Do the more educated know more about health? Evidence from schooling and HIV knowledge in Zimbabwe

Jorge Agüero & Prashant Bharadwaj *

June 2011

Abstract

This paper explores a fundamental link between education and health: knowledge about health. Do the educated *know* more about how certain diseases are spread and how to prevent themselves against it? Using age specific exposure to an education reform in Zimbabwe, we find that people with more years of schooling know more about how HIV spreads. Attending some secondary school raises the probability of having comprehensive knowledge of HIV by nearly 30 percent and raises the probability of knowing that HIV can be spread during birth by 11 percent. We find access to news media outlets as a mechanism by which the more educated might obtain more information since attending secondary school is associated with a 60 percent increase in the likelihood of reading newspapers. Perhaps surprisingly then, but consistent with the current literature, we find that such knowledge does not translate into better observed behavior regarding HIV prevention by the more educated. Along key outcomes such as probability of getting tested for HIV and actual HIV status, our instrumental variables estimate suggest that the more educated are not any different from the less educated. We provide a few hypothesis for this discrepancy.

*University of California, Riverside; University of California, San Diego. Thanks to Veena Jeevanandam and Stephie Fried for research assistance. Much thanks to Josh Graff Zivin and Paul Niehaus for engaging discussions on the topic.Emails: jorgea@ucr.edu; prbharadwaj@ucsd.edu.

1 Introduction

The relationship between education and health has received a lot of attention from researchers. While the positive association between health and education is well noted, there is little work on whether more education is causally related to better health investments, and even less work exists on what the mechanisms might be behind this relationship. In the developing world, policy focus has been on improving education and health. In this context, it is important to understand whether education enables people to make better health investments. This paper explores these questions in the important setting of HIV in Africa. The mechanism we explore in this paper is a seemingly obvious one: knowledge of HIV. Do the educated know more about how HIV is transmitted? If so, does this knowledge result in safer sexual practices or lower HIV rates among the more educated? We study these questions in Zimbabwe, where the estimated adult HIV prevalence rate is a staggering 14.3 (UNICEF 2010).[1]

Obtaining a causal interpretation to the positive correlation between education and health is fraught with difficulty. As is well noted by now, other explanations for this correlation include the idea that health improves education rather than the other way around, and that omitted factors related to both education and health are driving the result (Cutler and Lleras-Muney 2008). In order to causally examine this relationship and the mechanisms behind it, we use an instrumental variables approach that uses age specific exposure to an education reform in Zimbabwe. In 1980, after independence and after the restrictions limiting the advancement of black Zimbabweans were eliminated, enrollment into secondary school increased dramatically (see Figure 1). Since secondary school enrollment begins at age fourteen, we observe a large difference in completed education for people who were 14 and below as of 1980 as compared to people who were just 15 and over in 1980 (and as a consequence were too old to be enrolled in secondary school). Children who were 14 and below in 1980 obtain 1.5 years extra schooling and are 20 percent more likely to have attended some secondary school than their slightly older counterparts.

Using this discontinuity as an instrument for completed education, we find strong evidence that the educated know more about HIV and also know more about how to mitigate its transmission. Having gone to secondary school increases the probability of having comprehensive knowledge about HIV by nearly 40 percent. People who went to

[1] UNICEF Zimbabwe country statistics website as of June 17, 2011. URL: www.unicef.org/infobycountry/

secondary school are 18 percent more likely to know that a condom reduces the chances of getting HIV, and are also 30 percent more likely to know that HIV can be transmitted during breastfeeding.

However, what is it about education that makes people have more knowledge about HIV? Secondary school (grades 7 and up usually) is typically the time when sex education is taught in schools in many parts of the world. However, it is unlikely that sex education, especially with regards to HIV, was being taught in secondary schools in Zimbabwe in the 1980's. By many accounts, the National AIDS Coordination Program (NACP) was only set up in 1987 and the country's first HIV and AIDS policy was announced in 1999. We explore this issue by examining the use of various media outlets by educational status. We find that attending secondary schooling increases the likelihood of reading newspapers by 67 percent. This is an important finding as during the late 1980's and early 1990's, most of HIV/AIDS awareness was via newspapers. Writing in 1993, Pitts and Jackson (1993) note, " [...] national newspapers are the most important source of information reliably reaching the largest number of people. Newspapers have, in fact, been cited by social workers as the major source of information on AIDS and HIV in Zimbabwe." The more educated are also 25 percent more likely to watch television on a regular basis. Thus, it appears that access to media outlets play an important role in how the educated mange to get more information on HIV.

While the literature on the correlations between education and health is vast, there is little work on the causal relationship between education and health in the context of a developing country. Even in the developed country context, there are few studies that examine this relationship. Lleras-Muney (2005) and de Walque (2007a) are two of the recent papers to use instrumental variables strategy (changes in compulsory schooling laws and the Vietnam draft avoidance) to account for the endogeneity of education. They both find that education is positively related to health outcomes and behaviors; however, to our knowledge this paper is the first to examine this relationship in a causal framework in the specific context of HIV and in a developing country.

Even fewer papers consider the mechanisms that could explain the relationship (or lack thereof) between education and health. Lange (2011) is one of the few recent papers that examines how and why educated women in the US might make different investment decisions with regards to breast cancer screenings (although the variation in education in Lange's (2010) context is not exogenous). He finds evidence that the educated are more likely to include variation in risk factors when reporting personal cancer risk.

This supported by the fact that the more educated appear more likely to believe in scientific evidence regarding breast cancer. We directly examine such a mechanism by examining whether the more educated possess more knowledge of HIV.[2] While many recent papers have examined the role of information in inducing behavior change with regards to HIV (Thornton 2008, Dupas 2010, de Walque 2007b), these papers do not tackle the question of whether the educated possess more information about HIV. Understanding this link is a first step towards uncovering the ways in which education can affect health. In addition we are able to provide some evidence of *how* the more educated obtain more knowledge about HIV. Linking the use of various media outlets to educational status in our instrumental variables framework is a novel addition of our paper.

In important work, de Walque (2009) highlights trends in the relationship between education, HIV status and knowledge. He finds that education is positively related with behaviors that might mitigate the transmission of HIV (using condoms, talking to a spouse about HIV etc), but also predicts more infidelity and lower abstinence. While no significant association between education and HIV status is found in the overall population, he does find a negative association in urban settings. Using our instrumental variables strategy we find supporting evidence regarding the lack of a relationship between HIV status and education, suggesting that the results in de Walque (2009) are not just driven by omitted variables. We find no evidence that the more educated actually undertake healthier actions. For example, there does not seem to be a significant relationship between years of education and whether a condom was used during the most recent sexual activity.[3] The more educated are also not more likely to have been tested for HIV. Hence, while the data shows that education increases knowledge about HIV, education and knowledge per se does not appear to induce healthier behavior. Using data on actual HIV status, we find no significant relationship between years of education and HIV status, although, we caution that for this result the sample size is small and our first stage is rather weak. Nevertheless, the broad result from this paper suggests that while the more educated know more about HIV, this does not appear to translate to healthier behavior or significant differences in HIV positive status. Unfortunately, data limitations prevent us from conducting a rigorous analysis of *why* we find this relationship. We offer some preliminary explanations for this finding at the end of

[2]We also examine self reported risk of HIV, and find that the more educated hold no different beliefs than the less educated.

[3]It is not the case that the educated are more likely to be married. In the sample we examine almost everyone is married.

Section 5.

Our results suggest that policies hoping to bring about sexual behavior change should aim at providing information that is beyond what is currently known even by people with higher education. This is in line with the results of Dupas (2010), who suggests that detailed relative risk information is more effective than simple abstinence programs. Also, such policies aiming to bring about behavior change should not target just the less educated, as our results suggest that education and HIV status are not strongly related.

The remainder of this paper is organized as follows. Section 2 first describes the educational changes that occurred in Zimbabwe in 1980 that allows us to construct our instrumental variable design. Section 3 describes the data on HIV knowledge and behavior that we use in our specifications. Section 4 lays out the details of the identification strategy. We then move on to discuss the results in Section 5 and conclude in Section 6.

2 The 1980 education reform in Zimbabwe

In April 1980, the newly elected government of Zimbabwe reformed the education system to break with the apartheid-like regime that prevailed in Rhodesia.[4] Before 1980, at least 25 percent of black school-aged children failed to enter primary school due to a lack of places (Riddell 1980). For example, in 1976, for every 1,000 black school-aged children, 250 never started school. Of those who went to school, 377 graduated from primary school, but only 60 of them transitioned into secondary educaton. Thereafter, 37 reached Form IV and less than 3 reached lower Form VI (Nhundu 1992, p.79).[5]

The 1980 education reform has been widely documented in the literature (e.g., Edwards 1995, Edwards and Tisdell 1990, Dorsey 1989). As reported in Nhundu (1992), the were four key initiatives undertaken by the new government: (1) the introduction of free and compulsory primary education; (2) the removal of age restrictions to allow over-age

[4]For a history of Rhodesia's education system and the policies dictating the quantity and quality of schooling Africans received, see Atkinson (1972) and OCallaghan and Austin (1977).

