THE EFFECT OF HIV/AIDS ON EDUCATIONAL ATTAINMENT

Jane G. Fortson*
Becker Center on Chicago Price Theory
University of Chicago

This Version: October 2007
First Draft: September 2007

Abstract

Using data from Demographic and Health Surveys for eleven countries in sub-Saharan Africa, I estimate the effect of local HIV prevalence on individual human capital investment. I find that the HIV/AIDS epidemic has reduced human capital investment: living in an area with higher HIV prevalence is associated with lower levels of completed schooling and slower progress through school. These results are consistent with a model of human capital investment in which parents and children respond to changes in the expected return to schooling driven by mortality risk.

* *Email:* jfortson@uchicago.edu
Address: Becker Center on Chicago Price Theory, 5807 S. Woodlawn Avenue, Chicago, IL 60637

Background paper prepared for meeting of the World Bank/UNAIDS Economics Reference Group, 8 & 9 November 2007. Much of the text is excerpted from an earlier paper, "Mortality Risk and Human Capital Investment: The Impact of HIV/AIDS in Sub-Saharan Africa" (Fortson, 2007a). In addition to the acknowledgements cited there, I am grateful to Anne Case for her continuing guidance, UNICEF for helpful comments, and to Daniel Hirschberg for excellent research assistance.

I. INTRODUCTION

Twenty-five years after AIDS was first reported by the CDC in 1981, 38.6 million people worldwide are estimated to be infected with HIV. Sub-Saharan Africa, the "global epicenter of the AIDS pandemic," is home to over 60 percent of HIV-infected people, with prevalence among adults ages 15-49 estimated to be 6.1 percent (UNAIDS, 2006). AIDS killed some two million Africans in 2005, and has reduced life expectancies in the most affected countries to their pre-1970 levels (US Census Bureau International Data Base, 2005). And yet the impact of the AIDS epidemic extends far beyond its first-order effect on longevity. In some areas, AIDS has wiped out a substantial fraction of prime-age adults, threatening economic stability and growth today and in the coming decades. One important dimension along which HIV may influence economic conditions is through its effect on human capital investment.

Most existing estimates of the effect of HIV on human capital investment have focused on orphans. Orphans in Africa are significantly less likely to be enrolled in school than non-orphans, and progress more slowly when enrolled (e.g., Case, Paxson, and Ableidinger, 2004; Case and Ardington, 2006; Evans and Miguel, 2007).[1] However, the HIV/AIDS epidemic may well affect human capital investment more broadly. For instance, changes in mortality could have a non-negligible effect on human capital investment (e.g., Meltzer, 1992). Because the presence of HIV/AIDS may increase mortality risk in adulthood, it could *ceteris paribus* reduce the expected return to schooling. To the extent that human capital investment responds to changes in the expected return to schooling, we might

[1] While a great deal of work finds adverse schooling effects of orphanhood, not all studies conclude that differences between orphans and non-orphans are substantial (e.g., Ainsworth and Filmer, 2002).

think that human capital investment would fall among both orphans and non-orphans.[2] Through this and other channels, the HIV/AIDS epidemic may have a widespread effect on human capital investment and, therefore, growth.

The contribution of this paper is twofold. First, I provide estimates of the overall impact of the HIV/AIDS epidemic on educational attainment. Second, I evaluate empirically several competing hypotheses for why HIV might affect human capital investment.

In my principal empirical tests, I use nationally-representative cross-sections from Demographic and Health Surveys (DHS) for Burkina Faso, Cameroon, Ethiopia, Ghana, Kenya, Lesotho, Malawi, Mali, Senegal, Tanzania, and Zambia to estimate the relationship between local HIV prevalence and the educational outcomes of successive birth cohorts. These surveys include newly-available estimates of HIV based on household testing – arguably the most accurate and geographically-specific estimates available. Using these cross-sectional data on local HIV prevalence in a difference-in-differences approach, my analysis exploits the substantial geographic and time variation in HIV prevalence to estimate the effect of HIV on human capital investment. I find that years of schooling, school attendance, primary school completion, and progress through school declined in places with higher levels of HIV, though they increased elsewhere. Relative to areas without HIV, post-1980 birth cohorts in areas with HIV prevalence of ten percent (today) received about 0.5 fewer years of schooling than pre-1980 cohorts.

I evaluate seven possible channels through which the epidemic could affect schooling: orphanhood, caretaking requirements, the provision of schooling, family

[2] There are indeed numerous other channels through which HIV could influence schooling. A related hypothesis, pursued by Kalemli-Ozcan (2006), is that HIV jointly influences fertility and human capital decisions, and that, as a form of insurance, parents may have more children but invest less in their education (a "quality-quantity trade-off").

resources, market wages, child-headed households, and mortality risk. These results provide inconclusive but suggestive evidence that the effect of local HIV prevalence on human capital investment is driven in large part by responses to declines in the expected return to schooling resulting from increases in mortality risk.

II. DATA

My empirical analysis uses nationally-representative cross-sections from Demographic and Health Surveys (DHS) for eleven countries in sub-Saharan Africa: Burkina Faso (2003), Cameroon (2004), Ethiopia (2005), Ghana (2003), Kenya (2003), Lesotho (2004), Malawi (2004), Mali (2001), Senegal (2005), Tanzania (2003), and Zambia (2001-2002).[3] In addition to providing demographic, economic, and fertility microdata, these waves of the DHS are linked to HIV test results which can be used to calculate local HIV prevalence.

HIV Prevalence.

In these eleven DHS surveys, some adult respondents to the survey were asked to provide blood samples for HIV testing.[4] I calculate local HIV prevalence as the adult HIV rate in each sector of residence (urban or rural) within each region (generally defined by a government administrative division). Table I.A shows the considerable variability in local HIV prevalence across and within countries.

Estimates of HIV prevalence from other sources are generally based on the testing of sub-populations (e.g., pregnant women or commercial sex workers). Because DHS HIV testing was conducted at the household level, respondents should be more representative of

[3] Source: DHS datasets, www.measuredhs.com, MEASURE DHS, Macro International Inc.
[4] The data appendix provides detail on sample selection and testing procedures.

area residents than other HIV testing samples. Not only are DHS HIV testing data representative, but they provide estimates of prevalence at a fine level of geographic detail. National HIV prevalence data, in contrast, obscure the substantial geographic variation within country.

