From Predictive to Protective?

The Changing Relationship of HIV and Education in Africa

April 2012

Abstract:

This study uses longitudinal HIV prevalence data from Mali, Tanzania, Kenya and Zambia to examine whether the positive association between educational attainment and HIV prevalence is changing. My findings support the hypothesis that the relationship between HIV and education has begun to reverse in Africa. Although I find that, at the regional level, the association is still positive in all age groups, it is much weaker for the youngest generation. Furthermore, when I examine the relationship at the individual level, using region and country fixed effects, I find no positive association between HIV risk and education in the youngest cohort of respondents. Secondarily, I test two commonly hypothesized explanations for such a change- the erosion of educational infrastructure in high prevalence areas and the adoption of protective knowledge among the educated. I find evidence consistent with the hypothesis that education is becoming protective as the epidemic matures; regions with higher levels of adult education at baseline experience larger drops in HIV prevalence over the time period. Further, this does not appear to be a result of either selective mortality by education level or migration. I do not find evidence that educational attainment is eroding over time in higher prevalence areas.

Introduction

Social scientists have generally viewed educational achievement to be protective against a wide array of disease risks. This negative relationship between education and disease is attributed both to the direct effects of education (those with more education are better able access health information) and to the strong association between education and wealth (wealthier people can afford health care, better nutrition etc.)[1, 2] One very notable exception to this generalized pattern has been HIV in sub-Saharan Africa. Data collected from Africa in the early decades of the epidemic demonstrated that there was a robust *positive* association between education and HIV infection (Gregson et al. 2001; Hargreaves & Glynn 2002; Fortson 2008). This association likely arose, in part, because many of the correlates of education are also risk factors for HIV. For example, areas with high numbers of educated people also tend to be urban, have more developed economies and transport networks, and larger migrant labor populations- all of which are associated with the spread of HIV. In addition people who are more educated tend to marry later and may have more sexual partners before marriage, which can potentially increase exposure to HIV risk (Singh & Samara 1996; Gregson et al. 2001; Bongaarts 2007; Fortson 2009).

However, a variety of scholars in the field have hypothesized that the present positive association may ultimately reverse as populations respond to the development of the epidemic (Gregson et al. 2001; Hargreaves & Glynn 2002; Fortson 2008). More specifically, as the

[1] More educated people tend to enjoy higher income, which can confer a host of health benefits such as access to better care, less physically demanding work, better nutrition and lower levels of psychological stress (Cutler & Lleras-Muney 2006; Ross & Wu 1995).

[2] The health impact of education independent of wealth is a subject of much scholarship. I do not attempt to disentangle them here, but focus on educational attainment measures. Readers may choose to interpret education as a measure of broader privilege without, I hope, changing the usefulness of the findings. I note, however, that Fortson (2008) uses DHS to test the associations between wealth and HIV risk versus education and HIV risk independently, and finds that education has a significant independent relationship with HIV, while wealth does not.

information about how to prevent infection becomes more widely available, the more educated may access the information sooner, and consequently change behavior earlier or at a higher rate. This would ultimately lead HIV risk and education to develop the more common inverse relationship. If such a reversal is, in fact, taking place, it could signal a shift of the burden of the epidemic away from more educated and affluent communities and towards poorer and more marginalized populations.

The aims of this analysis are twofold, first to use longitudinal data to see whether there is evidence that the relationship between education and HIV risk is changing; and second, to see if the descriptive evidence can shed light on what the causes of such a change might be. As such, the analysis begins by asking -*Is there evidence for a reversal of the association between education and HIV risk in sub-Sahran Africa?* A reversal of the education/HIV risk relationship would have important practical consequence in that it implies that new populations are at potential risk. If less educated and more marginalized populations are now assuming comparatively higher risk of HIV infection, then prevention and outreach strategies in sub-Saharan Africa may also have to adjust accordingly. For example, prevention campaigns that rely on print messaging or classroom-based HIV education programs may be considerably less effective.

Moreover, for policy makers interested in influencing the course of the epidemic, the mechanisms driving such a change in the education/HIV association are also a subject of deep interest. If more educated populations are reducing their risk behaviors, it suggests that the educational outreach campaigns of the last decades have had some success. If, on the other hand, the risk relationship is shifting because of other changes- such as high mortality, selective migration or reduced education levels in high prevalence areas- then the same change does not carry nearly as promising of a message. The second part of this analysis concerns itself with this question and tests evidence for several pathways through which the association might be changing.

My findings support the hypotheses of previous scholars that the relationship between HIV and education has begun to reverse in sub-Saharan Africa. Although I find that, at the regional level, the relationship is still positive across age groups, it is much weaker for the youngest generation. I find even stronger evidence at the individual level. Using country by region fixed effects, I find that for the youngest cohort, there is no positive association between an individual's HIV risk and educational attainment.

I also find evidence consistent with the hypothesis that education may be becoming protective in the present state of the epidemic; regions with higher levels of adult education at baseline experience larger drops in HIV prevalence. Further, this does not appear to be a result of either selective mortality or migration. I do not, however, find evidence that children's educational attainment is eroding over time in higher prevalence areas.

Background

Logic suggests at least two mechanisms that could reverse the education/HIV relationship at a population level; either communities with high HIV rates suffer losses in educational advantage, or communities with high education levels comparative HIV rates come down. [3] These two pathways for change are not mutually exclusive or exhaustive, but are the two that are highlighted in the literature and are, consequently, the focus here as well. The first

[3] Note that the changes described need not be driven by changes in comparatively high education/HIV communities. Since they are comparative measures, they could equally driven by increases in education or increases in HIV in previously low education/HIV communities.

pathway, which I label the *'educational erosion pathway'* explains the change via erosion of the educational system in high HIV prevalence areas. The second pathway, which I call the *'protective knowledge pathway'* theorizes that more educated people may change their behaviors earlier as knowledge about HIV prevention disseminates.

Educational Erosion

For some time, observers and scholars have warned that the particular characteristics of the HIV epidemic– that it targets working age adults, that it clusters in households, that the rates of infection are higher among the educated – mean that the effects of the disease could go far beyond the afflicted households and result in widespread societal deterioration (Gachuhi 1999; Hacker 2002; Bennel, Hyde and Swainson 2002; Akbulut-Yuksel and Turan, 2010). This theory anticipates that schools, hospitals and civil institutions will deteriorate as a result of the deaths of a large swath of the educated laboring population. In addition, parents in affected communities, who perceive a high risk of early mortality, may reduce investment in their children's education because they do not expect to live long enough to reap the benefits of the investments. Over time, this process could plausibly work to reverse the positive relationship between HIV and education. If enough erosion of educational infrastructure and investment were to occur, high HIV prevalence regions could eventually become associated with lower levels of education rather than higher.

Recent research lends some support to these bleak projections. Fortson (2011) uses Demographic and Health Surveys in sub-Saharan Africa to show that areas with higher HIV prevalence have larger schooling declines among those who were of school going age when the epidemic hit. Behrman et al (2010) show, in a high prevalence area of Malawi, that mothers who perceive a higher risk of HIV reduce investment in children's education. In addition, Akbulut-Yuksel and Turan (2010) use the DHS surveys to document that children in high HIV prevalence areas accumulate comparatively less education than their parents, even within households with no HIV infected members. Collectively, this research suggests that HIV is associated with losses in education that go beyond the individuals, or even the generation, directly infected.

Protective Knowledge

The second path through which the relationship between education and HIV might reverse is simply through the diffusion of information about the disease, resulting in people changing their behaviors. Historically, as disease mechanisms have become better understood, the more highly educated have adopted protective behaviors earlier and an education gradient in mortality and morbidity has subsequently emerged. For example, in the 1960s, there was little difference in mortality for older American men across education levels, but once the risks of smoking became widely understood (beginning with the release of the Surgeon General's report on smoking in 1964), better educated men's mortality dropped throughout the 70s and 80s, becoming significantly below that of men with less education (Cutler & Lleras-Muney 2006; Crimmins & Saito 2001; Preston & Elo 1995; Feldman et al. 1989). Some researchers have suggested that a similar process may take place with HIV if people with more education respond differently as information about the disease disseminates (De Walque 2007; Gregson et al. 2001).

In the case of HIV in Africa, it is plausible that this process is only recently taking place. Transmission pathways were imperfectly understood until the late 1980s and early 1990s. Even once transmission pathways were identified in the scientific community, disinformation and political resistance to the new knowledge slowed its impact in sub-Saharan Africa. Influential figures, such as the President and Health Minister of South Africa at the time, publicly questioned the scientific consensus, effectively clouding the public narrative about how HIV was spread and how it might be prevented. By the early 2000s, however, the public skepticism about the routes to HIV infection had been largely overcome, and there now seems to be widespread acceptance in both the scientific and popular press on the causes of HIV. Now that information on how HIV/AIDS spreads has become established knowledge in sub-Saharan Africa, I might expect to see the more highly educated adopting protective behaviors at higher rate, leading the relationship between education and HIV eventually to develop the more familiar negative association. Moreover, I would expect such differences in response by education level to be especially visible among the youngest generation, who will have been the first generation to reach their sexually active years in a period when knowledge about effective prevention was widely accepted and publicized. While the uninfected in the older age cohorts are also able to put new knowledge to use, for many of the older generation, access to protective information may have come too late, irrespective of their education level.

Recent work by Case and Paxson (2011) lends support to this theory. They examine behavioral changes among younger women and find that young women in high prevalence areas respond by decreasing teen sexual activity outside of marriage and by entering marriage earlier. More importantly for this analysis, they find that the behavioral effect on early marriage is magnified for highly educated girls. These findings suggest both that younger cohorts of Africans may be changing their behaviors in response to the epidemic, and that there may be some differential in response by education level. If the more highly educated adopt protective behaviors earlier and/or at higher rates, there may ultimately be a reversal in the positive correlation between HIV and education.

