off to one less year of completed education.

A possible complication is that, in this area where the majority of adults are not involved in formal employment, education may matter less: whether a young person has 9 or 8 years of education, she may still not get a job. On the other hand, we might argue that education has the same effect in informal as in formal job markets, so that a more educated person is more likely to find work as a housecleaner or to sell fruit at the market than a less educated person. Let us think about education in the local context using other information from the Household Socio-economic Survey. Based on very rough estimates, we do find evidence that more educated adults are financially more successful. For example, among those adults aged 20 to 59 who responded to each question using simple specifications controlling for age, age squared, age cubed, gender and year of interview, we find the following: A one year increase in education was associated with a 10% increase in the odds of earning an income (odds ratios from logistic regression, N=31,941), with greater improvements for women than for men. For each additional year of education, an individual earns on average 155 rand per month (results from OLS regression, N=3,966; median reported monthly earnings are 700 rand, or about 70 dollars), with women having larger earnings benefit from education than men. For each additional year of education, an individual is 21% more likely to be in a higher wealth quartile than in a lower one (odds ratio from ordered multionomial logistic regression, N=66,416). These are crude estimates, only controlling for basic characteristics and not considering the endogeneity between education and socio-economic status, but they might be taken as evidence that, in this area as well, education matters for long-term socioeconomic status.

BIBLIOGRAPHY

Ainsworth, Martha and Deon Filmer. 2002. "Poverty, AIDS and Children's Schooling: A Targeting Dilemma." *World Bank Policy Research Working Paper*.

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon. 2005. "Orphanhood and the long-run impact on children." *unpublished*.

Bennet, Janette. 2003. "Millions spent on maths and science." Pp. http://www/sundaytimes.co.za/2003/02/23/business/surveys/survey09.asp in *Sunday Times*. Durban.

Bicego, George, Shea Rutstein, and Kiersten Johnson. 2002. "Dimensions of the Emerging Orphan Crisis in Sub-Saharan Africa." *Social Science & Medicine* 56:1235-1247.

Case, Anne and Cally Ardington. 2006. "The Impact of Parental Death on School Outcomes: Longitudinal Evidence from South Africa." *Demography* 43:401-420.

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

Coombe, Carol. 2000. "Keeping the Education System Healthy: Managing the Impact of HIV/AIDS on Education in South Africa." *Current Issues in Comparative Education* 3:1-13.

Stata Corporation. 2003. *Stata 8 Reference*, Edited by S. Press. College Station.

Dunteman, George H. 1989. *Principal Components Analysis*, vol. 69. Thousand Oaks, CA: SAGE University Press.

Evans, David and Edward Miguel. 2005. "Orphans and Schooling in Africa: A Longitudinal Analysis." *unpublished*.

Ford, K. & Hosegood, V. (2005) AIDS mortality and the mobility of children in KwaZulu Natal, South Africa. *Demography*. 42(4):757-768.

Filmer, D. and L. Pritchett. 1999. "The Effect of Household Wealth on Educational Attainment: Evidence from 35 Countries." *Population & Development Review.* 25:85-120.

Hosegood, V. 2002. "Basic Descriptives." Unpublished. Mtubatuba: Africa Centre for Health and Population Studies.

Hosegood, V., Benzler, J. & Solarsh, G. (2005) Population mobility and household dynamics in rural South Africa: implications for demographic and health research. *Southern African Journal of Demography.* 10(1&2):43-67

Hosegood, V. and K. Ford. 2003. "The impact of HIV/AIDS on children's living arrangements and migration in rural South Africa." Presented at Conference on African Migration in Comparative Perspective, Johannesburg.

Hosegood, V.and I.M. Timæus. 2006. "Household composition and dynamics in KwaZulu Natal, South Africa: Mirroring social reality in longitudinal data collection." Pp. 58-77 in *African Households: Censuses and Surveys*, edited by Etienne van de Walle. Armonk, NY: M. E. Sharpe.

Hosegood, V., A.-M. Vanneste, and I.M. Timæus. 2004. "Levels and causes of adult mortality in rural South Africa: the impact of AIDS." *AIDS* 18:663-671.

Hosegood, Floyd, Marston, Hill, McGrath, Isingo and Zaba. 2006.

Itano, Nicole. 2003. "Government Defends School Fees Policy." Pp. www.warmafrica.com/index/geo/5/cat/3/a/a/artid/432 in *VOANEWS*, vol. 2004. Johannesburg.