[5]Zimbabwe's education system consists of primary education, secondary education and tertiary education. The primary level is a seven-year cycle and the official entry-age is six years. It runs from Grade 1 through Grade 7. Primary education leads to a Grade 7 certificate. Secondary education is divided into three two-year levels: junior, middle and high/advanced. Entering high/advanced secondary school requires the student to pass the "O"-level examinations.

children to enter school; (3) community support for education and; (4) automatic grade progression, in particular from primary to secondary school. An immediate impact of these steps was an enormous increase in school enrollment. Between 1979 and 1985, total enrollment rose from 885,801 to 2,698,878: a 205 percent increase (Nhundu 1992, p.82). The greatest expansion took place in secondary education where enrollment grew by 628 percent during the same period. As Figure 1 shows, gross enrollment in secondary schools climbed from 66,215 in 1979 to 482,000 in 1985, peaking at a little over 700,000 in 1991.

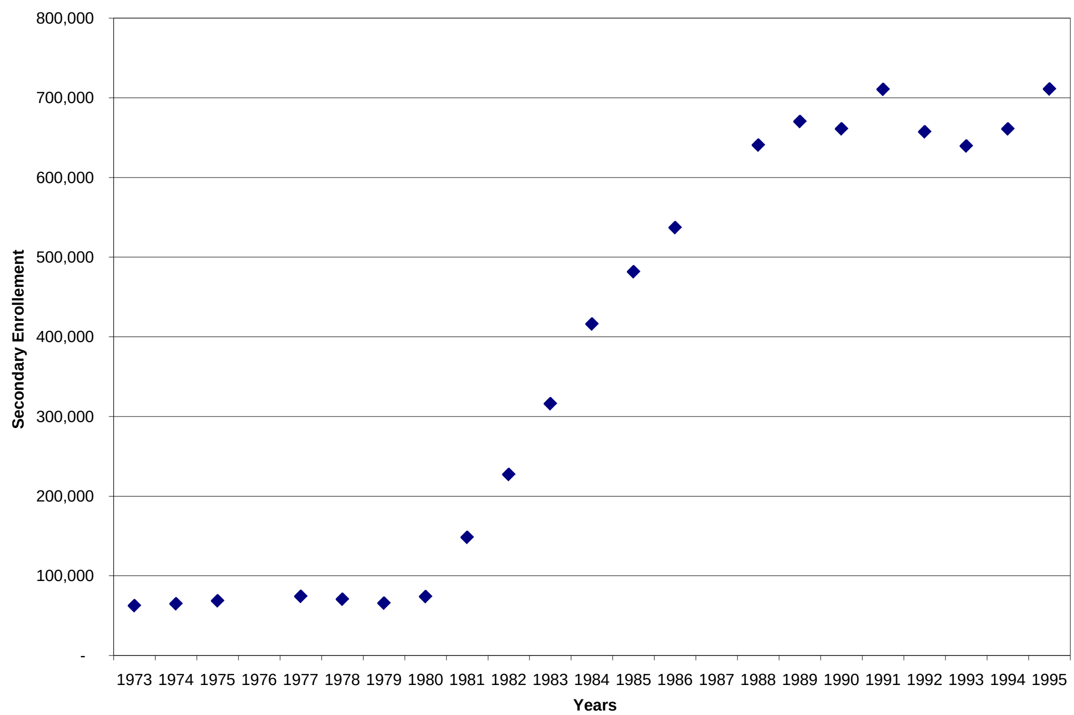


Figure 1: Secondary School Enrollments in Zimbabwe: 1973-1995
Source: United Nations, *Statistical Yearbook*, 1975, 1980, 1982, 1984, 1985-1989, 1992, 1994, 1995, and 1997.

To accommodate the increased demand, the government built new schools and undertook extensive reconstruction and expansion of existing facilities. This increase is shown in Figure 2. Between 1979 and 1983, the overall number of schools grew by 90 percent. Primary schools increased by 65 percent. Again, the largest increased is found in secondary schools: they grew by 575 percent since 1979. These figures are consistent with an increase in the budget allocated to education. In the fiscal year of 1979-80, the share of education was 11.6 percent in the national budget. It almost doubled in 1980-81

(22.1 percent), and remained at about 17 percent until 1986-87 (Dorsey 1989). The early years of the reform focussed on opening new secondary schools especially in rural areas. The target was to provide a secondary school within walking distance of all rural pupils, especially where geographic and demographic factors were conducive.

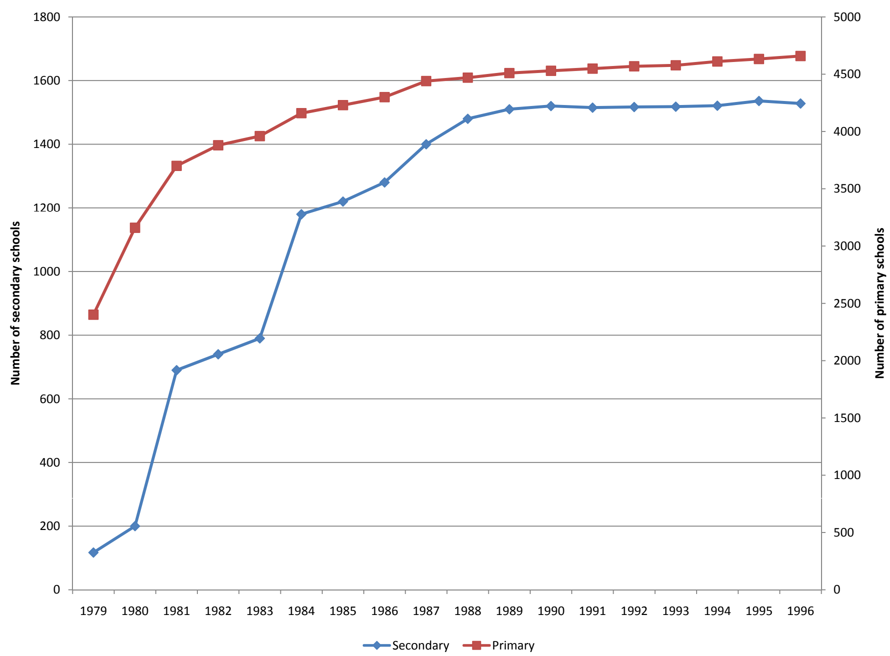


Figure 2: Trends in School Construction by Level, 1979-1996
Source: Zimbabwe Ministry of Education, Culture and Sports. *Annual Education Report*.

Mirroring the massive response in enrollments are the transition rates from primary to secondary in Zimbabwe's schools. As Figure 3 shows, the transition rate from Grade 7 (last grade of primary education) to Form I (first grade of secondary education) remained below 30 percent throughout the seventies. Beginning in 1980, the year of the reform, it jumped to 87 percent and averaged 70 percent for the rest of the decade.

Children start primary school at the age of seven, thus on-time completion of all primary grades would enable them to start secondary school at the age of 14. As shown in Figure 3 there is a clear discontinuity in the probability that a child would go to secondary school in 1980. A 14 year-old in 1979 had a one-in-five chances of enrolling in secondary school. A child only a year younger, a 14 year old in 1980, was four times as likely to enrol in secondary education. Therefore, the educational reform of 1980 provides a natural experiment, where for reasons exogenous to her choice, an adolescent would eventually acquire more schooling. We exploit the timing of the reform to ask whether the educated know more about HIV.