Empirical Strategy:

The empirical strategy of this analysis is to examine the descriptive evidence available in four African demographic and health surveys (DHS) over two waves to see if the relationship between HIV risk and education levels change over time. Because risk of HIV infection is hypothesized to relate both to individual behaviors and local risk environment, I examine this question at both the regional and individual level. My initial test is at the regional level- where I compare the trajectory of HIV prevalence in geographic areas with higher and lower average education. I split the individuals in the sample into three different age cohorts and look at how the regional HIV-education relationship changes over time in successively younger cohorts. In addition, where the data allows, I examine the same question at the individual level, comparing the risk of HIV infection among individuals within the same region to see if it changes differentially by education level.

In order to address the second research question on potential causes, I use multiple waves of HIV and education data to calculate the change in regional HIV prevalence over time and examine whether it relates to the initial level of education in the region or vice versa. For example, I ask whether areas with higher levels of education have larger declines in HIV prevalence or, alternately, do areas with very high HIV prevalence see larger declines in children's educational attainment?

Data and Measures

The data for this analysis is drawn from the Demographic and Health Surveys (DHS). The DHS are large nationally representative household surveys that cover household members' education, nutrition, fertility, and health. More recently, some African DHS have asked a subsample of adult respondents to provide blood samples for HIV testing[4]. The four countries examined here - Kenya, Mali, Tanzania and Zambia- are, to date, the only ones for which two rounds of HIV testing data are available. All the data used here was collected between 2001 and 2009. For Mali and Tanzania, the two waves of HIV and household data were collected approximately five years apart. For Kenya and Zambia, the two waves were collected six years apart.[5] I also use two measures of regional educational attainment. The first measure is mean years of completed education among adults in the region. However, because some of the respondents are still quite young and could still be completing education, I also use percent of adults aged 15-49 who have completed primary schooling as a secondary measure of regional educational levels. In the section testing the educational erosion hypothesis, I examine changes in children's schooling over time at the regional level by calculating regional attendance rates and mean years of completed education among children 7-14 years old.

Much of the analysis here rests upon regional measures of HIV prevalence and educational attainment calculated at two points in time. Regions are identified within each country by using the geographic regions as specified in the DHS and further dividing them by urban and rural sectors where applicable, resulting in 91 distinct regions. The number of

[4]Response rates were high and analysis on a subset of countries has shown that bias from non-response is small (Mishra et al. 2006).

[5] In Mali, HIV tests were collected in 2001/02 and 2006. In Tanzania they were collected during 2003/4 and again in 2008. In Kenya, HIV prevalence data was collected in 2003 and again in 2009. For Zambia, it was collected in 2001 and 2007.

observations available for analysis depends on the level of analysis and outcome being measured. Regional educational attainment measures are calculated using the full sample of 149,116 adults aged 15-49. I calculate regional HIV prevalence using information on the 39,913 adults aged 15-49 in the HIV-tested subsample.[6] In my analysis of the individual-level data, I am limited to the data for Tanzania and Kenya (the first wave HIV survey results in Mali and Zambia were not linkable to individual survey responses), leaving me with an analytic population of 24,943 individuals in the youngest and middle cohorts. In the section testing the educational erosion hypothesis, I examine changes in children's schooling over time at the regional level based on 82,304 children who are aged 7-14 in the sample households.

Table 1 presents summary statistics of regional HIV prevalence and regional educational outcomes by country and wave. Kenya, Tanzania and Zambia are fairly typical of the East and Central African region in that they all have relatively high levels of adult education and correspondingly high levels of HIV. Mali is the lone West African country and has a different pattern with lower levels of adult education and lower levels of HIV. Over the course of the two waves, there is a general trend of a slight increase in mean years of education completed and relatively little change in percentage of adults with primary education. There are also modest drops in mean regional HIV prevalence across all four sample countries over time.

[Table 1 approximately here]

As noted above, much of the analysis relies on regional aggregations of individual level data, and Table 2 contains basic descriptive information on the individuals from which the regional means are constructed. Within both the adult and child samples, the age and gender

[6] DHS surveys provide HIV sample weights to adjust the tested subsample to be representative of the sample as a whole, and I use them here to compute regional HIV prevalence.

distributions are comparable across both countries and waves. Similar to the regional patterns, the individual level data also shows an upward trend overall in educational attainment and Mali has lower measures of education for both adults and children than the other countries.

[Table 2 approximately here]

Analysis Part 1: Is the Relationship Between HIV and Education Changing?

I approach the question of whether the relationship between education and HIV is changing by splitting the sample into three different age cohorts and looking at the relationship in successively younger cohorts. I divide the adults into a 'youngest' cohort who are aged 15 to 24 in the first wave, a 'middle' cohort who are aged 25 to 34 in the first wave, and an 'oldest' cohort who are aged 35 to 44 in the first wave. I am then able to follow the cohorts as they age in the second wave, which is either five or six years later, depending upon the country. The youngest cohort is of particular interest because its members would have been born between 1976 and 1987 and entered adolescence during the 1990s and early 2000s. In effect, this group is the first that could have begun making decisions about sexual behaviors in a period when the transmission pathways of HIV were widely accepted and the means of prevention were known and publicized.

To begin examining the question of whether the relationship between HIV prevalence and education levels differs by cohort, I estimate a simple model of the relationship between regional HIV prevalence and education for each cohort as below:

$$(1) \qquad HIV_{rcw} = \beta_0 + \beta_1 Ed_{rcw} + \theta X_{rcw} + \varepsilon_{rcw}$$

where the dependent variable HIV_{rcw} is the regional HIV prevalence for adults in region *r*, within age cohort *c* at wave *w*. Independent variable Ed_{rcw} is the regional mean for the educational outcome (either 'years of education' or 'primary school completion rate') for adults in age cohort *c* and region *r* at wave *w*. X is a vector of control variables for the same region, cohort and wave, including year of the survey measurement and indicators identifying each country.

[Table 3 approximately here]

Table 3 shows the coefficients of ordinary least squares regressions of regional HIV prevalence on the mean regional education level for each cohort in the first wave (model (1)). Figures 1a and 1b show the same regression lines in graphical form. For both measures of education, the relationship is positive and statistically significant across age cohorts. However, the coefficient for the youngest cohort is less than half those for the older two cohorts. For adults between 25 and 34, one additional year in the mean education for adults in the region is associated with a correspondingly higher level of HIV prevalence of about 2.6 percentage points, whereas for 15-24 yr olds, the same difference in mean years of education is associated with only about 1.2 percentage point greater HIV prevalence. Further, this difference between cohorts is statistically significant (pval<0.01).

[Figures 1a and 1b approximately here]

A similar pattern exists when primary school completion is used to measure regional education levels. Whereas 25-34 year olds see an approximately 0.27 percentage point higher HIV prevalence associated with a one percentage point increase in the share of adults who completed primary school, the younger cohort has a much weaker association, with a 0.1 percentage point greater HIV prevalence associated with the each one percentage point increase in the share of adults who completed primary school.

This weaker relationship between HIV and education for the younger age cohort is consistent with the idea that the strong association between education and HIV is beginning to change. Assuming there is consistency in education levels within regions across cohorts, young people in highly educated areas are potentially still exposed to more risk, in part due to the residual higher prevalence in their communities. Nonetheless, the attenuation of the education-HIV association suggests that they may be also altering behaviors more to protect themselves.

However, protective behavioral changes are not the only possible explanation for the observed patterns. The difference between cohorts depicted in Figures 1a and 1b could also be consistent with an association that simply strengthens as people age. As time goes on, and the young cohort experiences a longer period of sexual activity, those in highly educated regions may contract the diseases at higher rates and ultimately the relationship could evolve to look exactly as it does for the older cohorts. Because I have two rounds of HIV testing data, I am able to examine this possibility by looking at the progression of the association for all the cohorts over a 5-6 year period.

[Figures 2a and 2b approximately here]

Figures 2a and 2b overlay the regression lines for the same three age cohorts when education and HIV prevalence is measured again in second wave. The solid lines are the same regression lines from Figures 1a and 1b. The dotted line represents model (1) repeated in wave 2 for the youngest cohort, now approximately aged 20-31. The dashed lines similarly represent the relationship for the older cohorts as they are measured in the second wave.[7] The corresponding coefficients are also presented in Table 3. If the relationship between HIV and education strengthens as the cohort ages, I would expect to see the slope of the dotted line begin to rotate up –approaching the slopes observed for the older cohorts -and the OLS coefficient on the education measure to increase from wave to wave. In fact, it does not. HIV prevalence goes up somewhat among the younger group at all education levels, but the slope of the line does not change in any substantive way. The coefficient estimate on 'years of education' which is 0.012 in the first wave decreases to 0.011 in the second wave and the difference between the two is non-significant. The pattern for the middle cohorts is similar with a small statistically non-significant change in the education coefficients.

The oldest cohort actually demonstrates the most change in the coefficient between waves. The coefficient on years of education drops from 0.032 to 0.020. This pattern could indicate that less educated people in the older group are contracting HIV at a much higher rate or that selective morality by education level is driving the observed change. While I test for selective mortality for all three age cohorts below, I am most concerned about it for this age group- who are more likely to have contracted HIV early in the epidemic and therefore the most

[7] Part of the oldest age group ages out of measurement over the course of the two waves (i.e. they pass 49 yrs of age and are no longer asked the individual questionnaires), As such, the wave 2 measurement is based on only the younger members of this cohort.

likely to have had the disease long enough to be experiencing high levels of AIDS mortality.[8] In addition, a small part of this cohort ages out of the survey between wave one and two. Because of these two factors, I treat the estimates from this cohort with caution, and the remaining analyses are presented both including and excluding this age cohort in order to ensure the results are robust to its exclusion. Further, I focus the remaining sections on the younger two cohorts, for whom I have complete information in both waves. Ultimately, the inclusion or exclusion of the oldest cohort does not change the substantive results of the analysis. The relationship between education level and HIV risk is substantially weaker for the youngest cohort and this cohort difference appears to persist over time.