Naidoo, Subashni. 2003. "Pupils "carded" for school fees." Pp. www.suntimes.co.za/2003/02/09/news/durban/ndbn01.asp in *Sunday Times*. Durban.

—. 2004. "Pupil 'threatened with demotion' for unpaid fees'." Pp. http://www.sundaytimes.co.za/Articles/TarkArticle.aspx?ID=946340 in *Sunday Times*. Durban.

Ntozi, James P M. 1997. "Effects of AIDS on children: the problem of orphans in Uganda." *Health Transition Review* 7:23-40.

Nyamukapa, C. and S. Gregson. 2004. "Extended Family's and Women's Roles in Safeguarding Orphans' Education in AIDS-afflicted rural Zimbabwe." *Social Science & Medicine* 60:2155-2167.

Paley, Dawn. 2004. "Election Day in South Africa." Pp. www.alternet.org/story.htm?StoryID18425, vol. 2004. Johannesburg: Alternet.org.

Pandor, Naledi. 2004. "Access and Equity in South African Education." Pp. http://www.info.gov.za/speeches/2004/04112912121001.htm in *World conference on the righ to and rights in education*. University of Amsterdam.

Pearce, Justin. 1999. "Education: A promise hard to keep." Pp. news.bbc.co.uk/1/hi/special_report/1999/05/99/south_africa_elections/352189.stm in *BBC News*. London.

—. 2004. "SA poor's education struggle." Pp. news.bbc.co.uk/2/hi/africa/3605275.stm in *BBC News*.

Pile, Jacqui. 2005. "Orphans' trials mount." Pp. http:wwww.sundaytimes.co.za/zones/sundaytimesNEW/basket16st/basket16st116594070 in *Sunday Times*. Durban.

Psacharopoulos, George and Harry Anthony Patrinos. 2004. "Returns to Investment in Education: A Further Update." *Education Economics* 12:111-135.

Solarsh, G., J. Benzler, V. Hosegood, and F. Tanser. Unpublished Draft. "Hlabisa Demographic Surveillance System: South Africa." Mtubatuba: Africa Centre for Population Studies and Reproductive Health, University of Natal, University of Durban, and South African Medical Research Council.

South African Government Information. 2006. "South African Government Information: School Calendar." http://www.info.gov.za/aboutsa/school_cal/.

UNICEF. "South Africa: Statistics." Pp. http://www.unicef.org/infobycountry/southafrica_statistics.html#15, vol. 2006.

Table 1: Summary statistics of children aged 6 to 19 the first time they were observed

Characteristic	Mean or count	SD or percent
DEPENDENT VARIABLES (mean and SD)		
Education gap in years	1.16	1.74
Girls' education gap in years	1.03	1.60
Boys' education gap in years	1.43	1.84
PARENTAL MEMBERSHIP AND SURVIVAL (count and percent)		
Both parents household members(omitted)	11,576	30.80%
One parent absent, of which:	12,187	32.29%
Mother member, Father not member	*9,910*	*81.32%*
Father member, Mother not member	*2,277*	*18.68%*
One parent deceased, of which:	4,379	11.60%
Mother member, Father deceased	*3,732*	*85.22%*
Father member, Mother deceased	*647*	*14.78%*
With neither parent, of which:	9,589	25.42%
Neither parent member	*5,357*	*55.87%*
Mother deceased, Father not member	*1,708*	*17.81%*
Father deceased, Mother not member	*1,535*	*16.01%*
Both parents deceased	*989*	*10.32%*
CHILD CHARACTERISTICS		
Male (count and percent)	18,719	49.61%
Age in years (mean and SD)	11.85	4.17
Age order among household children (mean and SD)	2.92	1.99
Multiple household memberships (count and percent)	3,213	8.52%
EDUCATION IN HOUSEHOLD (mean and SD)		
Education of the most educated household member (years)	10.70	2.48
HOUSEHOLD CHARACTERISTICS (count)		
Wealth quartile – poorest	9,754	
Wealth quartile – 2nd poorest	9,384	
Wealth quartile – 2nd wealthiest	9,741	
Wealth quartile – wealthiest	8,852	
EUCLIDIAN DISTANCES, KM (mean and SD)		
To primary school	1.26	0.7
To secondary school	1.97	1.03
Number of children	**37,731**	

Data Source: Africa Centre Demographic Information System, KwaZulu-Natal, South Africa: Children aged 6 to 19, 2000-2004.