However, local HIV rates could be measured with error. For about 19 percent of respondents, HIV test results are missing – primarily because of test refusal. Because HIV status may well be related to HIV test refusal and because data are missing at different rates across sectors and regions, missing HIV test data could bias my estimates of local HIV prevalence. However, the magnitude of any bias would most likely be quite small: Mishra et al. (2006) show that such non-response has an insignificant effect on national HIV prevalence estimates using these data.

All the same, local HIV prevalence data have their limitations. Sampling variability could introduce classical measurement error, attenuating my estimates. More importantly, estimates of local HIV prevalence over time are unavailable. My estimates of local HIV prevalence from DHS testing are applicable only to the survey year. To address this shortcoming, I adopt a difference-in-differences approach, comparing changes over time (across birth cohorts) across areas with different levels of HIV (in the survey year).

Individual Characteristics.

Data on age, educational attainment, and other individual characteristics are drawn from responses to the individual and household questionnaires of the DHS. For the most part, my analysis uses a sample of adults ages 15-49. I consider three educational outcomes for adults: years of schooling, school attendance (completed one or more years), and primary school completion.[5] In Section IV, I use a sample of children ages 7-14 to explore explanations for my main results. Because schooling for children may be incomplete, I use progress through school (years / (age – 6)) as an additional measure of schooling attainment for this younger group.

Child characteristics, including progress through school, are summarized in Table I.B while summary statistics for adult characteristics, including all three measures of education, are presented in Table I.C. (Survey sampling procedures, response rates, and variable definitions are described in detail in the data appendix.) In the countries under study, about three-quarters of people reside in rural areas. On average, adults have completed 4.2 years of schooling, and a little over half report having completed at least one year of school. Thirty-six percent of adults have at least a primary school education.

My data include information about the respondent's current region and sector of residence, but only limited information about his childhood residence. Because my ability to identify the effect of HIV on human capital investment relies on the assumption that adults living in a particular area are representative of those living in that area in childhood, migration (and mortality, for that matter) could be a source of bias. For simplicity, I assume that individuals have not moved from their childhood region/sectors of residence. Fortson (2007a) shows that the results are not driven by migration or mortality.

[5] I use these three educational outcomes to limit concern that incomplete schooling could bias my results. However, only 14 percent of adults ages 25-49 have more than 10 years of schooling, so any bias should be small. As an additional test, I top-code years of schooling at 10; the results are quite similar. Using a number of outcomes should also assuage concerns that my estimates are sensitive to transformations of the dependent variable (Meyer, 1995).

III. EMPIRICAL MODEL

My approach exploits geographic and time variation in HIV prevalence. Using fixed effects for region/sector of residence, I estimate the effect of local HIV prevalence on three educational outcomes: completed years of schooling, school attendance, and primary school completion. In particular, I estimate the following model:

$$S_{icr} = \beta_0 + \beta_1 HIV_r \times I(c \geq 1980) + \beta_2 F_{icr} + \gamma_c + \alpha_r + \varepsilon_{icr} \qquad (1)$$

where $S_{icr} =$ educational outcome for respondent i in cohort c and region/sector r,

$HIV_r =$ HIV prevalence in region/sector r in survey year,

$I(c \geq 1980) =$ indicator for whether respondent i was born in or after 1980,

$F_{icr} =$ indicator for whether respondent i in cohort c, region/sector r is female,

$\gamma_c =$ fixed effect for year of birth c, and

$\alpha_r =$ fixed effect for region/sector r.

γ_c allows for a flexible trend in educational outcomes over time, and α_r accounts for underlying differences in educational outcomes across region/sectors of residence. (Note that the level effects of current HIV rates will be absorbed by the region/sector indicators.) My sample includes adults ages 15-49 in the survey year, and is weighted using DHS-provided weights that account for the sampling and response probabilities. I compute Huber-White standard errors that adjust for possible correlations of outcomes within each region/sector.[6]

[6] If there is serial correlation in the outcome variable – which is quite possible for schooling – conventional difference-in-differences techniques tend to over-reject a null hypothesis of no effect. While I do not adopt a true difference-in-differences approach, my standard errors could still be plagued by this problem. Clustering that accounts for intraregional auto-correlation has been found to be an effective solution in some applications when the number of regions is sufficiently large (Bertrand, Duflo, and Mullainathan, 2004).

This is a difference-in-differences framework, comparing differences in educational outcomes of unaffected and affected cohorts across regions with differing levels of treatment intensity (HIV prevalence).[7] Region/sector fixed effects control for underlying differences in educational outcomes across regions. Unlike a simple difference-in-differences approach, I allow for a flexible time trend in schooling outcomes and control for sex. Difference-in-differences results without controls (not shown) are quite similar to these results.

This approach amounts to assuming that HIV had no effect on the educational outcomes of cohorts born before 1980, and a constant effect on cohorts born in or after 1980. My sample of adults ages 15-49 covers birth cohorts 1951-1989. Respondents in the unaffected (1951-1979) cohorts would have, for the most part, completed their schooling by the late 1980s, before HIV was widespread. Respondents in the affected (1980-1989) cohorts reached school-going age in the late 1980s and 1990s, when HIV was more prevalent.

Figure 1 shows the relationship between HIV prevalence (in the survey year) and the difference in mean educational attainment between younger (post-1980) and older (pre-1980) birth cohorts, where each observation represents a region/sector. The population-weighted least squares estimate suggests that there is indeed a negative relationship between HIV and the change in educational attainment.

Table II shows estimates of the difference-in-differences in completed years of schooling, separately by country. In all countries but Malawi and Senegal, the difference-in-differences estimate is negative; because there are only a small number of region/sectors

[7] This requires the assumption that relative prevalence is stable over time. While it would be possible to use a more disaggregated measure of HIV prevalence – such as prevalence at the region/sector/gender level – doing so would require the additional assumption that relative prevalence has also been stable by gender, which we do not believe to be the case (e.g. UNAIDS and WHO, 1999; Oster, 2006).

within each country, the effect of HIV is not precisely estimated. Therefore, I pool results from all eleven countries, weighting each country by its population. Table III shows results from this pooled specification. HIV has a negative and significant impact on completed years of schooling, attendance, and primary school completion. Relative to a base case of no HIV, local HIV prevalence of 10 percent is associated with a decline in schooling of about 0.5 years, a 5 percentage point decrease in the probability of attending school, and a 6 percentage point fall in the probability of completing primary school. These results suggest that educational outcomes improved less over time in areas harder hit by HIV.