Testing the Change in the Education/HIV relationship at the Individual Level

As with most communicable diseases, the relationship between individual behaviors and HIV risk is partly due to individual choices but also related to the environment in which the individual resides. Two individuals in different regions who chose to engage in unprotected sex are making the same behavioral decision, but the one who lives in an area with high prevalence is at much higher risk of contracting HIV from a partner than the one who lives in an area with low prevalence. As noted previously, this could partly explain why the relationship between HIV and education weakens for the youngest cohort but does not disappear entirely in the regional level analysis. Because HIV risk has both a regional and an individual behavior component, I move from comparing averages across regions and compare individuals within the same region.

[8] If there is selective mortality of the highly educated, then the education levels of a cohort should decrease over time. Although it is not statistically significant, I do find a slight drop in mean education levels of the oldest cohort between waves. In contrast, the younger two cohorts have slightly positive changes, although they are also not statistically significant. A more detailed treatment of the selective mortality test is included in part 2 of the analysis.

As a second step, I ask, *within the same country and region* do more educated individuals have higher risk than the less educated, and does that change over time?

Using individual questionnaire responses from Kenya and Tanzania, I first test the underlying premise that adults in the sample with more education actually are more knowledgeable about HIV/AIDS prevention. I use the same two education measures (years of schooling the person completed and a binary measure of whether the individual completed primary school) and I find that both education measures are positively and statistically significantly associated with HIV knowledge.[9] More educated people are more likely to know that a healthy looking person can be infected with HIV and that HIV risk can be minimized by regular condom use, abstinence, and/or minimizing concurrent sexual partners. Education is also negatively associated with incorrect beliefs that HIV can be transmitted through mosquito bites or sharing food with an infected person. This association between education and HIV/AIDS knowledge is consistent across cohorts and across both waves.[10]

[Figure 3 approximately here]

I next test whether having more years of education translates to lower prevalence among the educated. As a first step, I compare the probability of testing HIV positive at each year of educational attainment for the two younger cohorts in each wave. Figure 3 shows this relationship for both cohorts in the first and second wave. Several things are evident from this simple cross sectional relationship. The first is that the youngest cohort is at lower risk across all

[9] As only 4% of the sample has more than 12 years of education, I top code years of education at 12 yrs, but results are robust to using 16 years of education as well.

[10] This analysis is not shown here, but is available on request. I run probit regressions of an individual's probability of correctly answering questions about HIV on the education measures. I control for individuals age, age squared, urban residence, country, region, gender and the year of measurement. All results reported are statistically significant at 5% level or higher.

education levels in both waves. The second is that there is no positive relationship between HIV and education among the younger cohort at the individual level. Furthermore, for individuals at the lower end of the educational scale, the prevalence of HIV increases considerably between waves in *both* cohorts. The middle cohort experiences increases in HIV prevalence among both those with high and low levels of education, although the increases are more dramatic at the lower end. The youngest cohort experiences substantial increases in infection rates from wave one to wave two only at the lower end of the educational spectrum. The results depicted in Figure 3 suggest that, for individuals in the youngest cohort, not only has the relationship between education and HIV ceased to be positive, it may actually be becoming negative. I test this relationship more directly by splitting the sample into adults with primary education or more, and those with less than a primary education. I then run a difference of differences model in each cohort by estimating,

$$(4)\quad HIV_{iw} = \beta_0 + \beta_1 Primary_{iw} + \beta_2 W2 + \beta_3 Primary_{iw} * W2 + \theta X_{iw} + \varepsilon_{iw}$$

where HIV_{iw} is the probability that individual *i* in wave *w* tests positive for HIV, and $Primary_{iw}$ is a binary indicator of whether that individual reported having completed primary school or more. $Wave2$ is an indicator for whether the observation is during the second wave of the survey, and X_{iw} is a vector of controls including age, gender, urban residence and country region fixed effects. The coefficient $\beta_3$represents the difference-in-difference estimator. In other words, β_3 is an estimate of how much the change in HIV prevalence between waves for people with primary education differs from the change in HIV prevalence for people with less than a primary education. I find that for the youngest cohort, β_3 is statistically significant and equal to -

0.016 (pval<.05). In other words, the probability of being HIV positive for an individual with less than a primary education increases by 1.6 percentage points more than it does for someone with a primary education or more. [11] For the middle cohort, the coefficient estimate of β_3 is -0.009 and is non-significant.

Collectively, I interpret the differences between cohorts illustrated in this section as early evidence that the relationship between HIV and education is undergoing a change. Further, this change is concentrated among younger people, who both have lower rates of HIV prevalence and a weaker association between regional education rates and regional HIV prevalence. [12] At the individual level as well, the evidence suggests the relationship is different for the youngest generation. Whereas for the older generations, education was a risk factor for HIV infection, having completed primary school or more now appears to be comparatively protective for the young. [13]

Analysis Part 2: Examining Potential Pathways for Change

The observation that the positive relationship between education and HIV is weakening across age cohorts, although interesting in itself, does not account for the causes or pathways for change. If the changes observed above reflect educated individuals adopting protective behaviors as a result of the dissemination of prevention knowledge, then they are a hopeful development in the efforts to curtail the epidemic. Conversely, if they are due to the deterioration of educational

[11] Estimates are based on a sample of 14,323 HIV tested individuals in the youngest cohort, and 10,620 HIV tested individuals in the middle cohort, *p*-values based on bootstrapped standard errors.
[12] In an additional analysis, which is not shown, I repeated the above separately for men and women but did not find substantively different results by gender.
[13] I also performed the same differences-in-differences estimation at the regional level (comparing regional changes in HIV prevalence across waves for primary educated and those with less than a primary education) and results were consistent with the individual level analysis.

infrastructure in previously highly educated areas, then they are hardly good news. Although their implications differ, there is no reason that these two mechanisms are exhaustive or mutually exclusive. Both could be happening simultaneously. The remaining sections of the paper begin to explore evidence for the possible pathways for these changes, looking first at the '*educational erosion*' pathway and secondly at the '*protective knowledge'* pathway.

Evidence for Educational Erosion

In this scenario, the causality runs from HIV to education, where high prevalence causes a widespread social decline, including the deterioration of educational infrastructure and investment.

As noted above, much recent work on HIV in sub-Saharan Africa has documented an association between high HIV prevalence and deterioration in human capital. To my knowledge, this data offers the first opportunity to examine both the changes in HIV and education longitudinally in several nationally representative datasets. If high HIV prevalence is causing a general deterioration in educational infrastructure and investment, I would expect to see education levels fall over time in high prevalence areas.[14] In order test for such an effect, I compare longitudinal changes in children's educational outcomes across regions with varying levels of prevalence. Specifically, I look at children aged 7-14 in both waves, and test whether the education levels of children in high HIV prevalence areas fall (or rise less) than those in lower prevalence areas.

[14] Note that high prevalence areas are generally more educated to begin with, so the overall levels of both adults and children will be higher there.

I concentrate on children's educational outcomes for two reasons; First because the educational attainment of adults is largely completed and unlikely to change much over the observed time period; and second, because of the potential bi-directionality of the HIV/education relationship among adults. School age children, in contrast, are far less likely to have contracted HIV themselves, and so losses in their educational attainment speak more to generalized erosion of schooling in communities than to the direct consequences of infection. Once again, I use two different measures of regional education levels. The first is a measure of the share of children aged 7-14 who are attending school. However, the corrosive effect of HIV might not be seen simply in attendance rates, but also in children's overall completion or progress through school. As such, I also use average years of completed education for the 7-14 age group. For both educational outcomes, I estimate the least squares regression for the model,

$$(2) \qquad \Delta Ed_{rw2-rw1} = \beta_0 + \beta_1 HIV_{rw1} + \theta X_{rw1} + \varepsilon_{rw1}$$

in which $\Delta Ed_{rw2-rw1}$ is the change in the regional mean for the educational outcome (either 'average years of education completed' or 'school attendance rate') for children in region *r* between wave 1 and wave 2. Independent variable HIV_{rw1} is the *adult* regional HIV prevalence for region *r* in wave 1, and X_{rw1} is the same vector of regional control variables from model (1). If high HIV prevalence is leading to a general erosion of educational opportunity, I would expect to see a drop in children's educational attainment between the waves (or perhaps smaller rises) in places with high adult HIV prevalence at baseline. Figures 4a and 4b show the regional changes in children's educational outcomes in relation to the adult HIV prevalence in wave 1. Table 4 shows the regression coefficients which correspond to the regression lines depicted in the figures.

[Table 4 and Figures 4a and 4b approximately here]

As you can see in the figures, the general educational trend is positive. Educational outcomes for children in most areas changed for the better over the period covered by the two waves. I do not find any statistically significant evidence that regions with higher HIV prevalence are experiencing smaller gains (or losses) in education when compared to lower prevalence regions. Although the coefficients on the OLS regressions are negative, they are not significant either in aggregate or when separated by gender.

In summary, contrary to the previous work cited, I don't find evidence for educational erosion in high HIV areas. I interpret these results with caution, however, since the sample is limited to 91 regions. It is possible that, as more longitudinal data becomes available, there will be stronger evidence either for or against the educational erosion pathway.

Evidence for Protective Knowledge

In this scenario, the causal pathway runs from education to HIV rates. It predicts that places with high numbers of educated people will see more pronounced declines in HIV prevalence until high education rates come to be associated with low prevalence, a reversal of the current pattern.