Note: Household characteristic values are scores based on eigenvalues from principal component analysis. Wealth quartiles were determined at the household level and are therefore not exactly uniformly distributed when matched to children.

Table 2: Estimates of the effect of parental survival and household membership on the education gap for all children aged 6 to 19, 2000-2004[a]
Coefficient estimates from random effects models

Characteristic	β	se
PARENTAL MEMBERSHIP AND SURVIVAL		
Both parents household members(omitted)		
One parent absent		
Mother member, Father not member	0.143	(0.019)**
Father member, Mother not member	0.178	(0.029)**
One parent deceased		
Mother member, Father deceased	0.073	(0.024)**
Father member, Mother deceased	0.182	(0.045)**
With neither parent		
Neither parent member	0.355	(0.022)**
Mother deceased, Father not member	0.347	(0.032)**
Father deceased, Mother not member	0.349	(0.034)**
Both parents deceased	0.390	(0.041)**
CHILD CHARACTERISTICS		
Male	0.415	(0.015)**
Age in years	0.108	(0.010)**
Age quadratic	0.004	(0.000)**
Age order among household children	0.018	(0.006)**
Multiple household memberships	-0.013	(0.027)
EDUCATION IN HOUSEHOLD		
Education of the most educated household member, years	-0.112	(0.003)**
HOUSEHOLD CHARACTERISTICS		
Wealth quartile – poorest (omitted)		
Wealth quartile – 2nd poorest	-0.071	(0.014)**
Wealth quartile – 2nd wealthiest	-0.146	(0.016)**
Wealth quartile – wealthiest	-0.260	(0.019)**
EUCLIDIAN DISTANCES, KM (mean and SD)		
To primary school	0.024	(0.011)*
To secondary school	-0.002	(0.007)
Number of children	**37,731**	

Data Source: Africa Centre Demographic Information System, KwaZulu-Natal, South Africa: Children aged 6 to 19, 2000-2004.

Notes: Constant term included. Statistical significance denoted: + p<.10; * p<.05; ** p<.01.

The model also controls for year of observation and for the following household composition variables: elderly household members, pensionable elderly members, age of elderly members and other school-aged child members, child members younger than 6.

[a] The education gap is measured as the difference between age-appropriate and actual years of education.

Table 3: Effects of household composition on the education gap for boys and girls [a]
Coefficient estimates from random effects models with interactions between gender and household composition variables

Characteristic	Main effects (estimates for girls) [b]		Interaction effects for gender (additional effect for boys) [c]		Significant gender differences [d]
	β	**se**	**β**	**se**	
Male	-0.004	(0.003)			
Both parents household members (omitted)					
One parent absent					
Mother member, Father not member	0.104	(0.026)**	0.091	(0.034)**	Yes
Father member, Mother not member	0.162	(0.042)**	0.065	(0.057)	
One parent deceased					
Mother member, Father deceased	-0.040	(0.032)	0.231	(0.044)**	Yes
Father member, Mother deceased	0.105	(0.066)	0.164	(0.090)+	
With neither parent					
Neither parent member	0.234	(0.030)**	0.268	(0.040)**	Yes
Mother deceased, Father not member	0.289	(0.045)**	0.126	(0.062)*	
Father deceased, Mother not member	0.213	(0.048)**	0.290	(0.066)**	Yes
Both parents deceased	0.260	(0.058)**	0.264	(0.079)**	Yes
	B	**se**	**β**	**se**	
Constant	0.432	(0.084)**			
Observations		58,477			
Number of Children		37,731			

Data Source: Africa Centre Demographic Information System, KwaZulu-Natal, South Africa: Children aged 6 to 19, 2000-2004.

Notes: Statistical significance denoted: + p<.10; * p<.05; ** p<.01.

The model also controls for child's sex, age, age squared, age order among household children, whether he or she had multiple simultaneous household memberships, and whether he or she was school-aged after 1998 law change; years of education of the most educated household member, household

wealth quartile, Euclidean distance to schools, and year of observation fixed effects. . It also controls for other household composition variables: elderly household members and other children.

[a] The education gap is measured as the difference between age-appropriate and actual years of education.

[b] The estimated effect of the household composition characteristics for girls' education gap is the coefficient estimate shown in Column 2.