Because levels of completed schooling are low in these countries, these effects are in fact quite large: between pre- and post-1980 birth cohorts, areas with local HIV prevalence of ten percent (relative to areas without HIV) experienced a 12 percent reduction in completed years of schooling, a 9 percent reduction in the probability of attending school, and a 15 percent reduction in the probability of completing primary school.

IV. CHANNELS

Section III shows evidence of a robust negative relationship between local HIV prevalence and human capital investment. The observed link may, however, arise through numerous channels. Local HIV prevalence may be related to the probability of orphanhood, which has been shown to adversely affect educational outcomes. In addition, local HIV prevalence might reduce completed schooling because it affects the supply of teachers, economic wellbeing of families, responsibilities of children (to care for sick relatives), or the health of children (UNAIDS, 2002). Here, I test empirically for the influence of six possible determinants of human capital investment: orphanhood, caretaking requirements, the provision of schooling, family resources, market wages, and child-headed households.[8] Then, I discuss the possibility that the decline in human capital investment is driven by responses to the increased mortality risk due to HIV.

Orphanhood.

As noted in the introduction, several previous studies have shown that orphans are less likely to be enrolled in school than non-orphans. In theory, the negative relationship that I estimate between HIV and educational outcomes could reflect the fact that orphans have a greater likelihood of living in high-HIV areas. Using the difference-in-differences estimation strategy from Section III, I test whether the effect of local HIV prevalence on educational outcomes is driven by orphans. In particular, I compare the effect of local HIV prevalence over time between a sample of all children and a sample of non-orphans alone. Because orphan status is known only for sample children (but not adults), I estimate the effect of local HIV prevalence on a measure of progress through school (grades completed relative to potential) for children ages 7-14. My approach compares changes in progress between pre-1992 and post-1992 birth cohorts across areas with different levels of HIV. Table IV presents the results for the full sample of children, which shows that local HIV prevalence is associated with declines over time in progress. Relative to areas without HIV, later cohorts in areas with prevalence of ten percent complete about 0.03 fewer grades per

---

[8] We might also expect worsening child health to reduce schooling. While HIV may have had a substantial impact on child health (particularly through mother-to-child transmission of the virus), the fixed effects estimates reported in Section III likely reflect other channels. Children ill with HIV may be less likely to attend school or may progress more slowly through school. But because the effects in Section III are estimated on a sample of adults born before 1990, very few of these respondents would have been infected at birth, and any who were infected would be unlikely to survive to adulthood. And because the primary post-infancy mode of transmission in Africa is heterosexual sex, it is unlikely that infection after birth would directly affect completed years of schooling, especially since schooling levels are so low.

year.[9] Excluding double orphans (children who lost both parents), the effect of HIV on changes in progress through school is nearly unchanged. When I exclude both single and double orphans, again the estimate is negative and significant, and of the same magnitude as the comparison results. The results in Table IV indicate that the effect of HIV of human capital investment is not limited to orphans – there are in fact much broader effects of the HIV/AIDS epidemic.

Caretaking Requirements.

Children may also miss school to care for sick family members, contributing to the large negative effect of local HIV prevalence on schooling. To test whether this is driving my results, I exclude from the analysis children belonging to households known to have HIV-infected members (based on the DHS HIV test). Because I do not know HIV test results at the household level for Mali and Zambia, the sample is restricted to households undergoing HIV testing in Burkina Faso, Cameroon, Ethiopia, Ghana, Kenya, Lesotho, Malawi, Senegal, and Tanzania. I then test, on this sample of non-orphans with no known HIV-infected household members, whether the local HIV rate is negatively and significantly associated with progress through school. The results, shown in Table V, indicate that my estimates are not driven by family caretaking requirements, as the effect of HIV on changes in progress is negative and significant in this restricted sample.

Provision of Schooling.

Another possibility is that reductions in the provision of schooling, perhaps driven by teacher death, are responsible for the estimated effect in Table III. To estimate the

---

[9] This difference is rather small. However, because children ages 7-14 born before 1992 (with birth years 1987-1991) may also have been affected by local HIV prevalence, this is likely to be an underestimate of the effect of HIV. The estimate, the difference-in-differences between pre- and post-1992 cohorts, instead shows us how changes in HIV over this period are related to changes in progress through school.

contribution of changes in provision to the decline in human capital investment, I test whether the effect of local HIV prevalence on completed schooling differs for men and women. Though there are differences in educational attainment between men and women, these are small enough that we would not expect the effect of HIV to differ much if provision were driving the effect. The results, shown in Table VI, show that local HIV prevalence has a significantly larger (negative) effect on the educational outcomes of men: the effect of HIV on men's primary school completion is over seven times as large as the effect for women, even though men are only 35 percent more likely to have completed primary school. If provision were the primary channel through which HIV reduced schooling, we would not expect to see a large differential by sex.[10] In fact, the results show that the effect of local HIV prevalence on educational attainment operates almost entirely through its effect on men's educational outcomes.[11]

Family Resources.

If local HIV prevalence reduces family resources even among households without illness or death, this could in theory drive the estimated decline. However, if this were the case, we would expect to see comparable declines in schooling for men and women, assuming that the returns to schooling are likewise comparable (as indicated by Schultz, 2004).[12] However, as shown in Table VI, there are substantially larger reductions in

[10] Furthermore, recent survey data for Uganda show that only 1.8 percent of parents cited poor school quality as a reason for why their children dropped out of primary school (Uganda Bureau of Statistics and ORC Macro, 2001). If provision changes were a primary determinant of the decline in schooling, we would expect a greater proportion of parents to cite school quality as a reason for attrition.

[11] In results not shown, I find that these results are robust to the inclusion of gender-by-birth-cohort fixed effects.

[12] If parents' preferences prioritize boys before girls when it comes to schooling, we might expect to find differences by sex. But in this situation, we would expect girls' schooling to fall by more than boys' schooling – the opposite of what we find.

completed schooling for men. These results suggest that the effects of HIV on family resources do not explain a large part of the estimated decline.

Market Wages.

It has been proposed that increases in mortality due to AIDS may drive up wages among survivors (Young, 2005). If this is the case, it could be that the relatively larger declines in schooling in high-HIV areas represent children's responses to increased wages. And while I find different effects of HIV by sex, this is not necessarily inconsistent with this hypothesis. If boys and girls are not substitutes in the labor market, then the greater educational declines among men could reflect a relatively larger increase in men's wages. However, if changes in wages are driving the schooling decline, we might not expect to find differences in the effect of HIV across other subgroups. For example, Muslims and non-Muslims are arguably substitutes in the labor market, but may have very different expectations of the risk of contracting HIV in adulthood. In results not shown, I find much larger HIV-related declines in educational attainment among non-Muslims than among Muslims.[13] Such large differences in the effect of HIV by religion would be unlikely to arise if market wages were the primary channel through which HIV reduced schooling.