If education has become protective as a result of dissemination of knowledge about HIV, one might expect to see more highly educated people adopt protective behaviors first and, consequently, see HIV prevalence fall or rise more slowly in more highly educated areas. To examine this possibility, I ask whether areas with higher average education levels experience greater decreases in HIV prevalence over time than area with lower education levels. To do this, I observe the change in the adult HIV prevalence between the waves in relation to the mean adult education levels in the region at baseline. I use least squares regression to estimate the equation,

$$(3) \qquad \Delta HIV_{rcw2-rcw1} = \beta_0 + \beta_1 Ed_{rcw1} + \theta X_{rcw1} + \varepsilon_{rcw1}$$

where $\Delta HIV_{rcw2-rcw1}$ is the change in HIV prevalence between wave 1 and wave 2 for adults in region *r*, within age cohort c . Ed_{rcw1} is the regional mean for the educational outcome (either 'years of education' or 'primary school completion rate') for adults in age cohort *c* and region r at wave 1. X is the vector of regional control variables from model (1) including survey year and country indicators. Table 5 reports coefficient β_1 for model 3 for both education measures. Results are presented for the population as a whole and independently for the younger and middle age cohorts. There is a statistically significant association between the education level at baseline and the trajectory of HIV prevalence for both measures of education. One additional year of mean adult education is associated with a -0.7 percentage point difference in the change in regional HIV prevalence between the two waves. In other words, regions with higher adult education experience larger falls (or smaller rises) in HIV prevalence (as shown graphically in Figures 5a and 5b) and this pattern appears consistent across age cohorts and in both high and low prevalence countries.[15] [16]

[15] The coefficient on regional mean adult years of education and mean primary completion for Mali (the lone low prevalence country) are not statistically significant on their own as there are only 16 regions in Mali (and therefore very low statistical power) but the pattern is the same. i.e. the coefficients are negative and only slightly smaller than those for the high prevalence countries.

[Table 5 and Figures 5a and 5b approximately here]

While this pattern is supportive of the theory that education has begun to become protective, it is equally consistent with selective mortality by education level. If, as we know, highly educated people also experienced higher HIV rates early in the epidemic, then it is possible that there is differential mortality by education. That is, those more educated people that first contracted the disease are dying at a higher rate, bringing prevalence down among the educated via mortality rather than via averted cases. Unfortunately, I cannot distinguish HIV mortality from other types of mortality with this data, so it is difficult to examine this possibility directly. However, I can test for selective mortality by education by looking at whether education levels drop among the same age group over time. Since respondents don't normally lose years of education once they have accumulated them, the average education levels of a cohort should only be able to rise or stay the same as time goes on. If the average education level of a cohort drops, then it is very likely that more educated people are dying at a higher rate. As mentioned above, I might expect to find this effect most among the older cohorts since the youngest are less likely to have begun experiencing high mortality in the relatively short time period covered by the two waves.

I find that, for all three cohorts, there is no statistically significant change in average completed education over the two waves. For the youngest cohort, both the mean years of education and primary school completion increase slightly, by .39 years and .06 percentage points respectively. However, the change is not statistically significant. The middle cohort sees virtually no change in education. For the oldest cohort, the mean regional education values

[16] The finding is also consistent for both men and women when model (3) is run separately by gender.

decrease by -.19 years of education and -.037 percentage points for primary completion although the decreases are not statistically significant. I take this as evidence that selective mortality by education is unlikely to be driving the observed changes in the education-HIV relationship. However, I also calculate model (3) for just the younger two cohorts to account for the possibility that my results are driven by differential mortality by education in the oldest cohort which is not captured adequately by the test here. The findings are robust to the exclusion of the oldest group and the correlation between high adult education levels and reduction in regional HIV prevalence is statistically significant and of similar magnitude when only the younger cohorts are analyzed.

Another potential explanation for the patterns observed here is selective migration by HIV status. Prior research has shown that there is often a positive relationship between HIV status and migration experience (Anglewicz 2007) and that HIV positive individuals commonly return home to rural areas when they become sick (Clark et al. 2007). It is possible that the patterns described thus far are driven by people who are infected with HIV leaving places of high average education (such as metropolitan areas) to return home. If this were the case, the falls in prevalence in these areas have less to do with high average adult education levels than with high baseline HIV rates causing significant out-migration.[17] In order to address the possibility that my results are driven by selective migration, I excluded from the sample the 17 % of observations where the individual reported migrating, either within the country or internationally, between waves of the survey. I then re-calculated the regional HIV and education variables without the migrants and re-ran the analysis to see if the relationships described above held in the absence of

[17] I note that much of the change in the HIV education relationship is concentrated in the younger cohort. It seems unlikely that large enough percentages of this young age group would be in the advanced stages of AIDS and therefore migrating due to illness, but since the tests administered by the DHS identify only HIV status and not whether an individual has progressed to AIDS, the observation is necessarily speculative. As such, I test to exclude the possibility that migration is driving the results anyway.

migrants. I find that they do. While the exclusion of those who migrated between waves results in minor changes in the coefficients, all the coefficients reported in Tables 3-5 are in the same direction as reported and remain statistically significant. [Appendix] This suggests that the patterns described here are unlikely to be solely a result of selective migration away from areas with high average adult education. Finally, it is also possible that highly educated and HIV positive people could migrate out of the country entirely, which would not be identifiable in this cross sectional data, but would result in similar patterns as those seen here. In practical terms, that would be similar to selective mortality by education level in that it would mean the result is driven by the exit of a highly educated HIV infected individuals from the data. As such, the prior empirical test for losses in average education, which argued against selective mortality, similarly undermines an explanation based on large-scale out-migration by education level.

VI. Conclusion

The analysis above takes advantage of some of the first available nationally representative and longitudinal evidence on both HIV prevalence and educational trends. My findings support the hypotheses of previous scholars that the relationship between HIV and education has begun to reverse in sub-Saharan Africa. Although I find that, at the regional level, the relationship is still positive across age groups, it is much weaker for the youngest generation. For the youngest cohort in Tanzania and Kenya, I also find that, within the same country and region, individuals with more education do not have a higher risk of HIV infection.

In my examination of potential pathways for this change, I find support at both the regional and individual level for the idea that education is becoming more protective, perhaps as knowledge about the disease disseminates. Regions with the highest levels of adult education saw the largest drops in HIV prevalence over time and younger people with a primary education or higher experienced less increase in HIV prevalence than those in the same age cohort with less education. I do not, however, find evidence of general erosion in education in high prevalence regions. Although the evidence on the pathways is not conclusive, it does support a more hopeful interpretation of the changes in the education-HIV relationship. Rather than being evidence of general societal deterioration as a result of AIDS, the changes may reflect some success for the educational campaigns of the last two decades.

Appendix:

In order to address the possibility that the patterns observed could be caused by HIV positive people moving out of areas of high education (i.e. migration as a confounder), I re-ran the analysis to see if the findings would be robust to exclusion of migrants entirely. I eliminated individuals from the data who reported moving during the period covered by the two waves (about 17 % of the data), recalculated all regional education and HIV prevalence means, and then re-estimated the coefficients in Tables 3-5. If the coefficients changed significantly, then it would be plausible evidence that migration during the period had a substantively meaningful impact on the patterns observed above. Ultimately, the coefficients changed very little in the absence of migrants. The tables are reproduced below.

Table 6: OLS of Regional HIV Prevalence on Regional Education by Cohort and Wave (without migrants)

	Regional HIV Prevalence Youngest Cohort		HIV Prevalence Middle Cohort		HIV Prevalence Oldest Cohort	
	Wave 1 (15-24)	*Wave 2 (20-30)*	*Wave 1 (25-34)*	*Wave 2 (30-40)*	*Wave 1 (35-44)*	*Wave 2 (40-49)*
Completed Years Education for Youngest Cohort in Region	0.012 ** (0.004)	0.012 ** (0.004)				
Completed Years Education for Middle Cohort in Region			0.026 *** (0.004)	0.023 *** (0.005)		
Completed Years Education for Oldest Cohort in Region					0.029 *** (0.006)	0.021 * (0.009)
Proportion with Primary Education in the Region among Youngest Cohort	0.103 ** (0.035)	0.166 ** (0.057)				
Proportion with Primary Education in the Region among Middle Cohort			0.270 *** (0.042)	0.243 *** (0.051)		
Proportion with Primary Education in the Region among Oldest Cohort					0.308 *** (0.065)	0.288 * (0.117)
N	91	91	91	91	90	90

Based on adults aged 15-44 in first wave, regional HIV prevalence calculated using DHS provided HIV weights, regional education means calculated using provided sample weights, robust standard errors are in parentheses and additional controls are included for year of measurement and country.

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 7: OLS of Change in Regional Educational Outcomes of Children 7-14 on Regional HIV Prevalence of Adults (without migrants)

	Change in Percent Attending School in Region		
	All Children	Girls	Boys
HIV Prevalence among 15-49 Year Olds in Region in Wave 1	-0.265 (0.145)	-0.251 (0.194)	-0.232 (0.171)
Number of Regions	91	91	91
	Change in Average Years of Completed Education in Region		
	All Children	Girls	Boys
HIV Prevalence among 15-49 Year Olds in Region in Wave 1	-0.901 (0.702)	-0.216 (0.915)	-1.466 (0.908)
Number of Regions	91	91	91

Educational outcomes of children aged 7-14 taken from DHS household quesitonnairre and calculated using provided sample weights, Regional HIV prevalence calculated based on adults 15-49 and using HIV survey weights, robust standard errors in parentheses and additional controls for year of measurement and country is included but not shown; * p<0.05, **p<0.01, ***p<0.001

Table 8: Change in Regional HIV Prevalence on Regional Education Levels (without migrants)

	Regional Change in HIV Prevalence: *All Cohorts* (15-44 in Wave1)	Regional Change in HIV Prevalence: *Youngest Cohort* (15-24 in Wave1)	Regional Change in HIV Prevalence: *Middle Cohort* (25-34 in Wave1)
Completed Years Education of Adults 15-44 in First Wave	-0.006 * (0.002)		
Completed Years Education of Adults 15-24 in First Wave		-0.008 * (0.002)	
Completed Years Education of Adults 25-34 in First Wave			-0.003 (0.003)
Proportion of 15-44 yr olds with Primary Education in First Wave	-0.049 * (0.022)		
Proportion of 15-24 yr olds with Primary Education in First Wave		-0.067 * (0.031)	
Proportion of 25-34 yr olds with Primary Education in First Wave			-0.035 (0.031)
N	91	91	91

Based on adults aged 15-49, regional HIV prevalence calculated using DHS provided HIV weights, education means calculated using provided sample weights, robust standard errors in parentheses and additional controls included for year of measurement and country. * p<0.05, ** p<0.01, *** p<0.001

References

Akbulut-Yuksel, Mevlude, and Belgi Turan. "Left Behind: Intergenerational Transmission of Human Capital in the Midst of HIV/AIDS." *SSRN eLibrary*. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1675637.