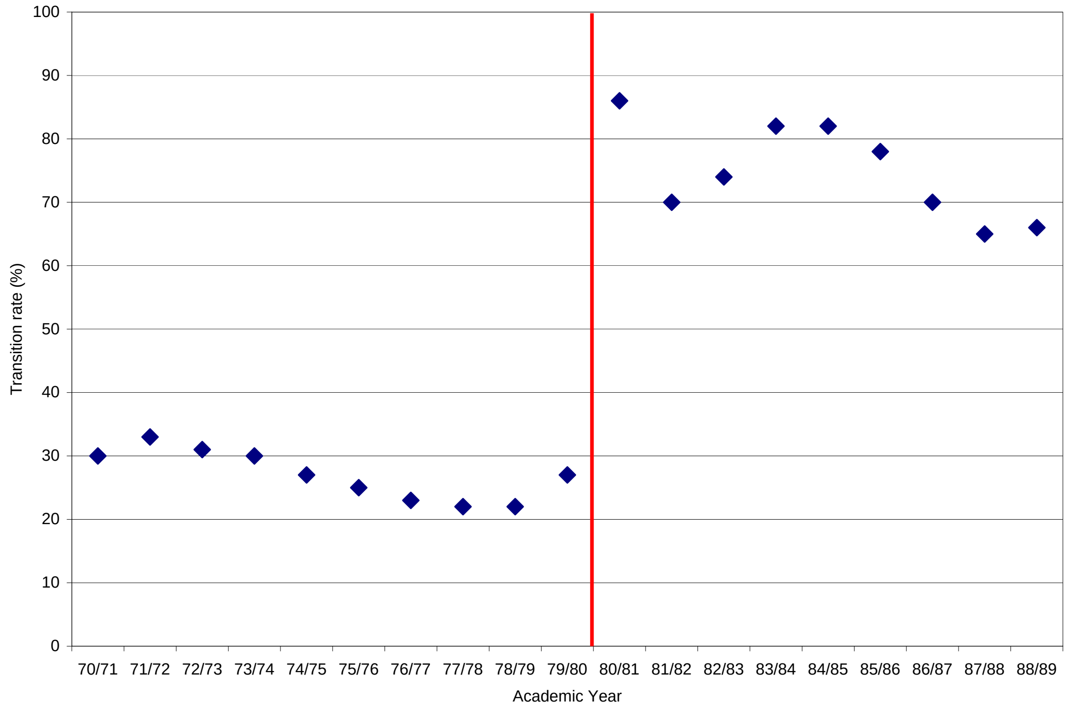


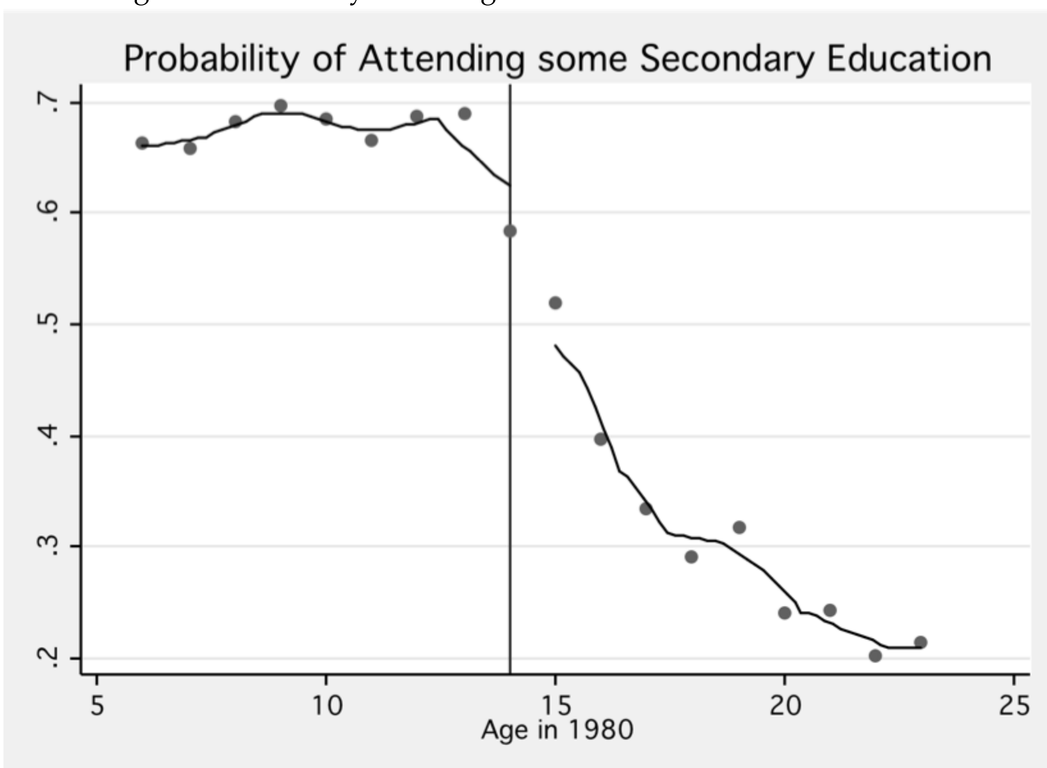
Figure 3: Grade 7 to Form 1 Transition Rates: 1970/71-1988/89
Source: Riddell and Nyagura (1991) Table 1.1. Note: Grade 7 is the last year of primary education and Form I is the first year of secondary education.

3 Data

The data sources for knowledge about HIV are the Demographic and Health Surveys (DHS) of Zimbabwe. The DHS are standardized nationally representative (cross-sectional) household surveys in developing countries. Women between the age of 15 and 49 answer a long questionnaire about their birth history, fertility preferences, family planning, their socio-economic and marital status, among other characteristics. Since the mid-1990s, the DHS asked women and their partners about their knowledge of HIV as well sexual practices. In the case of Zimbabwe only the 1999 and 2005-06 DHS contain this information and we focus on these surveys. Specifically, two components of the DHS from each year was used - the male recode and the individual recode (which records the responses of females). These four components were combined and used in the analysis.

Since the data sets contain year of birth and completed education as of the survey year, we are able to construct age as of 1980 (the beginning of the reform) and relate it to completed schooling for men and women. We find that people 14 or below as of 1980 obtain significantly more years of education as well as a higher probability of attending secondary school, in line with Figure 3 above.

Figure 4: Secondary schooling in Zimbabwe



In particular, the set of questions on HIV knowledge and related behavior asked in 1999 and 2005 were used. These include, for example, asking whether mosquitos are a source of HIV transmission, whether using condoms reduces the chances of getting HIV et cetera. For a full list of exact questions please see the Table 3. Actual HIV status is only available in the 2005 survey. Since our strategy involves examining ages around 14 years as of 1980, we use a bandwidth of eight years. While the entire sample (1999 and 2005 included) contains data on approximately 24,500 individuals, when restricted to individuals between ages 6-22 in 1980, we obtain our working sample of approximately 8,500. After dropping observations that report years of schooling strictly greater than 16 years, we are left with 8,432 individuals. In this sample on average, individuals have 7.7 years of education, 52 percent have secondary education or more, a third reside in urban areas and nearly all are married. In addition, we perform robustness checks for a bandwidth of 5 years as well (this halves our sample size to around 5000 individuals). Given data limitations, the larger bandwidth provides more precise estimates and more power in the first stage regressions (see Results section).

4 Identification Strategy

To obtain causal estimates of whether education leads to better learning about HIV transmission, we exploit the exogenous variation in schooling generated by the post-independence Zimbabwean reform discussed in section 2 above.

Consider the following relation between education and HIV-related knowledge:

$$H_i = \alpha + \beta S_i + x_i'\theta + e_i \tag{1}$$

where H_i is a measure of HIV knowledge for person i. We consider multiple measures for H_i as described in the previous section. S_i represents person i's education attainment (i.e., years of completed schooling or the probability that the highest education level is at least secondary). Thus β is the parameter of interest. Equation (1) includes a vector of characteristics (x_i') such as age in 1980, sex, survey fixed-effects, province fixed-effect and a binary variable equal to one if the person lives in an urban area and zero for rural.

The primary concern is that OLS estimates of β are biased. First, higher levels of ability lead to more education and better knowledge of health and other subjects. Thus, the observed correlation between education and knowledge is an artifact of the difficulty to control for ability. This implies an upward bias in the estimates of β using OLS. Second, if health improves education rather than the other way around it will also bias our estimates for (1). An ideal solution is to have a source of exogenous variation in schooling, which can be then used to test for a *causal* effect.

The 1980 reforms dramatically altered the schooling opportunities available to men and women young enough to take advantage of them. Figure 3 illustrates the sudden discontinuity in the transition rates induced by the reform allowing automatic admission to Form I after Grade 7. In itself, this implies a clearly different schooling experience for pupils aged fourteen or younger in 1980 compared to their seniors. However, the discontinuity in the probability of secondary-school enrollment is more *fuzzy* than sharp. [6]

This timing provides the source of exogenous variation in schooling to test for the causal effect on knowledge about health. Our estimation strategy compares the HIV-related knowledge of individuals aged 14 in 1980 to the corresponding knowledge of those aged 15 in 1980. As in van der Klaauw (2002), the indicator for the age cut-off serves as an excluded instrument in a Two-Stage Least-Squares (TSLS) regression. Formally, our first stage equation estimates the following relationship:

$$\begin{aligned} S_i &= \beta_1 DumAge_i + \beta_2 DumAge_i * (Age80_i - 14) \\ &\quad + \beta_3 (1 - DumAge_i) * (Age80_i - 14) + x_i'\theta + \epsilon_i \end{aligned} \tag{2}$$

Where $DumAge$ is a dummy variable taking on the value of 1 if age in 1980 is less than or equal to 14, and 0 otherwise. We use linear approximations on either side of the cutoff. Visual representation of the regression discontinuity (Figure 4) suggests that this might be a good fit.[7] Predicted values of S_i from this regression are then used to estimate our second stage:

$$\begin{aligned} H_i &= \beta \hat{S}_i + \beta_2 DumAge_i * (Age80_i - 14) \\ &\quad + \beta_3 (1 - DumAge_i) * (Age80_i - 14) + x_i'\theta + \epsilon_i \end{aligned} \tag{3}$$

Validity of the design rests on the assumption that children just below 14 years of age in

[6] Since the reforms also relaxed entry-age restrictions, they could have caused some overage enrollment at all grades of school. While this does not rule out the possibility of children older than 14 in 1980 attending secondary school, Figure 3 shows that is highly unlikely that overage students enroll.

[7] Results do show sensitivity to higher order polynomials. However, this is due to the levels of education observed by 14 and 15 year olds. Dropping these from the regression makes the results robust to higher order polynomials. This is not our preferred estimation strategy however, as we ideally would like to compare people close together in age.