Child-Headed Households.

Another possibility is that HIV is associated with the probability of living in a child-headed household, which in turn affects educational investment. The sample used in Tables II and III includes adults that grew up in child-headed households. However, if local HIV prevalence affects educational attainment through household composition, I would not

[13] These data show that Muslim men are more likely to be circumcised, which should reduce their risk of contracting HIV. However, within a region/sector, Muslims and non-Muslims do not show statistically different probabilities of HIV infection, controlling for sex.

expect to find substantial differences in the effect by sex, as I do in Table VI. Furthermore, child-headed households are not sampled in the DHS (though adults who grew up in child-headed households are sampled). Therefore, if local HIV prevalence affects educational attainment through child-headed households, I would not expect to find an effect of HIV among children in the sample, as I do in Tables IV and V.

Mortality Risk.

However, there is one possible channel that these tests cannot rule out. Local HIV prevalence, through its effect on expected longevity, may influence human capital investment. Completing an additional year of schooling delays the age at which earnings commence but increases the value of those earnings; thus, individuals face a tradeoff between present and future consumption. By lowering the probability of survival, HIV lowers the value of future consumption. Hence, changes in mortality risk, such as those generated by changes in local HIV prevalence, could imply a decline in educational investment. The results presented here are not inconsistent with this channel. For instance, in Table VI, I find a much larger effect of HIV on educational attainment among men, consistent with the higher mortality risk among men in the early part of the HIV/AIDS epidemic (UNAIDS and WHO, 1999; Oster, 2006).

VI. CONCLUSION

Exploiting regional and time variation in local HIV prevalence, I estimate the effect of the local HIV rate on human capital investment. I find that local HIV prevalence is associated with substantially worse educational outcomes: cohorts born after 1980 in areas with prevalence of 10 percent today completed about 0.5 fewer years of schooling than pre-1980 cohorts, relative to areas without HIV. Children in areas with higher levels of HIV fare worse along a number of dimensions – they are less likely to attend school, less likely to complete primary school, and progress more slowly through school. These results are robust to numerous sensitivity checks, including tests for omitted variable bias, sample selection bias, and attrition (see Fortson, 2007a).

I extend the analysis to empirically distinguish between channels from HIV to education. I evaluate six channels – orphanhood, caretaking requirements, provision, family resources, market wages, and child-headed households – and conclude that none of these drive the effect. My empirical findings are consistent with the idea that parents reduce their children's schooling because HIV lowers its payoffs. These results suggest that, when making schooling decisions, parents and children respond to changes in the expected return to investment due to mortality risk. My results provide some hope that expanded access to antiretroviral therapy would increase longevity among the HIV-infected and thus mitigate this effect.[14]

But absent marked improvements in longevity, the HIV/AIDS epidemic may impose a substantial economic cost. Based on these estimated effects and overall HIV prevalence figures, I would expect average eventual educational attainment among children today in sub-Saharan Africa to be about 0.3 years fewer than it would have been in the absence of HIV, with much larger declines in areas hardest hit by HIV. Moreover, HIV has not only endangered human capital investment and GDP (Krueger and Lindahl, 2001), but it has deprived a large fraction of adults of their productive prime, a disproportionate fraction of them relatively well-educated (Fortson, 2007b). The substantial reductions in human capital investment estimated here, coupled with the high levels of mortality among adults

---

[14] Indeed, new longitudinal evidence from Kenya shows that the availability of antiretrovirals increases school-going among children whose parents receive treatment (Graff Zivin, Thirumurthy, and Goldstein, 2006).

today, have reduced both the stock and flow of human capital in the region, contributing to sub-Saharan Africa's economic woes.

REFERENCES

Ainsworth, Martha and Deon Filmer (2002), "Poverty, AIDS, and Children's Schooling: A Targeting Dilemma," World Bank Policy Research Working Paper 2885.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004), "How Much Should We Trust Difference-in-Differences Estimates?" *Quarterly Journal of Economics*, 119(1), pp. 249-275.

Case, Anne, and Cally Ardington (2006), "The Impact of Parental Death on School Outcomes: Longitudinal Evidence from South Africa," *Demography*, 43(3), pp. 401-420.

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

Evans, David K., and Edward Miguel (2007), "Orphans and Schooling in Africa: A Longitudinal Analysis," *Demography*, 44(1), pp. 35-57.

Fortson, Jane (2007a), "Mortality Risk and Human Capital Investment: The Impact of HIV/AIDS in Sub-Saharan Africa," working paper, University of Chicago.

Fortson, Jane (2007b), "The Gradient in Sub-Saharan Africa: Socioeconomic Status and HIV/AIDS," working paper, University of Chicago.

Graff Zivin, Joshua, Harsha Thirumurthy, and Markus Goldstein (2006), "AIDS Treatment and Intrahousehold Resource Allocations: Children's Nutrition and Schooling in Kenya," NBER Working Paper 12689.

Kalemli-Ozcan, Sebnem (2006), "AIDS, Reversal of the Demographic Transition and Economic Development: Evidence from Africa," NBER Working Paper 12181.

Krueger, Alan B., and Mikael Lindahl (2001), "Education for Growth: Why and For Whom?" *Journal of Economic Literature*, 39(4), pp. 1101-1136.

Meltzer, David (1992), "Mortality Decline, the Demographic Transition, and Economic Growth," Ph.D. Dissertation, University of Chicago.

Meyer, Bruce D. (1995), "Natural and Quasi-Experiments in Economics," *Journal of Business & Economic Statistics*, 13(2), pp. 151-161.

Mishra, Vinod, Martin Vaessen, J. Ties Boerma, Fred Arnold, Ann Way, Bernard Barrere, Anne Cross, Rathavuth Hong, and Jasbir Sangha (2006), "HIV Testing in National Population-Based Surveys: Experiences from the Demographic and Health Surveys," *Bulletin of the World Health Organization*, 84(7), pp. 537-545.

Oster, Emily (2006), "Measuring the Magnitude of HIV in Africa," working paper, University of Chicago.