Anglewicz. 2007. Migration, Risk Perception , and HIV infection in Malawi. January 1. http://repository.upenn.edu/dissertations/AAI3292004/.

Behrman, J., Kohler, H.-P. & Castro, R., 2010. Perception of HIV risk and the quantity and quality of children: The case of rural Malawi. *PSC Working Paper Series*. Available at: http://repository.upenn.edu/psc_working_papers/23.

Bennell, P., Hyde, K., and N. Swainson 2002."The Impact of the HIV/AIDS Epidemic on the Education in Sector in sub-Sahran Africa" Center for International Education. University of Sussex

Bongaarts, John. 2007. "Late Marriage and the HIV Epidemic in Sub-Saharan Africa." *Population Studies* 61 (1) (March 1): 73-83.

Case, A. & Paxson, C., 2011. The Changing Role of Education on the AIDS Crisis in Africa.

Case, Anne & Paxson, Christina, 2011. The Impact of the AIDS Pandemic on Health Services in Africa: Evidence from Demographic and Health Surveys. *Demography*, 48(2), pp.675-697.

Clark, Samuel J., Mark A. Collinson, Kathleen Kahn, Kyle Drullinger, and Stephen M. Tollman. 2007. "Returning home to die: Circular labour migration and mortality in South Africa 1." *Scandinavian Journal of Public Health* 35 (69 suppl): 35 -44. doi:10.1080/14034950701355619.

Crimmins, E.M. & Saito, Y., 2001. Trends in healthy life expectancy in the United States, 1970-1990: gender, racial, and educational differences. *Social Science & Medicine*, 52(11), pp.1629-1641.

Cutler, D.M. & Lleras-Muney, A., 2006. Education and Health: Evaluating Theories and Evidence. *National Bureau of Economic Research Working Paper Series*, No. 12352. Available at: http://www.nber.org/papers/w12352 [Accessed March 31, 2011].

De Walque, D., 2007. How does the impact of an HIV/AIDS information campaign vary with educational attainment? Evidence from rural Uganda. *Journal of Development Economics*, 84(2), pp.686-714.

Feldman, J.J. et al., 1989. National trends in educational differentials in mortality. *American Journal of Epidemiology*, 129(5), pp.919 -933.

Fortson, J., 2011. Mortality risk and human capital investment: The impact of HIV/AIDS in sub-Sahran Africa. *The Review of Economics and Statistics*, 93(1), pp.1-15.

Fortson, J., 2009. HIV/AIDS and fertility. *American Economic Journal: Applied Economics*, 1(3), pp.170–194.

Fortson, J., 2008. The gradient in sub-Saharan Africa: socioeconomic status and HIV/AIDS. *Demography*, 45(2), pp.303–322.

Gachuhi, D. 1999. The impact of HIV/AIDS on the education system in the Eastern and Southern Africa region. Johannesburg: prepared for UNICEF.

Gregson, S., Waddell, H. & Chandiwana, S., 2001. School education and HIV control in sub-Saharan Africa: from discord to harmony? *Journal of International Development*, 13(4), pp.467–485.

Haacker, M. 2002. The economic consequences of HIV /AIDS in Southern Africa. IMF Working Paper, Vol. , pp. 1-41, 2002. Available at SSRN: http://ssrn.com/abstract=879415

Hargreaves, J.R. & Glynn, J.R., 2002. Educational attainment and HIV-1 infection in developing countries: a systematic review. *Tropical Medicine & International Health*, 7(6), pp.489-498.

Mishra, V. et al., 2006. HIV Testing in National Population Based Surveys: Experience from the Demographic and Health Surveys. *Bulletin of the World Health Organization*, 84(7), pp.537-545.

Preston, S.H. & Elo, I.T., 1995. Are Educational Differentials in Adult Mortality Increasing in the United States? *Journal of Aging and Health*, 7(4), pp.476 -496.

Ross, C.E. & Wu, C.-ling, 1995. The Links Between Education and Health. *American Sociological Review*, 60(5), pp.719-745.

Singh, S. & Samara, R., 1996. Early Marriage Among Women in Developing Countries. *International Family Planning Perspectives*, 22(4), pp.148-175.

Table 1: Regional HIV and Educational Outcomes by Country and Wave

		Kenya (2003)	Kenya (2009)	Mali (2001)	Mali (2006)	Tanzania (2003)	Tanzania (2008)	Zambia (2001)	Zambia (2007)	Total (Wave 1)	Total (Wave 2)
HIV Rates	Mean	0.076	0.063	0.015	0.012	0.081	0.069	0.173	0.154	0.087	0.075
	Standard Error	0.010	0.015	0.003	0.002	0.010	0.009	0.018	0.014	0.008	0.007
	25th Percentile	0.041	0.033	0.004	0.007	0.041	0.031	0.102	0.117	0.035	0.021
	75th Percentile	0.114	0.079	0.021	0.017	0.107	0.097	0.224	0.199	0.117	0.116
Adult Years of Completed Education	Mean	7.28	8.14	2.15	2.25	6.08	6.08	6.90	7.24	5.75	5.97
	Standard Error	0.66	0.59	0.37	0.36	0.18	0.19	0.39	0.38	0.25	0.25
	25th Percentile	6.45	7.24	0.86	1.04	4.86	5.09	5.39	5.57	4.68	5.01
	75th Percentile	9.53	9.67	3.00	3.25	7.13	7.24	8.32	8.59	7.29	7.56
Proportion Adults with Primary Education	Mean	0.57	0.65	0.19	0.19	0.69	0.68	0.61	0.64	0.57	0.58
	Standard Error	0.06	0.06	0.02	0.02	0.03	0.03	0.05	0.04	0.03	0.03
	25th Percentile	0.44	0.52	0.06	0.07	0.55	0.57	0.43	0.43	0.43	0.43
	75th Percentile	0.76	0.81	0.28	0.29	0.80	0.79	0.79	0.79	0.77	0.79
Number of Regions		15	15	16	16	42	42	18	18	91	91

All regional means calculated based on men and women aged 15-49, HIV prevalence calculated using DHS supplied HIV weights, Educational means calculated using DHS sample weights

Table 2: Sample Demographics by Country and Wave

	Kenya 2003	Kenya 2009	Mali 2001	Mali 2006	Tanzania 2003	Tanzania 2008	Zambia 2001	Zambia 2007	Total
Adults Aged 15-49									
Mean Age	27.90	28.33	28.96	28.50	28.45	28.09	27.54	28.11	28.30
	(0.07)	(0.07)	(0.06)	(0.06)	(0.08)	(0.08)	(0.07)	(0.08)	(0.02)
Mean Yrs Education	7.45	7.85	1.80	2.16	5.77	5.73	6.32	7.08	4.99
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.01)
Completed Primary Education	0.58	0.61	0.15	0.18	0.64	0.65	0.55	0.62	0.44
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
N	17,464	16,937	25,925	29,442	13,742	14,167	16,469	14,970	149,116
Percent of Total Sample	12	11	17	20	9	10	11	10	100
Children Aged 7-14									
Mean Age	10.49	10.41	10.26	10.22	10.34	10.40	10.33	10.44	10.34
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.01)
Mean Yrs Education	2.42	2.97	1.02	1.47	2.46	2.06	2.11	2.50	1.96
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)
Percent Attending School	0.72	0.83	0.35	0.42	0.76	0.74	0.67	0.76	0.60
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
N	8,313	8,608	15,645	17,087	7,369	7,656	8,951	8,675	82,304
Percent of Total Sample	10	10	19	21	9	9	11	11	100

Standard errors in parentheses, Means calculated using DHS sample weights

Table 3: OLS of Regional HIV Prevalence on Regional Education by Cohort and Wave

	Regional HIV Prevalence Youngest Cohort		HIV Prevalence Middle Cohort		HIV Prevalence Oldest Cohort	
	Wave 1 (15-24)	*Wave 2 (20-30)*	*Wave 1 (25-34)*	*Wave 2 (30-40)*	*Wave 1 (35-44)*	*Wave 2 (40-49)*
Completed Years Education for Youngest Cohort in Region	0.012 **	0.011 **				
	(0.003)	(0.003)				
Completed Years Education for Middle Cohort in Region			0.026 ***	0.021 **		
			(0.004)	(0.005)		
Completed Years Education for Oldest Cohort in Region					0.030 ***	0.020 **
					(0.006)	(0.007)
Proportion with Primary Education in the Region among Youngest Cohort	0.103 **	0.126 ***				
	(0.035)	(0.034)				
Proportion with Primary Education in the Region among Middle Cohort			0.270 ***	0.259 ***		
			(0.037)	(0.052)		
Proportion with Primary Education in the Region among Oldest Cohort					0.313 ***	0.232 ***
					(0.056)	(0.084)
N	91	91	91	91	90	90

Based on adults aged 15-44 in first wave, regional HIV prevalence calculated using DHS provided HIV weights, regional education means calculated using provided sample weights, robust standard errors are in parentheses and additional controls are included for year of measurement and country. One region with a small population had no individuals in the oldest age group tested for HIV in the first wave, that region was also dropped from the second wave analysis.