[c] To calculate the estimated effects of the household composition characteristics for boys' education gap, we add the interaction coefficient estimate in Column 3 to the corresponding main effect coefficient estimate in Column 2. For example, the estimated effect of being a double orphan is a gap of 0.26 years for girls and 0.524 (0.26+0.264) years for boys. As a result, we can predict that female double orphans are .26 years behind and male double orphans are .520 (0.260+0.264-0.004) years behind compared with children living with both parents. The parameter estimate on the interaction term is statistically different from zero, suggesting that male orphans are further behind than female orphans.

[d] Effects of this characteristics are significantly different from boys compared with girls at the .01 level.

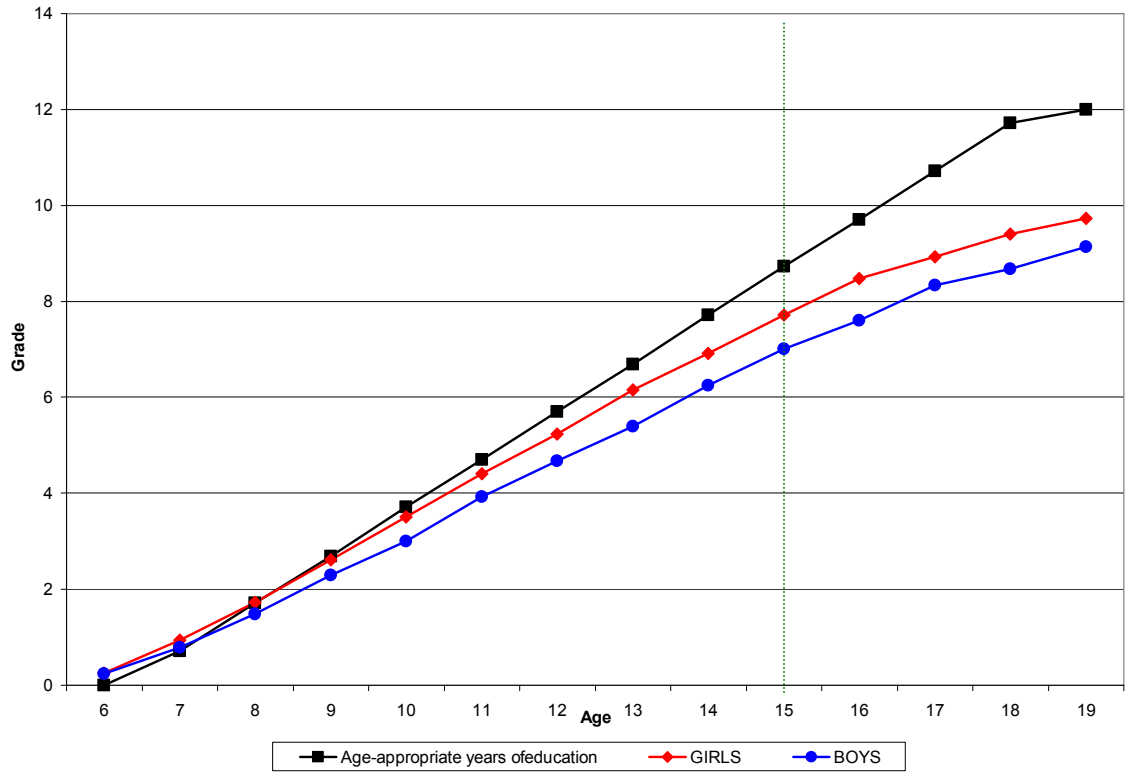
Table 4: Effects of parental survival and household membership on four alternative measures of progress in education
Odds ratios from logistic regressions with individual-level random effects