Rutstein, Shea Oscar and Guillermo Rojas (2003), *Guide to DHS Statistics*. Calverton, MD: Demographic and Health Surveys, ORC Macro.

Schultz, T. Paul (2004), "Evidence of Returns to Schooling in Africa from Household Surveys: Monitoring and Restructuring the Market for Education," *Journal of African Economies*, 13(AERC Supplement 2), pp. ii95-ii148.

Uganda Bureau of Statistics and ORC Macro (2001), *Uganda DHS EdData Survey 2001: Education Data for Decision-Making*. Calverton, Maryland, USA: Uganda Bureau of Statistics and ORC Macro.

UNAIDS (2002), *2002 Report on the Global HIV/AIDS Epidemic*. Geneva, Switzerland: UNAIDS.

UNAIDS (2006), *2006 Report on the Global AIDS Epidemic*. Geneva, Switzerland: UNAIDS.

UNAIDS and WHO (1999), *AIDS Epidemic Update: December 1999*. (Joint United Nations Programme on HIV/AIDS and World Health Organization), Geneva, Switzerland: UNAIDS and WHO.

US Census Bureau International Data Base (2005), U.S. Census Bureau, Population Division, International Programs Center, Data Update April 26, 2005.

Young, Alwyn (2005), "The Gift of the Dying: The Tragedy of AIDS and the Welfare of Future African Generations," *Quarterly Journal of Economics*, 120(2), pp. 423-466.

DATA APPENDIX

Data for the analysis come from Demographic and Health Surveys (DHS), which are available from ORC Macro (http://www.measuredhs.com). The sample includes eleven countries: Burkina Faso, Cameroon, Ethiopia, Ghana, Kenya, Lesotho, Malawi, Mali, Senegal, Tanzania, and Zambia. The analysis here uses only the most recent wave of the DHS for each country (2001 for Mali; 2001/2002 for Zambia; 2003 for Burkina Faso, Ghana, Kenya, and Tanzania; 2004 for Cameroon, Lesotho, and Malawi; and 2005 for Ethiopia and Senegal).[15] However, in some robustness checks in Fortson (2007a), I use previous waves of the data.

The DHS has several survey components, including a household questionnaire, women's questionnaire, and men's questionnaire. HIV testing was also conducted in the most recent wave in the eleven countries in my sample. In Burkina Faso, Cameroon, Ethiopia, Ghana, Kenya, Lesotho, Malawi, Senegal, and Tanzania, these HIV test results can be linked to other respondent characteristics. In Mali and Zambia, HIV test results are unlinked to the survey data; however, test results can be used to calculate HIV prevalence for various subgroups, including local prevalence. Local HIV prevalence is estimated at the region and sector (urban or rural) level, and is calculated as HIV prevalence among adults ages 15-49 (for comparability across countries).

All households selected for the survey were asked to respond to a household questionnaire, which provides information about the age, sex, educational attainment, and school enrollment status of household members and visitors (in addition to other

[15] The 2003 DHS for Tanzania is also referred to as the HIV/AIDS Indicator Survey (AIS), and covers only mainland Tanzania. The data used in this analysis from that survey as well as from Ethiopia (2005) and Senegal (2005) are from preliminary releases of the data. The Ethiopia survey lists 1997 as the survey year, because the Ethiopian calendar differs from the Gregorian calendar; however, when calculating year of birth, this analysis refers to the Gregorian calendar.

information). In Burkina Faso, Malawi, Mali, Senegal, and Zambia, HIV tests were conducted in a one-in-three subsample of households. In Cameroon, Ethiopia, Kenya, and Lesotho, HIV tests were conducted in a one-in-two subsample of households. In Ghana and Tanzania, HIV tests were conducted in all households. The analysis uses data on individual characteristics from the household questionnaire.

Each respondent eligible for HIV testing was asked to provide a blood sample for testing. In Burkina Faso, Cameroon, Ethiopia, Ghana, Kenya, Lesotho, Malawi, Mali, Senegal, and Tanzania, HIV testing was conducted on dried blood spot specimens collected by finger prick. In Zambia, the dried blood spot specimen came from a venous blood specimen.

Survey and HIV test non-response rates are shown in Appendix Table A.I. Response rates for the household questionnaire are quite high. Response rates for the HIV test components are somewhat lower.

This analysis uses household weights in most specifications, since information is drawn from the household questionnaire.[16] These weights adjust for the household sampling probability and household response rate (Rutstein and Rojas, 2003). In regressions which pool data from multiple countries, I multiply these household weights by population estimates from the 2007 CIA World Factbook. Local HIV prevalence is calculated using DHS-provided HIV weights, which adjust for individual sampling probabilities and test non-response (separately by sex).

[16] Results in columns (2) and (3) of Table V use the product of HIV weights from the cluster, adjusted by population size and rescaled by the country mean for non-orphan 7- to 14- year-olds in tested households.

TABLE A.I
SURVEY RESPONSE RATES

Country	Household Questionnaire	HIV Test		
		Men	Women	Total
Burkina Faso	99.4	85.8	92.3	89.3
Cameroon	97.6	89.8	92.1	91.0
Ethiopia	98.5	75.6	83.4	79.6
Ghana	98.7	80.0	89.3	84.9
Kenya	96.3	70.3	76.3	73.4
Lesotho	95.2	68.0	80.7	74.7
Malawi	97.8	63.3	70.4	67
Mali	97.9	75.6	85.2	80.7
Senegal	98.5	75.5	84.5	80.4
Tanzania	98.5	77.0	83.5	80.5
Zambia	98.2	73.3	79.4	76.5

Notes. Results are from the DHS for Burkina Faso (2003), Cameroon (2004), Ethiopia (2005), Ghana (2003), Kenya (2003), Lesotho (2004), Malawi (2004), Mali (2001), Senegal (2005), Tanzania (2003), and Zambia (2001/2002). Percent surveyed or tested among eligible respondents.