Table 4: OLS of Change in Regional Educational Outcomes of Children 7-14 on Regional HIV Prevalence of Adults

	Change in Percent Attending School in Region		
	All Children	Girls	Boys
HIV Prevalence among 15-49 Year Olds in Region in Wave 1	-0.194 (0.114)	-0.161 (0.161)	-0.185 (0.142)
Number of Regions	91	91	91

	Change in Average Years of Completed Education in Region		
	All Children	Girls	Boys
HIV Prevalence among 15-49 Year Olds in Region in Wave 1	-0.907 (0.584)	-0.152 (0.817)	-1.550 (1.270)
Number of Regions	91	91	91

Educational outcomes of children aged 7-14 taken from DHS household questionnaire and calculated using provided sample weights, Regional HIV prevalence calculated based on adults 15-49 and using HIV survey weights, robust standard errors in parentheses and additional controls for year of measurement and country are included but not shown; * p<0.05, **p<0.01, ***p<0.001

Table 5: Change in Regional HIV Prevalence on Regional Education Levels

	Regional Change in HIV Prevalence: *All Cohorts* (15-44 in Wave1)		Regional Change in HIV Prevalence: *Youngest Cohort* (15-24 in Wave1)		Regional Change in HIV Prevalence: *Middle Cohort* (25-34 in Wave1)	
Completed Years Education of Adults 15-44 in First Wave	-0.007	***				
	(0.002)					
Completed Years Education of Adults 15-24 in First Wave			-0.006	*		
			(0.003)			
Completed Years Education of Adults 25-34 in First Wave					-0.010	*
					(0.005)	
Proportion of 15-44 yr olds with Primary Education in First Wave	-0.059	**				
	(0.023)					
Proportion of 15-24 yr olds with Primary Education in First Wave			-0.054	*		
			(0.031)			
Proportion of 25-34 yr olds with Primary Education in First Wave					-0.100	
					(0.053)	
N	91		91		91	

Based on adults aged 15-49, regional HIV prevalence calculated using DHS provided HIV weights, education means calculated using provided sample weights, robust standard errors in parentheses and additional controls are included for year of measurement and country. '+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

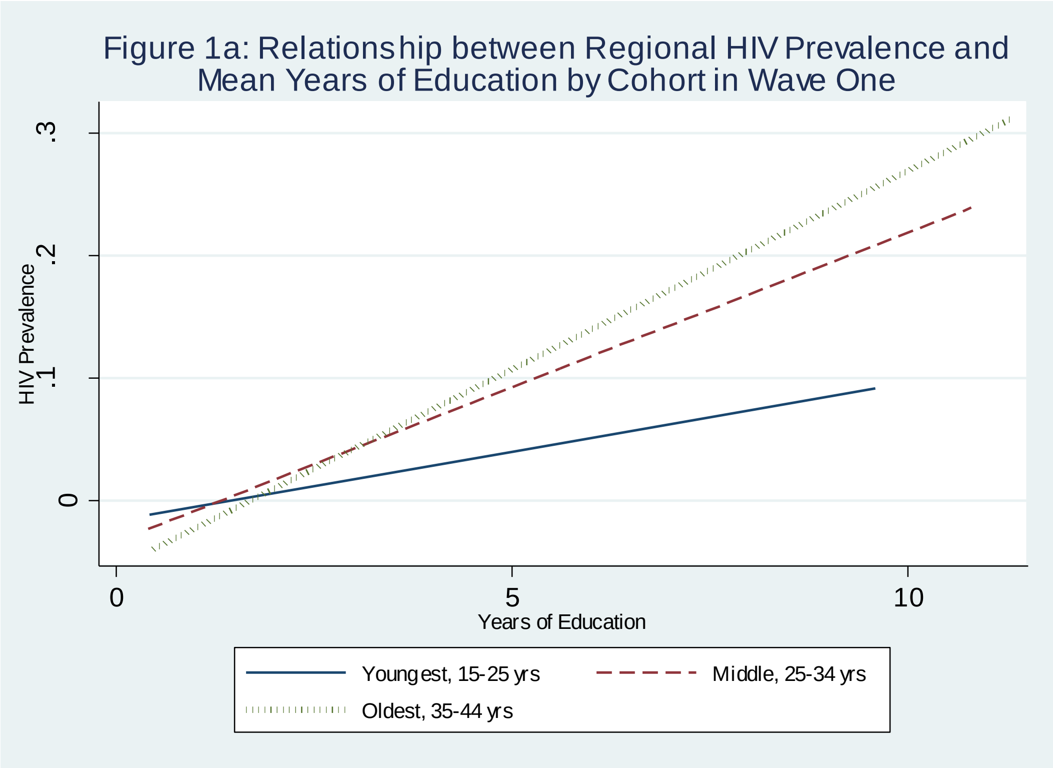
Figure 1a: Relationship between Regional HIV Prevalence and
Mean Years of Education by Cohort in Wave One
HIV Prevalence
0
.1
.2
.3
0
5
10
Years of Education
Youngest, 15-25 yrs
Middle, 25-34 yrs
Oldest, 35-44 yrs

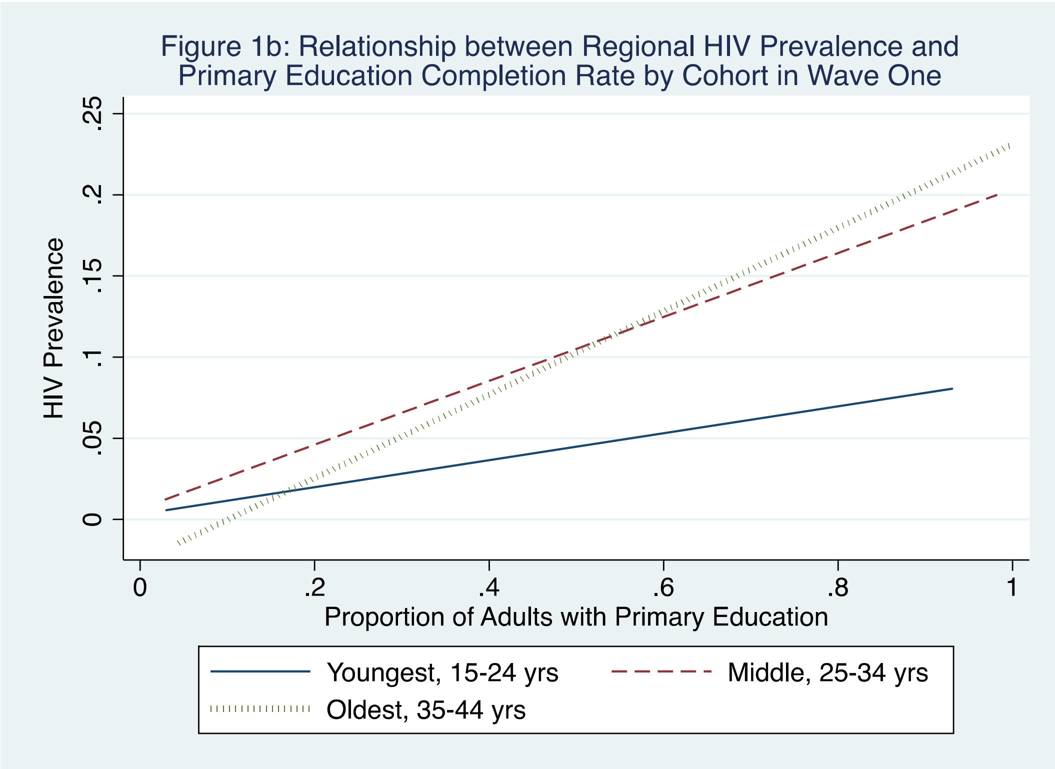
Figure 1b: Relationship between Regional HIV Prevalence and
Primary Education Completion Rate by Cohort in Wave One
HIV Prevalence
0
.05
.1
.15
.2
.25
0
.2
.4
.6
.8
1
Proportion of Adults with Primary Education
Youngest, 15-24 yrs
Middle, 25-34 yrs
Oldest, 35-44 yrs