	Ahead in school				Behind in school			
	Model 1: Entered school before required age [a]		Model 2: Completing more than required grades [b]		Model 3: One SD behind average for age [a]		Model 4 Dropped out [c]	
	OR	Se	OR	se	OR	se	OR	se
Both parents household members(omitted)								
One parent absent								
Mother member, Father not member	0.745	(0.058)**	0.616	(0.070)**	2.113	(0.184)**	1.772	(0.134)**
Father member, Mother not member	0.637	(0.077)**	0.637	(0.110)**	2.595	(0.324)**	2.209	(0.230)**
One parent deceased								
Mother member, Father deceased	0.761	(0.078)**	0.623	(0.076)**	1.622	(0.176)**	1.398	(0.135)**
Father member, Mother deceased	0.660	(0.140)+	0.434	(0.090)**	2.412	(0.459)**	1.353	(0.261)*
With neither parent								
Neither parent member	0.597	(0.058)**	0.227	(0.028)**	4.940	(0.481)**	3.367	(0.259)**
Mother deceased, Father not member	0.605	(0.086)**	0.249	(0.040)**	4.970	(0.639)**	2.657	(0.284)**
Father deceased, Mother not member	0.498	(0.080)**	0.238	(0.037)**	4.401	(0.609)**	3.523	(0.373)**
Both parents deceased	0.454	(0.088)**	0.274	(0.047)**	3.961	(0.643)**	2.332	(0.327)**
At least one female member over 50	0.778	(0.117) +	1.322	(0.201)+	1.330	(0.182)*	0.798	(0.122)
At least one male member over 50	1.069	(0.138)	1.188	(0.177)	0.912	(0.115)	1.058	(0.148)
At least one member of pension-eligible age	1.011	(0.113)	1.139	(0.169)	0.891	(0.096)	0.964	(0.104)
Observations	58,477		15,567		58,477		41,910	
Number of Children	37,731		13,468		37,731		28,330	

Data Source: Africa Centre Demographic Information System, KwaZulu-Natal, South Africa: Children aged 6 to 19, 2000-2004.

Notes: Constant term included. Statistical significance denoted: + p<.10; * p<.05; ** p<.01. The models control for child's sex, age, age squared, age order, whether he or she had multiple simultaneous household memberships, education of the most educated household member, household wealth quartile, Euclidean distance to schools, and year of observation fixed effects. It also controls for other household composition variables: elderly household members and number of household children. [a] All children aged 6 to 19 in 2000-2004 [b] All children aged 16 to 19 in 2000-200 [c] All children aged 6 to 16 in 2000-2004

Figure 1: Age-appropriate years of education and actual years of education for boys and girls ages 6 to 19, 2000-2004



Data Source: Africa Centre Demographic Information System, KwaZulu-Natal, South Africa: Children aged 6 to 19, 2000-2004.

Appendix 1: Matching Individuals to Households and Residences

Assembling the dataset necessitated matching individuals to households and membership episodes. Specifically, I identify the household in which a child was a member at the time of each observation. In most cases this is straightforward, however the ACDIS allows individuals to belong to multiple households at one time, which requires the researcher to make a judgment as to which is the primary household for the purpose of assigning household socio-economic and household composition measures. In these analyses, I focus on membership rather than residence as the main mechanism by which an individual is tied to households and to others within households. This is an informative approach because an individual's fortune is normally tied to the household that considers him or her to be a member, and it is this household that would invest in a child's education. Furthermore, individuals are generally resident in a household in which they are member. In future analyses, I will conduct sensitivity analysis to assess whether the results are sensitive to this choice.

For those individuals who were members of more than one household during the period of observation, the household in which they are members on each visit date is used. For example, a boy may have the following household membership patterns:

Table A.1: Sample format of an individual's membership episodes

Individual	Household	Membership start	Membership end
1	A	January 1, 2000 (First interview -left censoring)	February 9, 2001 (End of membership)
1	B	February 10, 2001 (Membership begins)	September 25, 2004 (Membership ends)
1	A	September 26, 2004 (Membership begins)	December 12, 2004 (Last interview –right censoring)

Let us say that the Household Socio-Economic Survey (HSE) rounds for the relevant households were collected on the following visits:

Table A.2: Sample format of household socioeconomic data collection rounds

Survey round	Household	Interview date
1	A	February 1, 2000
1	B	January 10, 2000
2	A	June 10, 2002
2	B	August 12, 2002
3	A	December 1, 2004
3	B	September 15, 2004

To get a full picture of the boy, I link his membership episodes to the most relevant socioeconomic characteristics. For this complicated individual, there would be a total of 4 observations. The child is first seen in household A on February 1, 2000, then in household B on August 12, 2002, again in household B on September 15, 2004, then in household A on December 1, 2004. To look at household composition, in this paperI use the same methods to determine which other individuals – parents, older individuals, other children -- are members of each household at each HSE round. The panel is created through repetition of these steps.

Table A.3: Sample panel of individual records

Individual	Interview date	Household
1	February 1, 2000	A
1	August 12, 2002	B
1	September 15, 2004	B
1	December 1, 2004	A

The possible complication for individuals such as this one is that the education level of this boy will not have changed between September and December 2004, but households A and B might have different circumstances. It is unclear which household – or what combination of the two – is actually affecting the child's access to education. Most of the children in the data have simpler life histories, but some are even more complicated.