TABLE I.A

LOCAL HIV PREVALENCE IN SURVEY YEAR, DETAILED SUMMARY STATISTICS

HIV RATE	BF	CM	ET	GH	KE	LS	MW	ML	SN	TZ	ZM	Total
Mean	1.70	6.37	3.29	2.12	7.62	24.77	12.90	1.48	0.79	8.07	17.35	5.66
Std. Dev.	1.42	2.86	2.84	1.14	5.75	5.66	6.86	1.38	0.70	6.23	7.57	5.89
25th Pctile	0.69	4.38	0.89	1.29	4.07	19.52	5.95	0.40	0.35	4.08	10.19	1.22
Median	1.33	6.47	2.97	2.11	6.38	23.43	13.18	1.35	0.60	6.41	17.18	4.23
75th Pctile	2.47	8.26	5.75	2.60	11.35	29.00	18.51	2.11	1.16	10.71	22.37	7.38
Obs.	25	22	22	20	15	20	6	16	22	42	18	228

Notes. Results are from the DHS for Burkina Faso (2003, BF), Cameroon (2004, CM), Ethiopia (2005, ET), Ghana (2003, GH), Kenya (2003, KE), Lesotho (2004, LS), Malawi (2004, MW), Mali (2001, ML), Senegal (2005, SN), Tanzania (2003, TZ), and Zambia (2001/2002, ZM). Table shows detailed summary statistics for the local HIV rate in the survey year, which is estimated for each region and sector of residence (within each country). The unit of observation is a region/sector. These HIV rates are calculated from the DHS HIV data using a sample which includes men and women ages 15-49, weighted using appropriate HIV sample weights. In calculating summary statistics overall, country observations are weighted by population.

TABLE I.B
SAMPLE CHILD CHARACTERISTICS, SUMMARY STATISTICS

	N	Mean	Standard Deviation
Local HIV Prevalence in Survey Year	128,035	0.037	0.050
Progress	122,744	0.299	0.346
Year of Birth	128,035	1994	2.607
Female	128,021	0.491	0.500
Rural	128,035	0.811	0.391
Burkina Faso	128,035	0.064	0.244
Cameroon	128,035	0.065	0.247
Ethiopia	128,035	0.420	0.494
Ghana	128,035	0.043	0.202
Kenya	128,035	0.099	0.299
Lesotho	128,035	0.005	0.074
Malawi	128,035	0.062	0.241
Mali	128,035	0.060	0.237
Senegal	128,035	0.056	0.229
Tanzania	128,035	0.094	0.292
Zambia	128,035	0.032	0.177

Notes. Results are from the DHS for Burkina Faso (2003), Cameroon (2004), Ethiopia (2005), Ghana (2003), Kenya (2003), Lesotho (2004), Malawi (2004), Mali (2001), Senegal (2005), Tanzania (2003), and Zambia (2001/2002). All results are weighted using provided household sample weights adjusted by population size. Sample includes boys and girls ages 7-14 at the time of the survey. Local HIV Prevalence is estimated (based on DHS HIV data) HIV prevalence among men and women ages 15-49 in the region and sector in 2001 (Mali), 2001-2002 (Zambia), 2003 (Burkina Faso, Ghana, Kenya, and Tanzania), 2004 (Cameroon, Lesotho, and Malawi) or 2005 (Ethiopia and Senegal). Progress is the number of completed years of schooling divided by the potential (age – 6); values outside the [0,1] range are coded as missing. Year of Birth is the respondent's year of birth. Female is an indicator for whether the respondent is female.

TABLE I.C
SAMPLE ADULT CHARACTERISTICS, SUMMARY STATISTICS

	N	Mean	Standard Deviation
Local HIV Prevalence in Survey Year	237,869	0.042	0.052
Years of Schooling	236,611	4.163	4.333
Years of Schooling > 0	236,611	0.594	0.491
Completed Primary School	236,827	0.360	0.480
Year of Birth	237,869	1976	9.611
Female	237,867	0.519	0.500
Rural	237,869	0.724	0.447
Burkina Faso	237,869	0.062	0.241
Cameroon	237,869	0.073	0.261
Ethiopia	237,869	0.392	0.488
Ghana	237,869	0.044	0.206
Kenya	237,869	0.112	0.315
Lesotho	237,869	0.007	0.084
Malawi	237,869	0.059	0.236
Mali	237,869	0.055	0.227
Senegal	237,869	0.063	0.244
Tanzania	237,869	0.098	0.298
Zambia	237,869	0.034	0.180

Notes. Results are from the DHS for Burkina Faso (2003), Cameroon (2004), Ethiopia (2005), Ghana (2003), Kenya (2003), Lesotho (2004), Malawi (2004), Mali (2001), Senegal (2005), Tanzania (2003), and Zambia (2001/2002). All results are weighted using provided household sample weights adjusted by population size. Sample includes men and women ages 15-49 at the time of the survey. Local HIV Prevalence is estimated (based on DHS HIV data) HIV prevalence among men and women ages 15-49 in the region and sector in 2001 (Mali), 2001-2002 (Zambia), 2003 (Burkina Faso, Ghana, Kenya, and Tanzania), 2004 (Cameroon, Lesotho, and Malawi) or 2005 (Ethiopia and Senegal) multiplied by an indicator for whether the respondent is in an affected (post-1980) birth cohort. Years of Schooling is completed years of schooling. Years of Schooling > 0 is an indicator for whether the respondent completed at least one year of schooling. Completed Primary School is an indicator for whether the respondent completed primary school. Year of Birth is the respondent's year of birth. Female is an indicator for whether the respondent is female.

TABLE II

DIFFERENCE-IN-DIFFERENCES REGRESSION: SEPARATELY BY COUNTRY

YEARS	(1) BF	(2) CM	(3) ET	(4) GH	(5) KE	(6) LS
Local HIV Prevalence × Post-1980	-5.149	-9.022	-16.454*	-10.543	-6.649	-5.394*
	(11.336)	(4.849)	(7.550)	(17.861)	(4.318)	(1.802)
Female	-0.923*	-1.335*	-1.543*	-2.068*	-0.824*	1.734*
	(0.128)	(0.140)	(0.224)	(0.093)	(0.147)	(0.243)
Constant	1.608*	6.375*	2.590*	7.905*	6.466*	3.917*
	(0.101)	(0.353)	(0.210)	(0.678)	(0.309)	(0.345)
Controls	Region × Sector FEs, Birth Year FEs					
Observations	24,834	23,123	29,967	11,363	17,398	19,477

YEARS	(7) MW	(8) ML	(9) SN	(10) TZ	(11) ZM
Local HIV Prevalence × Post-1980	0.239	-24.873	26.644*	-1.550	-3.237*
	(3.052)	(20.758)	(10.981)	(2.129)	(0.833)
Female	-1.538*	-1.054*	-1.012*	-0.852*	-1.324*
	(0.160)	(0.272)	(0.090)	(0.049)	(0.084)
Constant	5.455*	1.688*	2.717*	3.408*	6.599*
	(0.337)	(0.314)	(0.171)	(0.574)	(0.341)
	Region × Sector FEs, Birth Year FEs				
Observations	24,795	25,912	29,115	14,167	16,459