1980 are similar in unobservable ways to children just above the age of 14 in 1980. Our IV approach will be invalid if other variables (except for education) exhibit a discontinuity around the cut-off point. However, to some extent we can test the validity of our design. For example, the education reform in Zimbabwe was targeted to the black population who suffered the apartheid-style regime before 1980. Thus, we should expect a discontinuity at the age of 14 in 1980 for blacks but not for white Zimbabweans. Unfortunately, the DHS did not collect information about ethnicity. However, Agüero and Ramachandran (2011), using a 10 percent micro sample of the 2002 population census in Zimbabwe validate this point. They show that for whites there is no discontinuity at the cut-off point.

The increase in educational attainment due to the reform should affect only those in Zimbabwe and not those in other African countries. Agüero and Ramachandran (2011), again using the 2002 census, show that foreigners living in Zimbabwe in 2002 also failed to show a discontinuity at 14 in 1980. We expand this idea and explore whether other neighboring countries exhibit a discontinuity for men and women at the age of 14 in 1980. In figure 5 we reproduce figure 4 for Zambia using the 1996 DHS. Observing a discontinuity at the cut-off for Zambia will reduce the validity of our identification and suggest that factors beyond the local education reform in Zimbabwe are driving the results in Figure 4. As observed in figure 5, there is no evidence of such discontinuity in Zambia. Figure 6 shows a lack of discontinuity for the 1998 South Africa DHS (but restricted to women only). These cases allows us to reject the possibility that regional factors in southern Africa explain the gains in education for those aged 14 or less in 1980 in Zimbabwe.

Figure 5: Schooling in Zambia

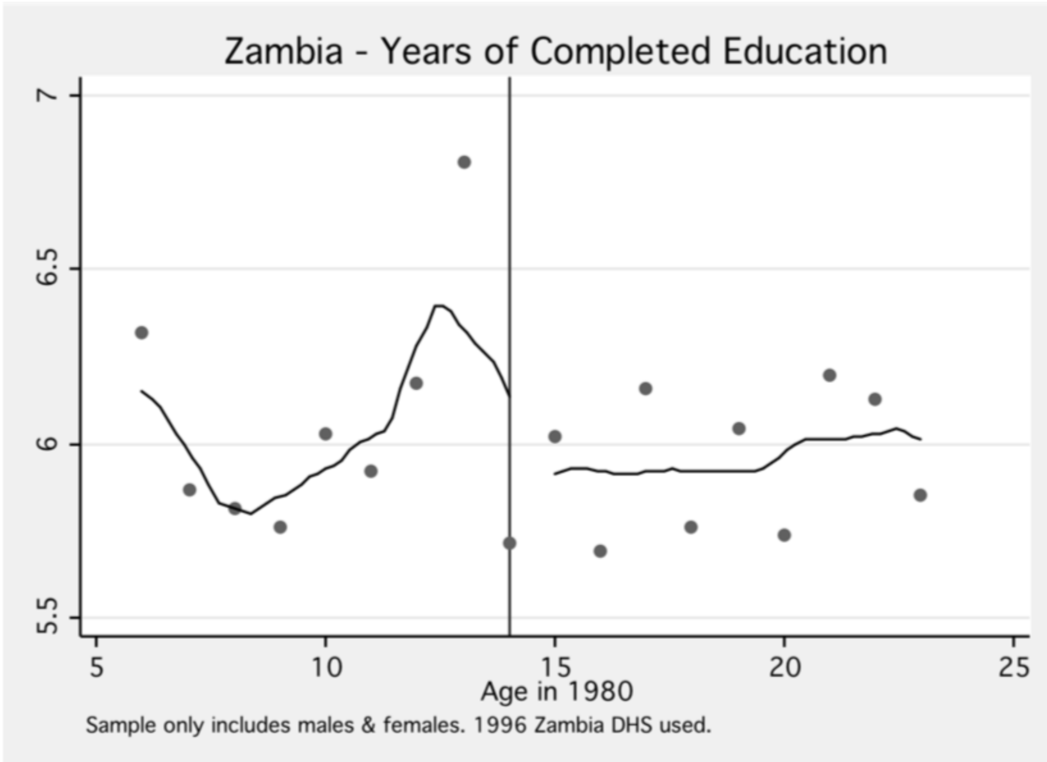
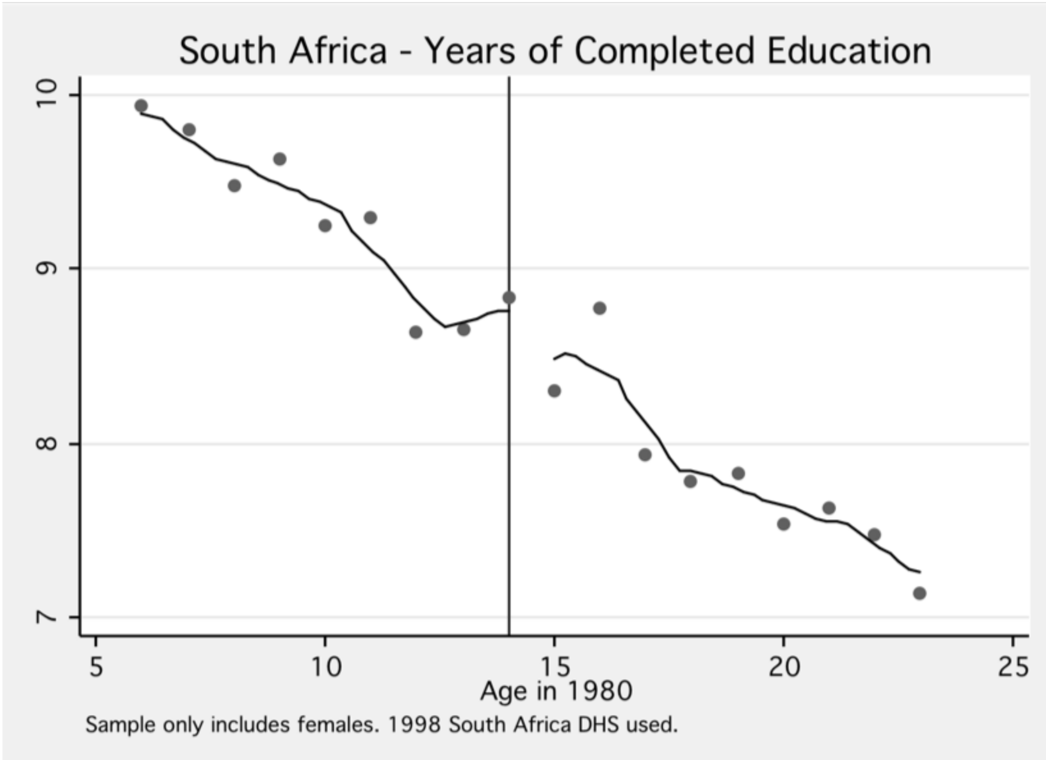


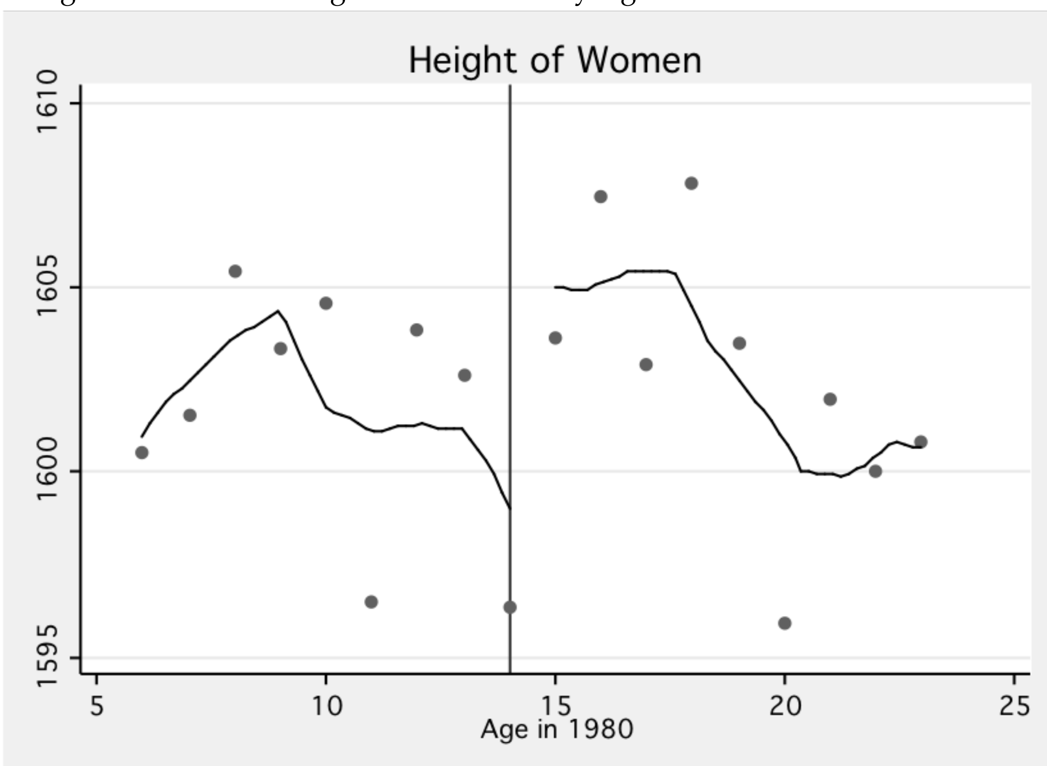
Figure 6: Schooling in South Africa



Finally, it is clear that there were other reforms in the country taking place at the same time (Palmer 1990, Kumaranayake, Lake, Mujinja, Hongoro, and Mpembeni 2000). Thus, it is possible that our estimation strategy is capturing the confounding effect of all other changes in addition to the educational reform. For example, it has been noted that the provision of health services increased with independence (Thomas and Maluccio 1996). Thus, the creation of new hospitals or medical centers might seem to confound the effect of the reform. However, this is unlikely to be the case because a hospital would not be able to discriminate those aged 15 in 1980 compared to those aged 14. Therefore, it is the fact that the reform affected those aged 14 or less in 1980 in a disproportional way relative to those aged 15 or more that allows us to identify the effects of education on knowledge. Nevertheless, we examine a proxy for health status around the age cut off to ensure that our results are not driven by changes other than education reform.