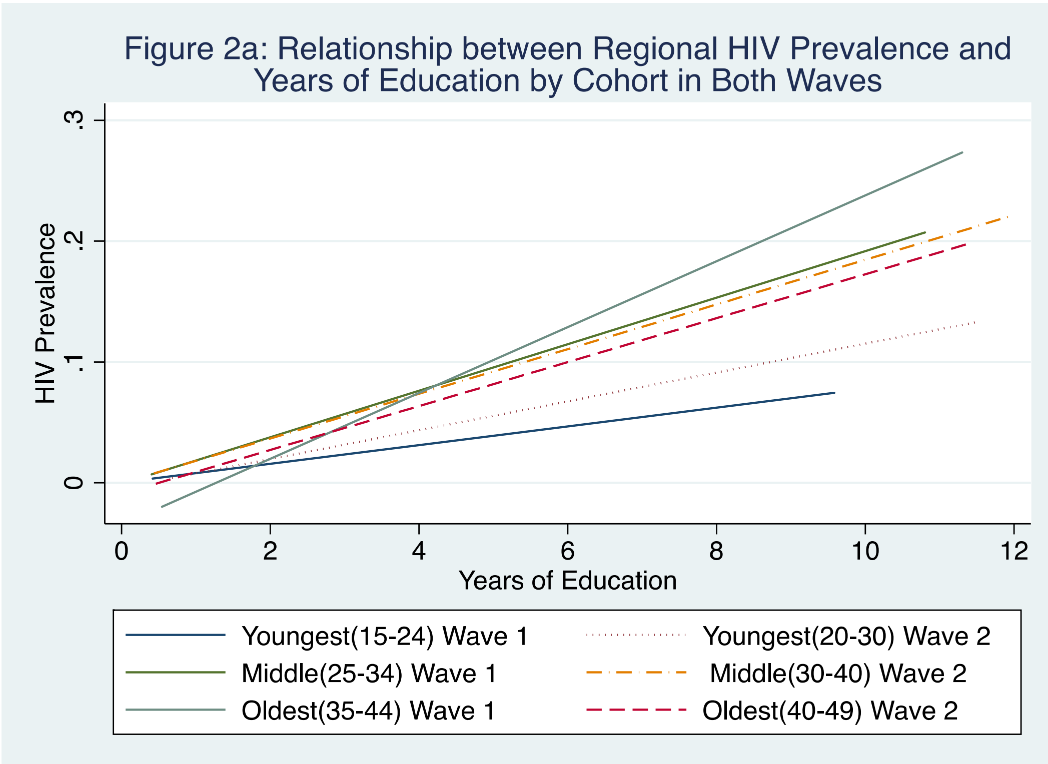

Figure 2a: Relationship between Regional HIV Prevalence and Years of Education by Cohort in Both Waves
HIV Prevalence
0
.1
.2
.3
0
2
4
6
8
10
12
Years of Education
Youngest(15-24) Wave 1
Middle(25-34) Wave 1
Oldest(35-44) Wave 1
Youngest(20-30) Wave 2
Middle(30-40) Wave 2
Oldest(40-49) Wave 2

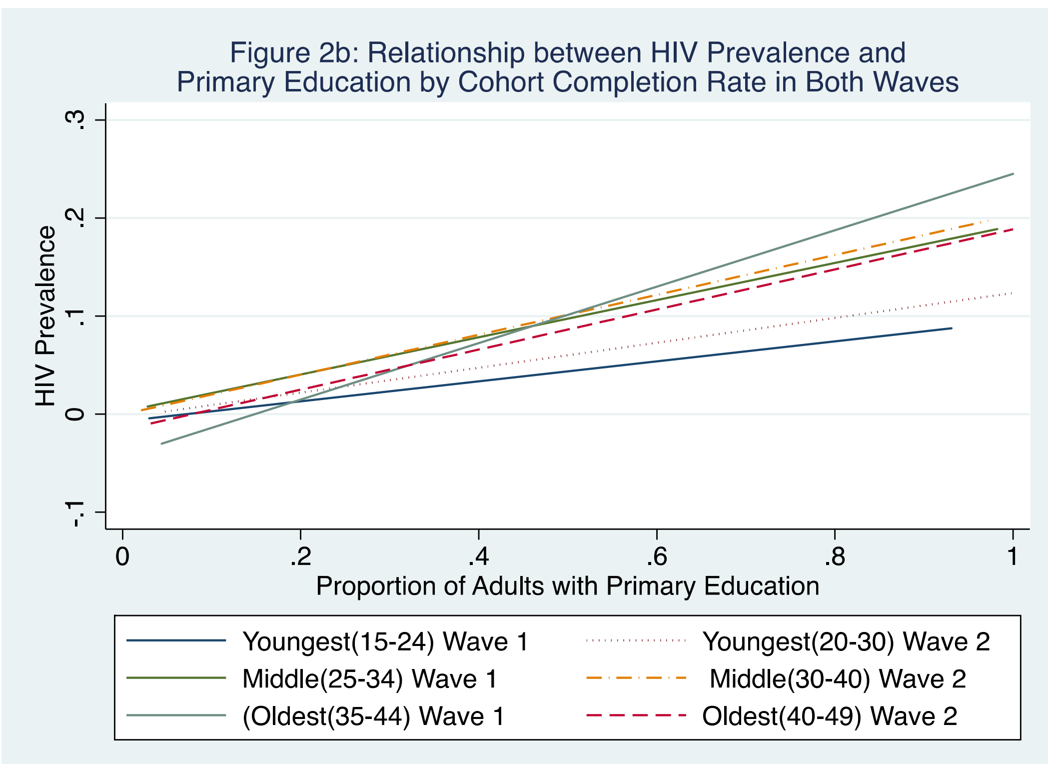

Figure 2b: Relationship between HIV Prevalence and Primary Education by Cohort Completion Rate in Both Waves
HIV Prevalence
-.1
0
.1
.2
.3
0
.2
.4
.6
.8
1
Proportion of Adults with Primary Education
Youngest(15-24) Wave 1
Middle(25-34) Wave 1
(Oldest(35-44) Wave 1
Youngest(20-30) Wave 2
Middle(30-40) Wave 2
Oldest(40-49) Wave 2

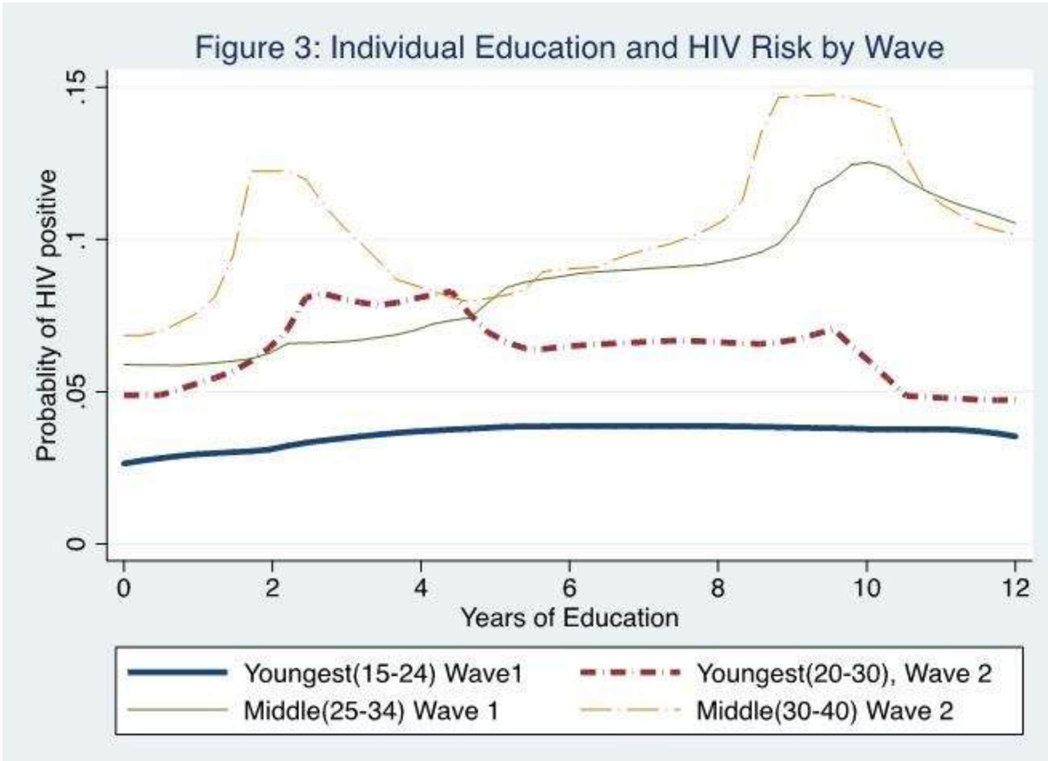
Figure 3: Individual Education and HIV Risk by Wave
Probablity of HIV positive
0
.05
.1
.15
0
2
4
6
8
10
12
Years of Education
Youngest(15-24) Wave1
Youngest(20-30), Wave 2
Middle(25-34) Wave 1
Middle(30-40) Wave 2

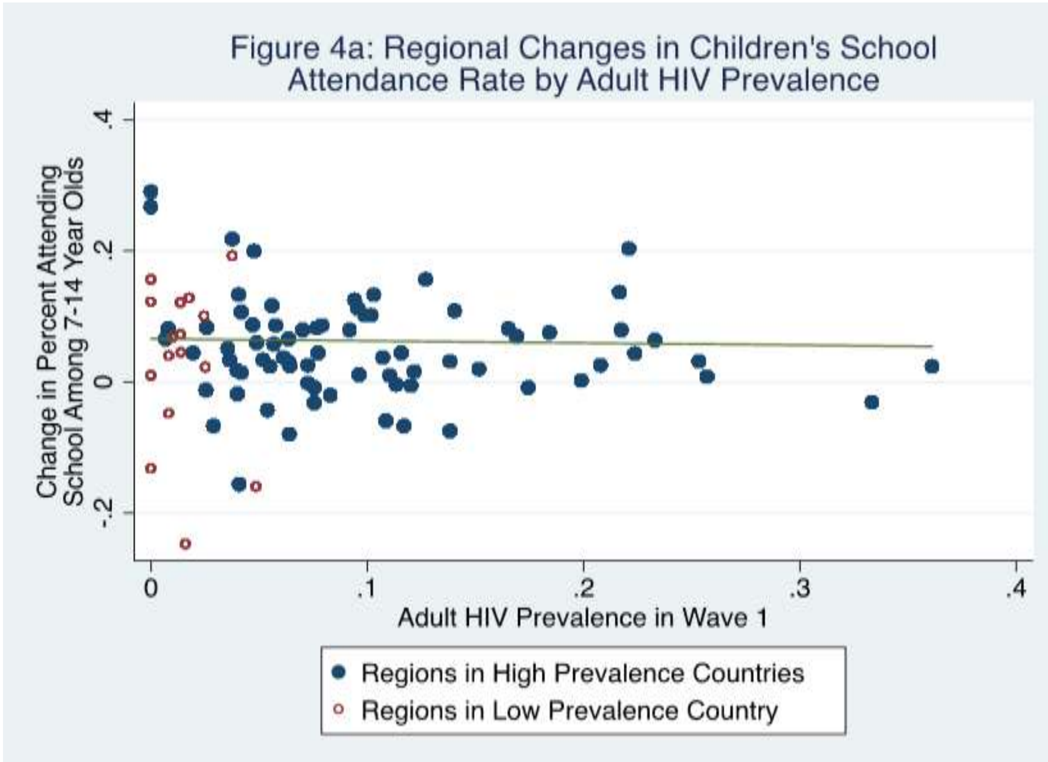
Figure 4a: Regional Changes in Children's School Attendance Rate by Adult HIV Prevalence
Change in Percent Attending School Among 7-14 Year Olds
-.2
0
.2
.4
0
.1
.2
.3
.4
Adult HIV Prevalence in Wave 1
Regions in High Prevalence Countries
Regions in Low Prevalence Country