Notes. Results are from the DHS for Burkina Faso (2003), Cameroon (2004), Ethiopia (2005), Ghana (2003), Kenya (2003), Lesotho (2004), Malawi (2004), Mali (2001), Senegal (2005), Tanzania (2003), and Zambia (2001/2002). All regressions are weighted least squares regressions with region by sector fixed effects and year of birth fixed effects, weighted using provided household sample weights with clustering on the region/sector. Sample includes men and women ages 15-49. The dependent variable is completed years of schooling. The coefficient of interest is on the interaction between local HIV prevalence and a post-1980 cohort indicator. HIV is estimated (based on DHS HIV data) HIV prevalence among men and women ages 15-49 in the region and sector in 2001 (Mali), 2001-2002 (Zambia), 2003 (Burkina Faso, Ghana, Kenya, and Tanzania), 2004 (Cameroon, Lesotho, and Malawi) or 2005 (Ethiopia and Senegal) and is multiplied by an indicator for whether the respondent is in an affected (post-1980) birth cohort. Female is an indicator for whether the respondent is female. Huber-White standard errors are in parentheses. * = p-value < .05

TABLE III
DIFFERENCE-IN-DIFFERENCES REGRESSION: POOLED

COMPLETED SCHOOLING	(1) YEARS	(2) YEARS > 0	(3) PRIMARY
Local HIV Prevalence × Post-1980 Cohort	-5.115*	-0.526*	-0.573*
	(1.946)	(0.211)	(0.159)
Female	-1.275*	-0.155*	-0.111*
	(0.102)	(0.026)	(0.009)
Constant	4.030*	0.518*	0.339*
	(0.276)	(0.032)	(0.027)
Additional Controls	Region × Sector FEs, Birth Year FEs	Region × Sector FEs, Birth Year FEs	Region × Sector FEs, Birth Year FEs
Observations	236,610	236,610	236,825

Notes. Results are from the DHS for Burkina Faso (2003), Cameroon (2004), Ethiopia (2005), Ghana (2003), Kenya (2003), Lesotho (2004), Malawi (2004), Mali (2001), Senegal (2005), Tanzania (2003), and Zambia (2001/2002). All regressions are weighted least squares regressions with region by sector fixed effects and year of birth fixed effects, weighted using provided household sample weights (adjusted by population size) with clustering on the region/sector. Sample includes men and women ages 15-49. The dependent variable is completed years of schooling (1), an indicator for whether the respondent completed at least one year of schooling (2), or an indicator for whether the respondent completed primary school (3). Local HIV Prevalence is estimated (based on DHS HIV data) HIV prevalence among men and women ages 15-49 in the region and sector in 2001 (Mali), 2001-2002 (Zambia), 2003 (Burkina Faso, Ghana, Kenya, and Tanzania), 2004 (Cameroon, Lesotho, and Malawi) or 2005 (Ethiopia and Senegal) and is multiplied by an indicator for whether the respondent is in an affected (post-1980) birth cohort. Female is an indicator for whether the respondent is female. Huber-White standard errors are in parentheses. * = p-value < .05

TABLE IV
CHANNELS: ORPHANHOOD

PROGRESS	(1) FULL SAMPLE	(2) NO DOUBLE ORPHANS	(3) NO SINGLE ORPHANS
Local HIV Prevalence × Post-1992 Cohort	-0.337*	-0.332*	-0.350*
	(0.124)	(0.125)	(0.135)
Female	-0.005	-0.005	-0.004
	(0.008)	(0.008)	(0.008)
Constant	0.376*	0.407*	0.403*
	(0.023)	(0.015)	(0.015)
Additional Controls	Region × Sector FEs, Birth Year FEs	Region × Sector FEs, Birth Year FEs	Region × Sector FEs, Birth Year FEs
Observations	122,730	118,898	102,416

Notes. Results are from the DHS for Burkina Faso (2003), Cameroon (2004), Ethiopia (2005), Ghana (2003), Kenya (2003), Lesotho (2004), Malawi (2004), Mali (2001), Senegal (2005), Tanzania (2003), and Zambia (2001/2002). All regressions are weighted least squares regressions with region by sector fixed effects and year of birth fixed effects, weighted using provided household sample weights (adjusted by population size) with clustering on the region/sector. Sample includes boys and girls ages 7-14. The dependent variable is the number of completed years of schooling divided by the potential (age – 6); values outside the [0,1] range are coded as missing. Local HIV Prevalence is estimated (based on DHS HIV data) HIV prevalence among men and women ages 15-49 in the region and sector in 2001 (Mali), 2001-2002 (Zambia), 2003 (Burkina Faso, Ghana, Kenya, and Tanzania), 2004 (Cameroon, Lesotho, and Malawi) or 2005 (Ethiopia and Senegal) and is multiplied by an indicator for whether the respondent is in an affected (post-1992) birth cohort. Female is an indicator for whether the respondent is female. "Double Orphans" are respondents whose fathers and mothers are not known to be alive. "Single Orphans" are respondents whose fathers or mothers (or both) are not known to be alive. Huber-White standard errors are in parentheses. * = p-value < .05

TABLE V
CHANNELS: CARETAKING REQUIREMENTS

PROGRESS	(1) NO SINGLE ORPHANS	(2) TESTING SAMPLE	(3) NO HIV+ MEMBER
Local HIV Prevalence × Post-1992 Cohort	-0.350*	-0.362	-0.407*
	(0.135)	(0.185)	(0.199)
Female	-0.004	0.000	-0.000
	(0.008)	(0.009)	(0.009)
Constant	0.403*	0.433*	0.195*
	(0.015)	(0.019)	(0.014)
Additional Controls	Region × Sector FEs, Birth Year FEs	Region × Sector FEs, Birth Year FEs	Region × Sector FEs, Birth Year FEs
Observations	102,416	35,884	32,998