The DHS has objective measures of height for women (they do not collect this information for men). If women aged 14 or less were disproportionately benefiting from health or nutrition related services relative to those aged 15 or more in 1980 we should expect to observe height as an adult to exhibit a discontinuity around the cut-off. Figure 7 plots mean height (implied 1 decimal point in centimeters) by age as of 1980 for women in Zimbabwe. While there appears to be a drop in height for people exactly aged 14 years in 1980 (this drop is a mere 0.7 centimeters), the trends before and after age 14 do not show any sign of differential access to health services.

Figure 7: Women's Height in Zimbabwe by Age in 1980



5 Results

5.1 First stage results

Before turning to the results on knowledge and behavior, we first present some evidence that the education reform affected the relevant age groups (14 and below in 1980) in terms of increasing years of schooling as well as the probability of having attended secondary school. We have already shown some of this evidence in figures in sections 2 and 3; thus, we only discuss regression estimates in this section. As mentioned in section 4, the regression discontinuity approach dictates that we use the cutoff variable (1 if age 14 or less in 1980, 0 otherwise) as the instrument in our second stage regressions (equation 3).

Table 2 shows that the first stage is rather strong. We explore 2 different bandwidths, ages 6-22 in 1980 (columns 1 and 2) and ages 9-19 in 1980. As is usually the case with RD designs, for larger bandwidths, we get greater precision but we are also including points that are likely not comparable due to other factors. Hence, while we use the 6-22 age bandwidth for our main specifications, we also do robustness checks with the 9-19 age bandwidth. Columns 1 and 2 of Table 2 suggests that people who were 14 and below as of 1980 receive nearly 1.5 years of additional schooling and were 20 percent more likely to have attended secondary school. With a smaller bandwidth, these effects reduce to an additional 0.8 years of schooling and a 10 percent increase in the probability of having attended secondary school. In subsequent results we use the discontinuity as an instrument for the years of education completed or for the probability of having attended some secondary school. For all our specifications, the first stage F statistic is greater than the conventional threshold of 10. For some of the behavior related specifications however, we are forced to use smaller samples and the F statistic in these instances falls below 10. We note all fist stage F statistics in the notes accompanying the tables.

5.2 Education and Knowledge about HIV

Table 3 examines the impact of education on knowledge about HIV. We present results for questions that were the same across surveys to increase sample size (for smaller sample sizes we obtain a rather small first stage F-statistic). For each question about HIV knowledge, we present both OLS and IV results. As hypothesized in section 3, it is likely that the OLS estimates of how education affects knowledge will be tend to bias the estimates upwards if omitted variables like ability play an important role. A comparison of OLS and IV estimates in Table 3 suggests the *opposite* - OLS estimates in general appear to be smaller than the IV estimates. The returns to education literature also finds a similar pattern for the most part - OLS estimates on the returns to education tend to underestimate the returns as suggested by IV or other empirical methods. Hence, it is possible that other omitted variables play an important role in OLS estimation or that IV accounts for measurement error that could bias the estimates in OLS towards zero.

Table 3 in essentially shows that the education plays an important role when it comes to knowing the right answer to more difficult questions. For basic questions like "Have you heard about HIV?", OLS and IV estimates are small (the standard errors in the IV estimates are larger rendering it statistically insignificant) suggesting that education does not have a large causal impact for such simple questions.[8] However, as columns 3 and 4 in Table 3 reveal, not everyone has "Comprehensive Knowledge of HIV" (46 percent of the sample has correct comprehensive knowledge of HIV). Someone who correctly answers the following questions is deemed to have comprehensive knowledge:

1. Can a healthy person have HIV?
2. Does a condom reduce the chance of getting HIV?
3. Can HIV be passed on via mosquitos?
4. Is HIV transmission limited by having only 1 sexual partner?

When it comes to comprehensive knowledge, education seems to play a very important role. Indeed, the IV estimates are almost twice the OLS estimates, and even the OLS estimates are quite large compared to the mean. Having attended secondary school increases the probability of having comprehensive knowledge by 0.38. Considering a mean of 0.46, this is a substantial effect over the mean. Examining the individual questions that make up comprehensive knowledge (columns 5-10 and 15-16) confirms the idea that education does indeed play a causal role in knowing about HIV.

Education also plays an important role in knowing about HIV transmission and preg-

[8] There could be concerns about using a linear probability model for this particular dependent variable since it appears extreme valued. Probit and IV probit estimates however reveal the same pattern.

nancy. Columns 11 and 12 examine the question of whether HIV can be passed on during breast feeding. The IV estimates are much larger than the OLS estimates, and considering that only 68 percent of people know the right answer, the effect is nearly 50 percent over the mean. Overall, the results in Table 3 suggest a large and causal relationship between schooling and knowledge about HIV. Appendix Tables 1 and 2 examine the same outcomes as in Table 3 with different bandwidths and by dropping ages 14 and 15 from the analysis. WIth a smaller bandwidth, the results are more imprecise as we get a weaker first stage estimate. The results however, do indicate the the magnitudes do not change much. When we drop ages 14 and 15 from the analysis, we get a much stronger first stage and once again, the results are qualitatively similar to that in Table 3.

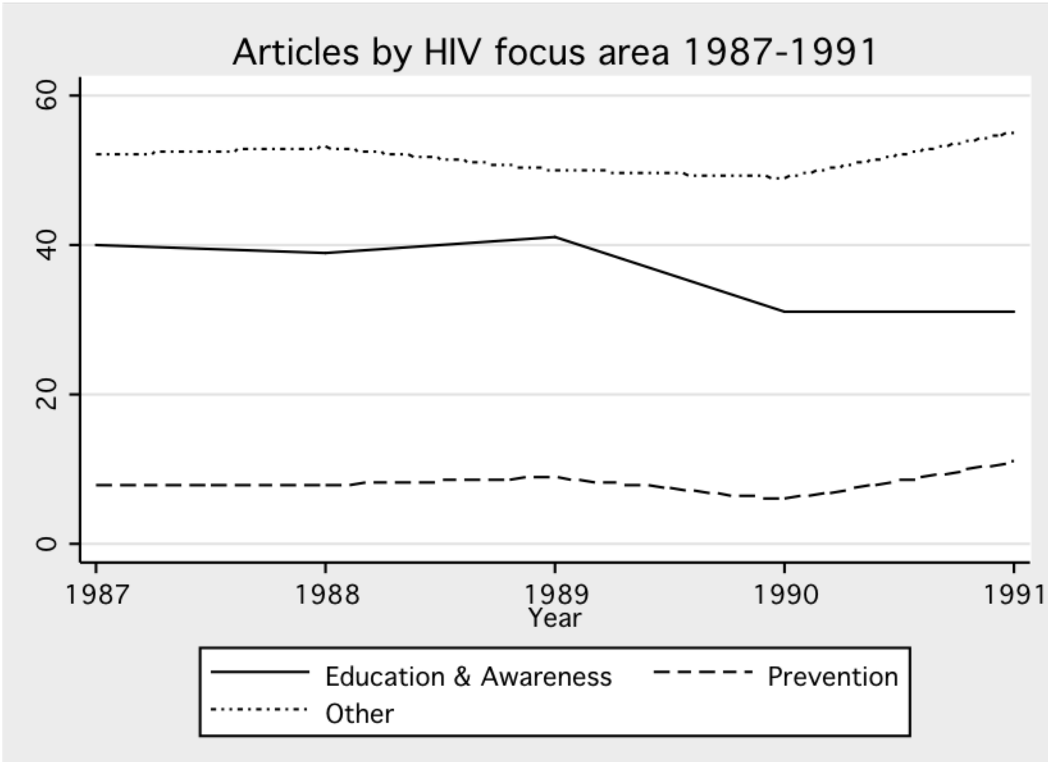
As mentioned in the introduction, it is not immediately clear *how* the more educated gather more information about HIV. It is unlikely that this information was learned in sex education classes in school in the 1980's. In fact the first reported cases of HIV in Zimbabwe date to the mid 1980's (USAID).[9] Table 4 examines the relationship between educational status and the use of various media outlets where one could learn about HIV.[10]

Attending secondary school significantly raises the possibility of reading newspapers (IV magnitudes suggest a 67 percent increase in the likelihood of reading newspapers). Moreover, the IV estimates are almost twice the OLS estimates. In one of the few studies on the topic from the late 1980's and early 1990's, Pitts and Jackson (1993) find that newspapers were the most important medium for conveying information about HIV/AIDS. Pertinent to our findings on knowledge and behavior, Pitts and Jackson (1993) note that the focus of these newspaper articles was more on education and awareness of HIV, rather than prevention. Below, we reproduce a graph from their study suggesting that the more educated were able to assimilate more information about HIV via such newspaper articles, but likely did not gain additional knowledge about prevention. This might explain the results we find with regards to behavior.