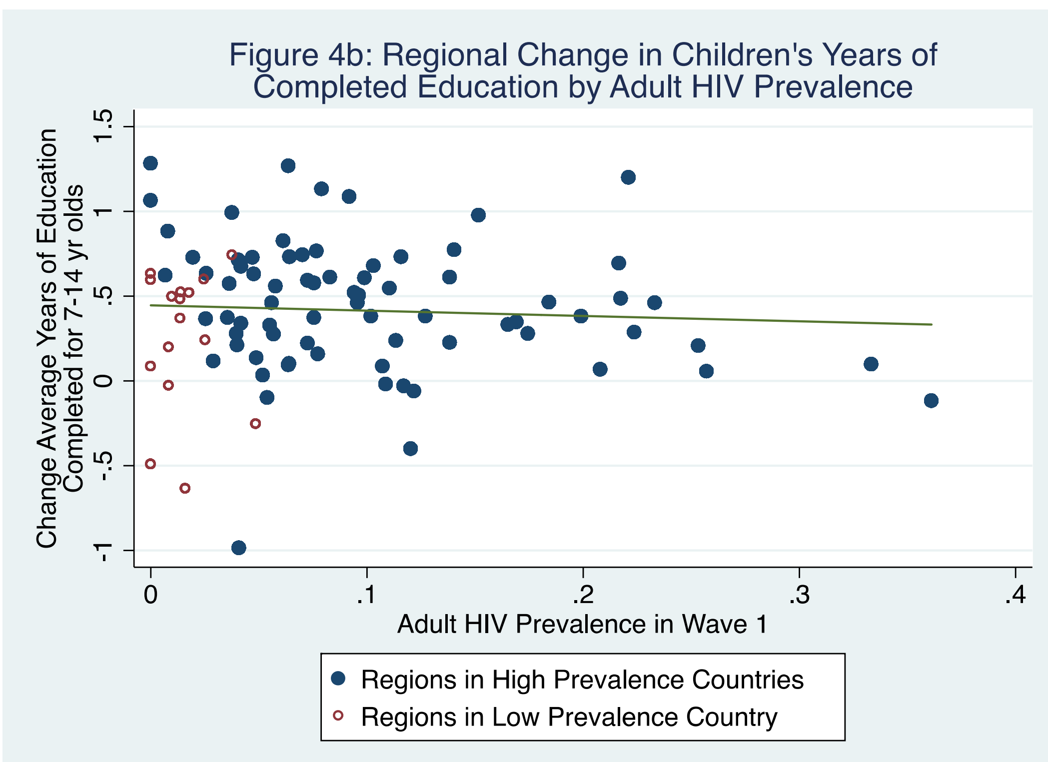
Figure 4b: Regional Change in Children's Years of Completed Education by Adult HIV Prevalence
Change Average Years of Education Completed for 7-14 yr olds
-1
-.5
0
.5
1
1.5
0
.1
.2
.3
.4
Adult HIV Prevalence in Wave 1
Regions in High Prevalence Countries
Regions in Low Prevalence Country

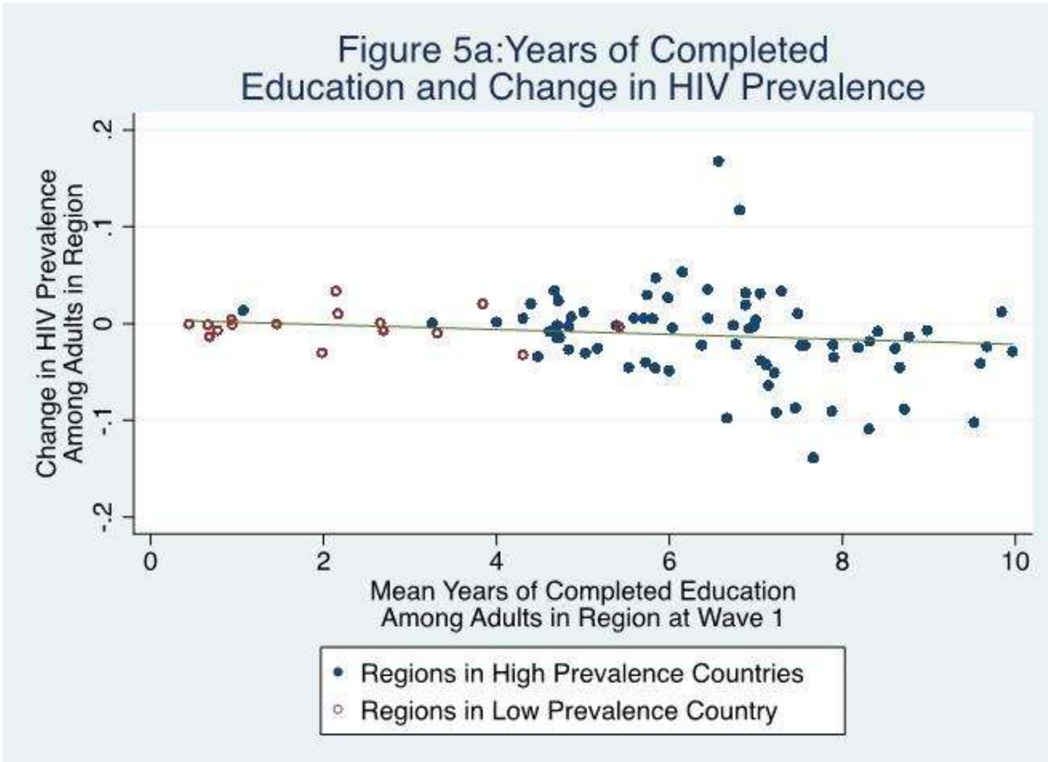
Figure 5a:Years of Completed
Education and Change in HIV Prevalence
Change in HIV Prevalence
Among Adults in Region
-.2
-.1
0
.1
.2
0
2
4
6
8
10
Mean Years of Completed Education
Among Adults in Region at Wave 1
Regions in High Prevalence Countries
Regions in Low Prevalence Country

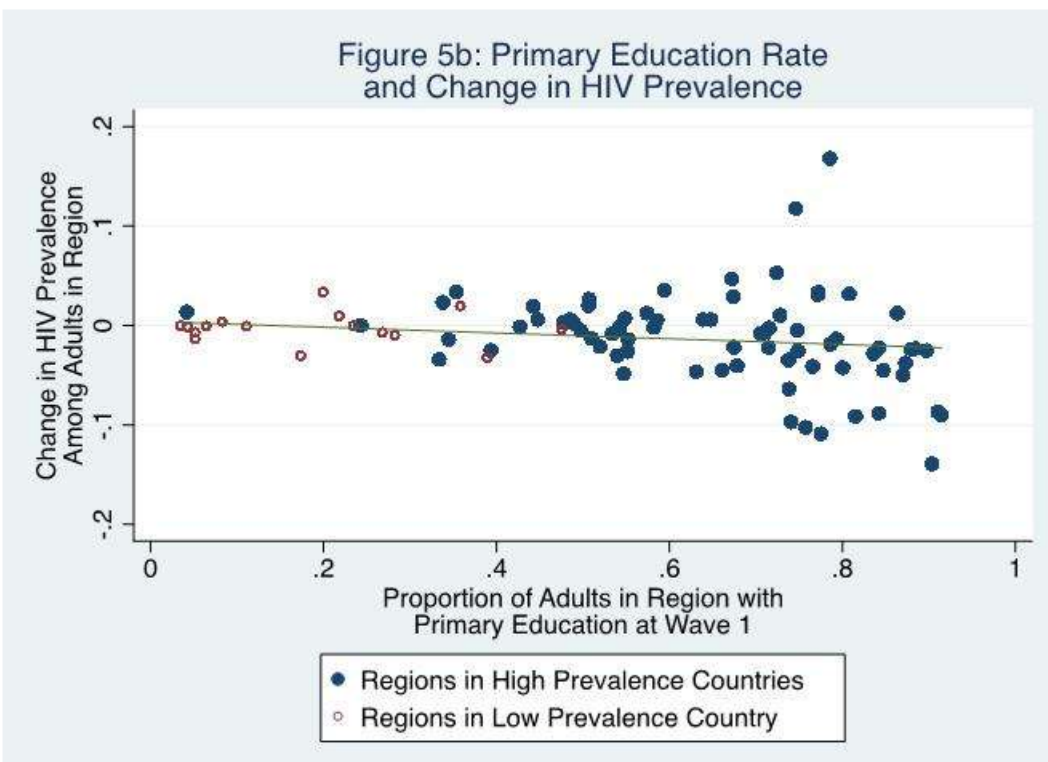
Figure 5b: Primary Education Rate
and Change in HIV Prevalence
Change in HIV Prevalence
Among Adults in Region
-.2
-.1
0
.1
.2
0
.2
.4
.6
.8
1
Proportion of Adults in Region with
Primary Education at Wave 1
Regions in High Prevalence Countries
Regions in Low Prevalence Country