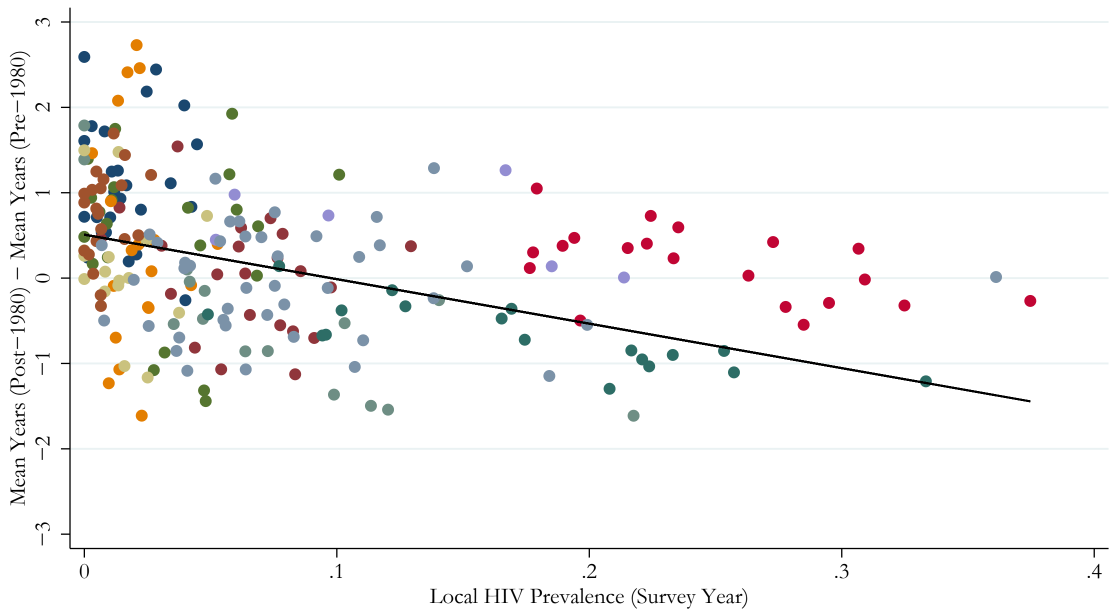
Notes. Results are from the DHS for Burkina Faso (2003), Cameroon (2004), Ethiopia (2005), Ghana (2003), Kenya (2003), Lesotho (2004), Malawi (2004), Mali (2001), Senegal (2005), Tanzania (2003), and Zambia (2001/2002). All regressions are weighted least squares regressions with region by sector fixed effects and year of birth fixed effect with clustering on the region/sector. Weights in column (1) are household sample weights (adjusted by population size). Weights in columns (2) and (3) are the product of HIV weights from the cluster, rescaled by the country mean and adjusted by population size. Sample includes boys and girls ages 7-14. The dependent variable is the number of completed years of schooling divided by the potential (age – 6); values outside the [0,1] range are coded as missing. Local HIV Prevalence is estimated (based on DHS HIV data) HIV prevalence among men and women ages 15-49 in the region and sector in 2001 (Mali), 2001-2002 (Zambia), 2003 (Burkina Faso, Ghana, Kenya, and Tanzania), 2004 (Cameroon, Lesotho, and Malawi) or 2005 (Ethiopia and Senegal) and is multiplied by an indicator for whether the respondent is in an affected (post-1992) birth cohort. Female is an indicator for whether the respondent is female. "Single Orphans" are respondents whose fathers or mothers (or both) are not known to be alive. "Testing Sample" includes households in which at least one member was tested for HIV, and excludes Mali and Zambia, in which HIV test results cannot be matched to the household. Column (3) excludes from the testing sample children living in households with at least one HIV-positive member. Huber-White standard errors are in parentheses. * = p-value < .05

TABLE VI
CHANNELS: DIFFERENCES BY SEX

COMPLETED SCHOOLING	(1) YEARS	(2) YEARS > 0	(3) PRIMARY
Local HIV Prevalence × Post-1980 Cohort × Male	-8.644*	-1.136*	-0.897*
	(1.329)	(0.256)	(0.167)
Local HIV Prevalence × Post-1980 Cohort	-0.998	0.015	-0.146
	(1.693)	(0.186)	(0.118)
Female	-1.432*	-0.176*	-0.128*
	(0.108)	(0.028)	(0.010)
Constant	4.107*	0.528*	0.346*
	(0.278)	(0.032)	(0.027)
Additional Controls	Region × Sector FEs, Birth Year FEs	Region × Sector FEs, Birth Year FEs	Region × Sector FEs, Birth Year FEs
Observations	236,610	236,610	236,825

Notes. Results are from the DHS for Burkina Faso (2003), Cameroon (2004), Ethiopia (2005), Ghana (2003), Kenya (2003), Lesotho (2004), Malawi (2004), Mali (2001), Senegal (2005), Tanzania (2003), and Zambia (2001/2002). All regressions are weighted least squares regressions with region by sector fixed effects and year of birth fixed effects, weighted using provided household sample weights (adjusted by population size) with clustering on the region/sector. Sample includes men and women ages 15-49. The dependent variable is completed years of schooling (1), an indicator for whether the respondent completed at least one year of schooling (2), or an indicator for whether the respondent completed primary school (3). Local HIV Prevalence is estimated (based on DHS HIV data) HIV prevalence among men and women ages 15-49 in the region and sector in 2001 (Mali), 2001-2002 (Zambia), 2003 (Burkina Faso, Ghana, Kenya, and Tanzania), 2004 (Cameroon, Lesotho, and Malawi) or 2005 (Ethiopia and Senegal) and is multiplied by an indicator for whether the respondent is in an affected (post-1980) birth cohort. Male is an indicator for whether the respondent is male. Female is an indicator for whether the respondent is female. Huber-White standard errors are in parentheses. * = p-value < .05

FIGURE 1
SCATTERPLOT: DIFFERENCE-IN-DIFFERENCES



Notes. Results are from the DHS for Burkina Faso (2003), Cameroon (2004), Ethiopia (2005), Ghana (2003), Kenya (2003), Lesotho (2004), Malawi (2004), Mali (2001), Senegal (2005), Tanzania (2003), and Zambia (2001/2002). Each observation is a region/sector, and shows the relationship between local HIV prevalence and the change in completed years of schooling between pre- and post-1980 birth cohorts. Sample includes men and women ages 15-49. HIV is estimated (based on DHS HIV data) HIV prevalence among men and women ages 15-49 in the region and sector in 2001 (Mali), 2001-2002 (Zambia), 2003 (Burkina Faso, Ghana, Kenya, and Tanzania), 2004 (Cameroon, Lesotho, and Malawi) or 2005 (Ethiopia and Senegal). Regression line is from a weighted least squares regression of these region/sector observations, weighted by the country's overall population.