For those individuals who were members of multiple households simultaneously, one household is selected randomly. A dummy variable is included to denote that the person was a member of multiple households simultaneously, with the expectation that children with multiple memberships may be different, both in their personal characteristics and because the circumstances of the household selected as their main household are not the only ones that affect them.

Appendix 3: More details on the other Explanatory Variables

The other explanatory variables capture characteristics of the child and the household. The child-specific variables are the child's sex, age, age order among household children, and whether the child was ever a member of multiple households simultaneously. Age and age order vary over time.

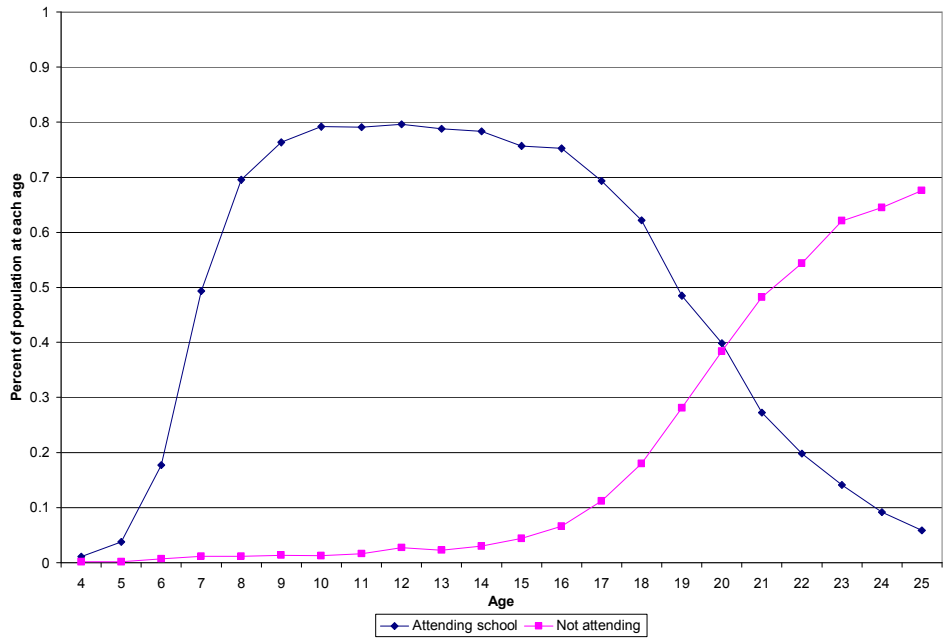
Age order in terms of age is the rank of the child among all household members younger than 20. The oldest individual under the age of 20 is coded as 1 and the youngest as 1+n. Age order may affect access to education in that younger children have to compete for household resources with more children than older children did at the same age. On the other hand, the marginal cost of sending younger children to school may be lower because they can make use of the uniforms, books, and transportation arrangements of older siblings. This measure is a generalization of the more commonly used birth order (see Behrman and Taubman 1986 for a review of how birth order might affect schooling). Age order may be a more informative measure than birth order in this setting because, in a household with several women of child-bearing age, the household must consider the needs of all member children together, rather than considering the children of each mother separately. Finally, some children may have siblings who are deceased or are independent and so are no longer included in the household's economic allocation process.

Some children in the area are considered members of several households. They may split their time between maternal and paternal households or between parents' and grandparents' households and be considered a member by the relatives in each household. As this is likely to affect educational attainment, either by providing

additional resources for school from multiple households or by disrupting regular school attendance, I include a dummy variable indicating that the child has multiple simultaneous household memberships. For the analyses, the child is randomly assigned to one of the households in which he or she is a member so that each child appears only once in the dataset at any given time period. Information on other members and other household variables pertain to the household to which the child is assigned.

I include age and age-squared variables because we can expect an age-specific pattern of school attainment, with children becoming enrolled by the time they reach age 7 and progressing until they drop out or complete secondary school. Children tend to fall behind with age because they are repeatedly exposed to the risk of failing a grade and because they face increasing opportunity costs in terms of formal or domestic employment as they mature. Based on the age-pattern of enrollment in the study area (shown in Figure 4.3) I only include individuals through age 19. While there are still individuals enrolled in school past that age, 19 is the age at which individuals should have completed secondary school if they progressed on or close to schedule and it is indeed the age at which a large number of individuals leave school.