[9] Web page as viewed on June 17, 2011. URL: www.usaid.gov/our_work/global_health/aids/Countries/africa/zimbabwe.html

[10] A more recent source of information on HIV are antenatal clinics. Antenatal clinics often provide pregnant mothers information about HIV transmission and its role during the birth and breastfeeding process. Unfortunately, the question of whether antenatal clinics provided information on HIV was only asked of women in the 2005 and as a result examining whether more educated mothers attend antenatal clinics where such information might be provided leads to extremely small sample sizes for IV estimation.

Figure 8: Articles in Zimbabwean Newspapers - Pitts and Jackson (1993)



A United Nations Population Fund (UNFPA) country fact sheet on HIV from ZImbabwe in 2009 reports that, "Historical mapping also found consistent high media coverage since the early 1990s and high coverage of STI treatment since the late 1980s."[11] Thus, it appears that the more educated might know more about HIV via more frequent media use.

5.3 Education and HIV related behaviors

Does the acquired knowledge, as shown in the previous section, translate into better preventive action by the more educated? As alluded to in the introduction, this is a rather difficult question to pin down as the questions about knowledge might not be

[11]URL as viewed on June 17, 2011. http://countryoffice.unfpa.org/zimbabwe/?publications=3216

the ones that people use to judge risk of transmission *et cetera*. Moreover, the detailed questions on HIV avoidance behavior is asked in the 2005-2006 rounds of the DHS, reducing our sample size by a significant amount. With these caveats in mind, we present in Table 5 results on education and HIV related behavior for a subset of questions even though for some specifications our first stage F statistic falls below conventionally accepted levels.

Using the IV estimates, it appears that having attended secondary school has a small and insignificant impact on number of lifetime sexual partners, but rather small and statistically insignificant effects for whether a condom was used the last time the person had sex, on whether the person was tested for HIV, or whether the person actually tested positive for HIV. The OLS and IV estimates are quite different and are not of the same pattern as in Table 3. For instance, it appears from the OLS estimates that the more educated are more likely to use a condom and are more likely to have been tested for HIV. The IV estimates suggest the opposite - not only are the standard errors larger, but the coefficients are much smaller and in all but one case they even change signs.[12]

The strong relationship between education and knowledge, but the lack of relationship between education and behavior is a little puzzling, however we believe there could be several explanations for this. Unfortunately, none of them are testable with our data or empirical strategy. Perhaps the questions that determine "HIV knowledge" in this context are not the factors used by people to assess risk regarding HIV. For example, if people believe that using a condom prevents HIV transmission, but do not believe it prevents transmission by much, then more educated people might respond correctly to the question on knowledge of condom use, yet not use a condom during intercourse. Another reason could be one of uniform environmental risk (mortality). If there are risks that are not preventable based on education levels, then while knowledge of specific diseases might be high among the more educated, investments or behaviors might not differ if the horizon is uniformly short. Yet another reason could be that the more educated have sex with people that are "safer" - hence, while knowledge might be higher, the more educated might not need to invest in safer practices since the transmission rates among the educated could be quite low (although correlations would suggest that this is not the case). Finally, our instrumental variables strategy examines people who on the margin, attended secondary school. If there are non linearities in how education

[12] Appendix Table 3 implements these regressions omitting ages 14 and 15 from the regression. While this raises the first stage F-statistic due to a better fit, it does not change the basic import of the results in the second stage.

affects behavior and HIV status, we will not be able to capture that. Despite the instrumental variables strategy, we recognize that the relationship we examine could be driven by various underlying mechanisms. We believe these are very important extensions to the research question at hand, but it is beyond the scope of the current paper to address these.

6 Conclusion

This paper examined the relationship between education and knowledge about health and health behaviors in the salient setting of HIV in Africa. By exploiting an education reform that affected enrollment into secondary school in 1980 in Zimbabwe, we are able to provide some of the first causal estimates on whether education matters for knowledge about health and whether education matters for health behavior with regards to HIV.

Our results reflect that the more educated are more likely to know complex and nuanced information about HIV. In particular, having had some secondary school has a large impact on having comprehensive knowledge about HIV. Moreover, the mechanism of greater knowledge appears to be via more frequent use of media outlets where information on HIV might be obtained. However, the more educated do not seem to undertake more preventive actions like using condoms or getting tested for HIV compared to less educated individuals. There also does not appear to be a difference in being tested HIV positive between people who went to secondary school and those who did not.

There are two broad implications that stem from this paper. First, the more educated while having more knowledge about HIV, do not appear to behave very differently or have lower HIV positive rates. It is likely then that information campaigns should focus on detailed information that even the educated might not possess. Second, given this discrepancy between knowledge and behavior, we think it is important to understand risk factors that people consider while making decisions about sexual practices. For example, simply knowing that a condom reduces the chances of contracting HIV is not enough to induce behavior if the perceived risks of not using a condom are very low. Given the high prevalence rates, this is unlikely, but more work needs to be done on assessing the risk factors that enter the decision making process for healthy behavior regarding HIV.

References

AGÜERO, J., AND M. RAMACHANDRAN (2011): "The Intergenerational Effects of Increasing Parental Schooling: Evidence from Zimbabwe," *Working Paper*.

ATKINSON, N. (1972): *Teaching Rhodesians: A History of Educational Policy in Rhodesia.* Longman, London.

CUTLER, D., AND A. LLERAS-MUNEY (2008): "Education and health: evaluating theories and evidence," *Making Americans Healthier: Social and Economic Policy as Health Policy*, pp. 29–60.

DE WALQUE, D. (2007a): "Does education affect smoking behaviors?:: Evidence using the Vietnam draft as an instrument for college education," *Journal of Health Economics*, 26(5), 877–895.

——— (2007b): "How does the impact of an HIV/AIDS information campaign vary with educational attainment? Evidence from rural Uganda," *Journal of Development Economics*, 84(2), 686–714.

——— (2009): "Does Education Affect HIV Status? Evidence from five African Countries," *World Bank Economic Review*.

DORSEY, B. J. (1989): "Educational Development and Reform in Zimbabwe.," *Comparative Education Review*, 33(1), 40–58.

DUPAS, P. (2010): "Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya," *American Economic Journal: Applied Economics*.

EDWARDS, G. (1995): "Rapid educational expansion and primary school efficiency in Zimbabwe, 1983-9," *Development Southern Africa*, 12(1), 87–95.

EDWARDS, G., AND C. TISDELL (1990): "Post-independence Trends in Education in Zimbabwe," *South African Journal of Economics*, 58(4), 298–307.

KUMARANAYAKE, L., S. LAKE, P. MUJINJA, C. HONGORO, AND R. MPEMBENI (2000): "How do countries regulate the health sector? Evidence from Tanzania and Zimbabwe," *Health policy and planning*, 15(4), 357.

LANGE, F. (2011): "Education and Allocative Efficiency. Evidence from Breast Cancer Screening," *Journal of Health Economics*.

LLERAS-MUNEY, A. (2005): "The relationship between education and adult mortality in the United States," *Review of Economic Studies*, 72(1), 189–221.

NHUNDU, T. J. (1992): "A Decade of Educational Expansion in Zimbabwe: Causes, Consequences, and Policy Contradictions.," *Journal of Negro Education*, 61(1), 78–98.

OCALLAGHAN, M., AND R. AUSTIN (1977): *Southern Rhodesia: The Effects of a Conquest Society on Education, Culture and Information.* Syndenhams Printers, Dorset.

PALMER, R. (1990): "Land reform in Zimbabwe, 1980–1990," *African Affairs*, 89(355), 163.

PITTS, M., AND H. JACKSON (1993): "Press coverage of AIDS in Zimbabwe: A five-year plan," *AIDS Care*, 5(2), 223–30.

RIDDELL, A. R., AND L. M. NYAGURA (1991): "What causes differences in achievement in Zimbabwe's secondary schools?," Policy Research Working Paper Series 705, The World Bank.

RIDDELL, R. (1980): *Education for employment*. Mambo Press, Gweru, Zimbabwe.

THOMAS, D., AND J. MALUCCIO (1996): "Fertility, contraceptive choice, and public policy in Zimbabwe," *The World Bank Economic Review*, 10(1), 189.

THORNTON, R. (2008): "The demand for and impact of learning HIV status: Evidence from a field experiment," *American Economic Review*, 98(5), 1829–1863.

VAN DER KLAAUW, W. (2002): "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach," *International Economic Review*, 43(4), 1249–1287.