Figure 3: Proportion of individuals aged 4 to 25 attending school full or part time



Data Source: Africa Centre Demographic Information System, KwaZulu-Natal, South Africa: Household Socio-Economic survey 1, 2000-2002.
Note: Proportions attending and not attending do not necessarily add to 1 because information on school attendance is missing for some individuals.

Gender is included in most models, and one model includes interaction terms between gender and the household composition variables. Access to education has often been documented to be more difficult for girls than for boys, especially in less developed countries.

In addition to child-specific variables, I also include household-specific variables. These variables are potentially time-varying as households acquire new members and assets or move. As discussed above, in order to maintain a more inclusive sample of children, I do not include parents' characteristics, such as education level.[20] Instead, I

[20] Survival status is known for 99% of mother and 84% of fathers, but detailed information about education, partnerships and presence is known for only about 40% of the fathers and 70% of mothers. I have also estimated models using only those children for whom at least the mother is known. In these

include a variable denoting the education level of the most educated individual in the household.

A measure of income or wealth is generally included in models of education because there is abundant evidence that income and wealth are associated with and determinants of child education (a thorough study focusing on this is Behrman and Knowles 1999). Households with higher wealth and income can support child school attendance: wealthier families can pay the school fees, buy uniforms and supplies and do not depend on the child's labor. They should also be less vulnerable to the need to allocate scarce resources towards the most productive members and should be less affected by negative shocks. Information on household income is not available in the data. In this area with high levels of unemployment and widespread informal work, such information would not be very useful: only 30% of women and 42% of men report that they are employed or earning money. Instead of income, I use information on the resources owned by the household as a proxy for wealth. I create an index of household wealth using principal component analysis. The index is based on the factor loadings of variables indicating house construction materials; household amenities like water source, toilet, and electricity; ownership of commodities; and farm production. The index is calculated separately for 2000-2002 and 2003-2004. The variables used to create the index are shown in Appendix 4.4. The wealth quartile is created by segmenting this index into quartiles. For details on the method, see Filmer and Pritchett (2001) and Dunteman (1989).

---

models, I was able to include mother's education (in place of the education of the most educated household member) and birth order (in place of age order). The results are similar to those presented in the paper.

Distance to school can be an important factor in attendance: children who live far away from a school may find it more difficult or costly to attend, having to hitchhike, pay taxi fare, or walk long distances to get to school and back every day. Distances to schools are calculated, using MapInfo software, as the linear, or Euclidean, distance based on GIS-coded locations of households and schools. I include the number of kilometers from the household to the nearest primary and the nearest secondary school.[21] While a child would clearly only attend primary or secondary school at any given point in time, distance to primary school affects whether the child stays enrolled long enough to make it to secondary school.

Finally, I control for some aspects of time. A series of dummy variables for each year of observation (2000-2004) removes purely inter-temporal variation in educational attainment rates. The other is a term indicating whether the child was eligible to be enrolled in school before or after 1998, the year in which a series of changes in age at school enrollment began.[22]

[21] This is only an approximation of the distance the child must actually travel to school. Calculated as a straight line between the homestead and the school, it provides a lower boundary of the actual distance.

[22] In addition to year-specific effects, there may be cohort effects in education, so that children born in some years are exposed to different circumstances from those born in other years. In an alternative specification, I substituted observation year with birth year. The substitution is necessary since the model would be over-specified if I included birth year, observation year and age. Results are consistent with those shown.

Appendix 3: Principal Component Analysis

A measure of income or wealth is generally included in models of education because there is abundant evidence that income and wealth are associated with and determinants of child education (a thorough study focusing on this is Behrman and Knowles 1999). Households with higher wealth and income can support child school attendance: wealthier families can pay the school fees, buy uniforms and supplies and do not depend on the child's labor. They should also be less vulnerable to the need to allocate scarce resources towards the most productive members and should be less affected by negative shocks. Information on household income is not available in the data. In this area with high levels of unemployment and widespread informal work, such information would not be very useful: only 30% of women and 42% of men report that they are employed or earning money. Instead of income, I use information on the resources owned by the household as a proxy for wealth. I create an index of household wealth using principal component analysis. The index is based on the factor loadings of variables indicating house construction materials; household amenities like water source, toilet, and electricity; ownership of commodities; and farm production. The index is calculated separately for 2000-2002 and 2003-2004. The variables used to create the index are shown in Appendix 4.4. The wealth quartile is created by segmenting this index into quartiles. For details on the method, see Filmer and Pritchett (2001) and Dunteman (1989).

We use principal component analysis to create an index of household wealth.[23] Many household resources are highly correlated, so including all of them as explanatory variables in the y_h vector of a regression would mask the effects of at least some of the variables. The index of household wealth is based on variables identifying house quality, objects owned, equipment and farm production, heating, electricity, and sanitation. A full list of the variables used, as well as their means and loadings, are shown below. The original variables indicate whether the household has each item and a count of the number of each if it is possible to have more than one. First, I convert all variables into dummy variables indicating whether the household had any of each resource, event, or activity. The exception is the number of rooms, which I leave as a count. The components created with principal component analysis are ordered in terms of the variance which they capture. For example, the principal component y_n is a linear combination of the original standardized variables x_n to $x_{p,}$ weighted by an optimal weight vector a_{n1} to a_{np}, which sums to 1. In this fashion, the variance captured by y_n is maximized:

$$y_n = a_{n1}x_1 + a_{n2}x_2 + ...a_{np}x_p = \sum_{i=1}^{p} a_{ni}x_i \qquad (2.3)$$

Principal component loadings are then calculated by multiplying the latent vectors (or eigenvectors) by the square root of the associated variance (eigenvalues).

To create the wealth index, I use the first principal component, which explains the greatest share of the variance (in this case 18%), to rank the households in terms of

[23] This discussion draws on Dunteman (1989) and Stata Corp (2003).

wealth, following Maluccio et al. (2004) and Filmer and Pritchett (2001). In my models, I use this index as a continuous variable, and also include a quadratic term to take account of nonlinearities in the associations.

Table A.4.4: Variables used in principal component analysis to create index of household wealth (continued)

	2000-2002				2003-2004		
Characteristic	Mean	Std. Dev.	First principal component loadings	Characteristic	Mean	Std. Dev.	First principal component loadings
Wheelbarrow	0.422	0.494	0.001	Wheelbarrow	0.444	0.497	0.019
Furniture	0.893	0.309	0.069	Furniture	0.937	0.243	0.099
Work vehicle	0.025	0.155	0.072	Work vehicle	0.035	0.184	0.024
Gas	0.095	0.293	0.045	Gas	0.075	0.264	0.038
Wood	0.400	0.490	-0.239	Wood	0.446	0.497	-0.262
Solar power	0.002	0.043	0.008	Solar power	0.002	0.050	0.017
Coal	0.003	0.059	-0.002	Coal	0.003	0.053	-0.006
Electricity	0.242	0.428	0.300	Electricity	0.250	0.433	0.306
Other power source	0.001	0.035	0.003	Other power source	0.001	0.038	-0.002
Generator	0.003	0.051	0.017	Generator	0.006	0.079	0.032
Paraffin	0.257	0.437	-0.052	Paraffin	0.213	0.409	-0.037
Sewing machine	0.158	0.364	0.064	Sewing machine	0.142	0.349	0.072
Cooker	0.717	0.451	-0.175	Cooker	0.661	0.473	-0.162
Radio	0.596	0.491	-0.083	Radio	0.726	0.446	0.117
Cooker oven	0.145	0.352	0.252	Cooker oven	0.182	0.386	0.265
Hotplate	0.218	0.413	0.184	Hotplate	0.223	0.416	0.171
Kettle	0.233	0.423	0.291	Kettle	0.249	0.432	0.290
Refrigerator	0.419	0.493	0.278	Refrigerator	0.482	0.500	0.287
Gas cooker	0.435	0.496	0.142	Gas cooker	0.384	0.486	0.112
Farm tools	0.781	0.414	-0.072	Farm tools	0.801	0.399	-0.052
Automobile	0.110	0.313	0.143	Automobile	0.134	0.341	0.163
Bike	0.119	0.324	-0.003	Bike	0.095	0.293	0.003
Block maker	0.104	0.306	0.016	Block maker	0.111	0.314	0.034

Data Source: Africa Centre Demographic Information System, KwaZulu-Natal, South Africa: Household Socio-Economic surveys 1 and 2, 2000-2004
Notes: The first principal component for 2000-2002 has eigenvalue of 6.22 and accounts for 13.2% of the variance. The first principal component for 2003-2004 has eigenvalue of 5.78 and accounts for 12.3% of the variance.