Table 1: Summary Characteristics

Sample construction		
Overall Sample size	24598	
Observations 6-22 years old in 1980	8512	
Drop observations with greater than 16 years education	8432	
Observations 9-19 years old in 1980	5068	
In sample 6-23 years old in 1980:		
Variable	*Mean*	*Std Dev*
Male	0.65	0.47
Years of Education	7.7	3.54
Secondary school	0.53	0.49
Urban residence	0.33	0.47
Married	0.94	0.22

Table 2: First stage - did the program affect completed education?

	Ages 6-22 in 1980		Ages 9-19 in 1980	
	Years of education completed (1)	Attended secondary school? (2)	Years of education completed (1)	Attended secondary school? (2)
Cutoff dummy (1 if age<=14 in 1980, 0 otherwise)	1.252***	0.172***	0.784**	0.105**
	[0.312]	[0.0405]	[0.317]	[0.0426]
(Age in 1980-14)*Cutoff dummy	-0.0395*	-0.00335	-0.0823*	-0.0144**
	[0.0201]	[0.00376]	[0.0395]	[0.00556]
(Age in 1980-14)*(1-Cutoff dummy)	-0.215***	-0.0326***	-0.346***	-0.0481***
	[0.0576]	[0.00617]	[0.0888]	[0.0124]
Constant	7.433***	0.454***	7.759***	0.448***
	[0.325]	[0.0410]	[0.358]	[0.0557]
Observations	8,432	8,432	5,068	5,068
R-squared	0.276	0.268	0.253	0.244

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Notes: Sample consists of 1999 and 2005 rounds of the Zimbabwe DHS.

Regressions also include region and year of survey fixed effects, and dummies for sex and type of residence (urban)

Table 3: Education and Knowledge about HIV

	1	2	3	4	5	6	7	8
LHS variables are 1 if the answer is correct, 0 if wrong. Bandwidth of 8 years around age 14 considered.	Heard of HIV?		Comprehensive knowledge of HIV		Can a healthy person have HIV?		Condom reduces chances of HIV?	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Attended secondary school (1,0)	0.0206***	0.0625*	0.171***	0.381***	0.0858***	0.0407	0.106***	0.180*
	[0.00353]	[0.0315]	[0.00990]	[0.120]	[0.00798]	[0.0806]	[0.00929]	[0.0881]
(Age in 1980-14)*Cutoff dummy	0.000919**	0.00170**	-0.000126	0.00322	0.00293***	0.00215	-0.00359***	-0.00209
	[0.000319]	[0.000635]	[0.00152]	[0.00255]	[0.000778]	[0.00163]	[0.00103]	[0.00167]
(Age in 1980-14)*(1-Cutoff dummy)	0.000710*	0.00289	-0.00176	0.0101	-0.00255**	-0.00547	-0.00328*	0.00115
	[0.000361]	[0.00169]	[0.00318]	[0.00820]	[0.00114]	[0.00492]	[0.00172]	[0.00563]
Constant	0.966***	0.948***	0.315***	0.227***	0.773***	0.804***	0.580***	0.562***
	[0.00626]	[0.0201]	[0.0275]	[0.0593]	[0.0151]	[0.0528]	[0.0261]	[0.0572]
Mean of dependent variable	0.98		0.46		0.88		0.76	
Observations	8,432	8,432	8,432	8,432	8,239	8,239	8,427	8,427
R-squared	0.019		0.053	0.015	0.055	0.048	0.050	0.043

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

	9	10	11	12	13	14	15	16
LHS variables are 1 if the answer is correct, 0 if wrong	HIV pass on by mosquito bites?		HIV passed on via breast feeding?		HIV passed on during delivery?		HIV is limited by having sex with only 1 partner	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Attended secondary school (1,0)	0.0850***	0.145*	0.0404***	0.300**	0.0614***	0.0664	0.154***	0.190
	[0.00905]	[0.0696]	[0.0103]	[0.124]	[0.0104]	[0.0470]	[0.0108]	[0.123]
(Age in 1980-14)*Cutoff dummy	0.00309***	0.00398***	-0.00141	0.00346	0.00336***	0.00260	0.00146	0.00118
	[0.000668]	[0.00135]	[0.00144]	[0.00336]	[0.000693]	[0.00151]	[0.00149]	[0.00272]
(Age in 1980-14)*(1-Cutoff dummy)	-0.00307**	0.000883	-0.00189	0.0123*	-0.00338**	-0.00307	0.00411*	0.00745
	[0.00112]	[0.00392]	[0.00214]	[0.00688]	[0.00148]	[0.00292]	[0.00223]	[0.00799]
Constant	0.982***	0.966***	0.454***	0.312***	0.216***	0.207***	0.423***	0.409***
	[0.0234]	[0.0406]	[0.0206]	[0.0676]	[0.0150]	[0.0336]	[0.0224]	[0.0762]
Mean of dependent variable	0.85		0.68		0.58		0.56	
Observations	7,931	7,931	7,870	7,870	7,890	7,890	8,423	8,423
R-squared	0.082	0.078	0.264	0.206	0.561	0.558	0.077	0.077

Robust standard errors in brackets, clustered at age level
*** p<0.01, ** p<0.05, * p<0.1

Notes: Sample consists of 1999 and 2005 rounds of the Zimbabwe DHS. Triangular weights used in all IV specifications.

First stage F-stat is around 21. Regressions also include region and year of survey fixed effects, and dummies for sex and type of residence (urban)

Table 4: Education and Media Use

	1	2	3	4	5	6
1 if ever engage in these activities, 0 otherwise. Bandwidth of 8 years around age 14 considered.	Read Newspapers		Listen to Radio		Watch Television	
	OLS	IV	OLS	IV	OLS	IV
Attended secondary school (1,0)	0.317***	0.676***	0.157***	0.176	0.179***	0.254**
	[0.0100]	[0.158]	[0.0107]	[0.150]	[0.0123]	[0.117]
(Age in 1980-14)*Cutoff dummy	0.00291	0.0103*	-0.000487	0.000541	0.00540***	0.00651***
	[0.00194]	[0.00503]	[0.00166]	[0.00227]	[0.00118]	[0.00211]
(Age in 1980-14)*(1-Cutoff dummy)	-0.00319	0.0147*	0.00125	0.00321	0.000695	0.00629
	[0.00304]	[0.00757]	[0.00149]	[0.00897]	[0.00148]	[0.00757]
Constant	0.362***	0.184**	0.556***	0.557***	0.545***	0.508***
	[0.0278]	[0.0690]	[0.0237]	[0.0947]	[0.0213]	[0.0750]
Mean of dependent variable	0.46		0.72		0.39	
Observations	8,421	8,421	8,416	8,416	8,424	8,424
R-squared	0.301	0.206	0.177	0.177	0.366	0.361

Robust standard errors in brackets, clustered at age level
*** p<0.01, ** p<0.05, * p<0.1

Notes: First stage F-stat is around 20. Triangular weights used in all IV specifications.
Regressions also include region and year of survey fixed effects, and dummies for sex and type of residence (urban)

Table 5: Education and HIV Preventive Behaviors

	1	2	3	4	5	6	7	8
	Total Number of Lifetime Partners		Used condom during last intercourse? (1=yes, 0=no)		Ever been tested for HIV? (1=yes, 0=no)		HIV Status (1=positive, 0=negative)	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Attended secondary school (1,0)	0.103*	-0.0926	0.0287***	-0.0229	0.102***	-0.0615	0.0176	0.257
	[0.0532]	[1.189]	[0.00577]	[0.0758]	[0.00884]	[0.0986]	[0.0116]	[0.237]
(Age in 1980-14)*Cutoff dummy	0.0258	0.0547	-0.00454***	-0.00500**	0.000423	-0.00343	0.00319	0.00867
	[0.0254]	[0.0361]	[0.00107]	[0.00187]	[0.00126]	[0.00248]	[0.00224]	[0.00536]
(Age in 1980-14)*(1-Cutoff dummy)	0.0289*	0.00644	-0.00292*	-0.00547	-0.00286	-0.0112*	-0.0138***	0.00115
	[0.0158]	[0.0566]	[0.00140]	[0.00423]	[0.00178]	[0.00563]	[0.00299]	[0.0135]
Constant	1.606***	1.785**	0.0942***	0.0981**	0.225***	0.312***	0.378***	0.182
	[0.155]	[0.750]	[0.0235]	[0.0452]	[0.0262]	[0.0636]	[0.0447]	[0.168]
First stage F-stat		7.04		17.62		19.9		5.3
Mean of dependent variable	2.73		0.11		0.19		0.3	
Observations	4,730	4,730	7,297	7,297	8,333	8,333	4,010	4,010
R-squared	0.276	0.285	0.016	0.009	0.051	0.017	0.011	

Robust standard errors in brackets, clustered at age level
*** p<0.01, ** p<0.05, * p<0.1

Notes: Regressions 3-6 use data from DHS rounds 1999 and 2005-06, while regressions 1,2,7 & 8 use data only from 2005-06 round. Triangular weights used
Regressions also include region and year of survey fixed effects, and dummies for sex and type of residence (urban)

Appendix Table 1: Education and Knowledge about HIV (bandwidth ages 9-19)

	1	2	3	4	5	6	7	8
LHS variables are 1 if the answer is correct, 0 if wrong	Heard of HIV?		Comprehensive knowledge of HIV		Can a healthy person have HIV?		Condom reduces chances of HIV?	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Attended secondary school (1,0)	0.0186***	-0.00676	0.164***	1.382	0.0842***	0.0333	0.102***	0.366
	[0.00388]	[0.176]	[0.0160]	[1.252]	[0.0102]	[0.560]	[0.0135]	[0.733]
(Age in 1980-14)*Cutoff dummy	0.00103	0.000630	-0.00130	0.0327	0.00227	0.000199	-0.00146	-0.000984
	[0.00110]	[0.00494]	[0.00480]	[0.0362]	[0.00295]	[0.0158]	[0.00405]	[0.0210]
(Age in 1980-14)*(1-Cutoff dummy)	-0.000132	-0.00421	-0.000799	0.0974	-0.00341	-0.00435	-0.00285	0.0190
	[0.00120]	[0.0147]	[0.00526]	[0.103]	[0.00352]	[0.0466]	[0.00460]	[0.0605]
Constant	0.973***	0.989***	0.350***	-0.268	0.775***	0.813***	0.601***	0.485
	[0.00649]	[0.0882]	[0.0380]	[0.633]	[0.0226]	[0.290]	[0.0313]	[0.367]
Observations	5,068	4,215	5,068	4,215	4,954	4,121	5,065	4,212
R-squared	0.017	0.011	0.048		0.051	0.041	0.046	

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

	9	10	11	12	13	14	15	16
LHS variables are 1 if the answer is correct, 0 if wrong	HIV pass on by mosquito bites?		HIV passed on via breast milk?		HIV passed on during delivery?		HIV is limited by having sex with only 1 partner	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Attended secondary school (1,0)	0.0889***	0.369	0.0481***	0.967	0.0626***	0.0819	0.167***	0.216
	[0.0112]	[0.609]	[0.0137]	[1.107]	[0.0109]	[0.546]	[0.0155]	[0.795]
(Age in 1980-14)*Cutoff dummy	0.00151	0.0129	-0.000579	0.0240	-5.32e-05	-0.00144	-0.00391	-0.00134
	[0.00332]	[0.0185]	[0.00396]	[0.0298]	[0.00322]	[0.0159]	[0.00466]	[0.0231]
(Age in 1980-14)*(1-Cutoff dummy)	-0.00178	0.0185	-0.00248	0.0681	-0.00155	0.00214	0.0122**	0.0108
	[0.00379]	[0.0485]	[0.00446]	[0.0900]	[0.00347]	[0.0449]	[0.00511]	[0.0653]
Constant	0.998***	0.868***	0.449***	-0.0246	0.200***	0.196	0.420***	0.407
	[0.0277]	[0.318]	[0.0303]	[0.581]	[0.0245]	[0.293]	[0.0375]	[0.402]
Observations	4,761	3,965	4,738	3,933	4,752	3,950	5,061	4,209
R-squared	0.084		0.268		0.573	0.537	0.080	0.076

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Notes: Sample consists of 1999 and 2005 rounds of the Zimbabwe DHS. Triangular weights used in all IV specifications.
First stage F-stat is around 13. Regressions also include region and year of survey fixed effects, and dummies for sex and type of residence (urban). Due to small number of age groups, standard errors are not clustered at the age level here.

Appendix Table 2: Education and Knowledge about HIV (dropping ages 14 and 15 in 1980)

	1	2	3	4	5	6	7	8
LHS variables are 1 if the answer is correct, 0 if wrong	Heard of HIV?		Comprehensive knowledge of HIV		Can a healthy person have HIV?		Condom reduces chances of HIV?	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Attended secondary school (1,0)	0.0210***	0.0796**	0.173***	0.226**	0.0884***	-0.0156	0.111***	0.278**
	[0.00389]	[0.0313]	[0.0107]	[0.0868]	[0.00865]	[0.0412]	[0.00878]	[0.0932]
(Age in 1980-14)*Cutoff dummy	0.000585	0.00123***	-0.000101	0.000252	0.00384***	0.00367***	-0.00332**	0.000490
	[0.000421]	[0.000347]	[0.00177]	[0.00131]	[0.000809]	[0.00111]	[0.00144]	[0.00220]
(Age in 1980-14)*(1-Cutoff dummy)	0.00108**	0.00490**	-0.00156	0.00124	-0.00321**	-0.0118***	-0.00329	0.00724
	[0.000464]	[0.00213]	[0.00338]	[0.00696]	[0.00133]	[0.00349]	[0.00199]	[0.00748]
Constant	0.963***	0.936***	0.310***	0.311***	0.777***	0.849***	0.574***	0.515***
	[0.00745]	[0.0230]	[0.0312]	[0.0537]	[0.0157]	[0.0271]	[0.0282]	[0.0637]
Mean of dependent variable	0.98		0.46		0.88		0.76	
Observations	7,664	6,515	7,664	6,515	7,485	6,365	7,659	6,511
R-squared	0.019		0.054	0.048	0.059	0.038	0.051	0.021

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

	9	10	11	12	13	14	15	16
LHS variables are 1 if the answer is correct, 0 if wrong	HIV pass on by mosquito bites?		HIV passed on via breast milk?		HIV passed on during delivery?		HIV is limited by having sex with only 1 partner	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Attended secondary school (1,0)	0.0883***	0.122	0.0435***	0.264**	0.0570***	0.0665	0.152***	0.221*
	[0.00917]	[0.0854]	[0.0112]	[0.119]	[0.0109]	[0.0397]	[0.0116]	[0.110]
(Age in 1980-14)*Cutoff dummy	0.00318***	0.00322	-0.00242	0.00102	0.00418***	0.00392***	0.000892	-0.00130
	[0.000987]	[0.00189]	[0.00156]	[0.00169]	[0.000686]	[0.00103]	[0.00181]	[0.00296]
(Age in 1980-14)*(1-Cutoff dummy)	-0.00301**	0.000125	-0.000609	0.0130	-0.00449**	-0.00472	0.00456*	0.0128
	[0.00135]	[0.00525]	[0.00218]	[0.00936]	[0.00157]	[0.00334]	[0.00250]	[0.00941]
Constant	0.978***	0.979***	0.454***	0.338***	0.223***	0.208***	0.417***	0.375***
	[0.0249]	[0.0503]	[0.0239]	[0.0690]	[0.0163]	[0.0340]	[0.0242]	[0.0811]
Mean of dependent variable	0.85		0.68		0.58		0.56	
Observations	7,210	6,129	7,153	6,077	7,172	6,090	7,655	6,506
R-squared	0.082	0.084	0.260	0.216	0.565	0.567	0.076	0.073

Robust standard errors in brackets, clustered at age level
*** p<0.01, ** p<0.05, * p<0.1

Notes: Sample consists of 1999 and 2005 rounds of the Zimbabwe DHS. Triangular weights used in all IV specifications.

First stage F-stat is around 93. Regressions also include region and year of survey fixed effects, and dummies for sex and type of residence (urban)

Appendix Table 3: Education and HIV Preventive Behaviors (dropping ages 14 and 15)

	1	2	3	4	5	6	7	8
	Total Number of Lifetime Partners		Used condom during last intercourse? (1=yes, 0=no)		Ever been tested for HIV? (1=yes, 0=no)		HIV Status (1=positive, 0=negative)	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Attended secondary school (1,0)	0.0624	-0.951	0.0316***	-0.0919	0.105***	-0.0345	0.0213	0.130
	[0.0517]	[1.055]	[0.00532]	[0.0699]	[0.00970]	[0.0849]	[0.0126]	[0.190]
(Age in 1980-14)*Cutoff dummy	0.00999	0.0447	-0.00480***	-0.00674**	-0.000266	-0.00671*	0.00292	0.00734
	[0.0283]	[0.0409]	[0.00135]	[0.00240]	[0.00168]	[0.00330]	[0.00317]	[0.00691]
(Age in 1980-14)*(1-Cutoff dummy)	0.0422**	-0.0408	-0.00241	-0.00901*	-0.00205	-0.00684	-0.0132***	-0.00495
	[0.0190]	[0.0512]	[0.00170]	[0.00439]	[0.00221]	[0.00581]	[0.00370]	[0.0104]
Constant	1.583***	2.460***	0.0954***	0.134**	0.211***	0.273***	0.413***	0.316**
	[0.178]	[0.717]	[0.0261]	[0.0462]	[0.0272]	[0.0413]	[0.0379]	[0.104]
First stage F-stat		18		32		42		16.8
Mean of dependent variable	2.73		0.11		0.19		0.3	
Observations	4,283	3,641	6,655	5,656	7,572	6,434	3,642	3,103
R-squared	0.270	0.248	0.017		0.052	0.026	0.013	0.002

Robust standard errors in brackets, clustered at age level
*** p<0.01, ** p<0.05, * p<0.1

Notes: Regressions 3-6 use data from DHS rounds 1999 and 2005-06, while regressions 1,2,7 & 8 use data only from 2005-06 round. Triangular weights used in
Regressions also include region and year of survey fixed effects, and dummies for sex and type of residence (urban)