The Education "Vaccine" Against HIV: Evidence from Zambia's Basic Education Sub-Sector Investment Program

An honors thesis for the Department of Economics
Alix Slosberg

Tufts University, 2013

Table of Contents

Abstract

The HIV and AIDS pandemic remains one of the most pressing health and developmental issues of the twenty-first century. As is the case with other public health concerns, policymakers and academics have looked to the role that education can serve in preventing the spread of the disease. However, despite education's usual role as a preventative factor, or "social vaccine," against contagion, within the discourse on the HIV and AIDS pandemic education's association with infection is contested. Additionally, in many analyses done relating education to HIV status, education is endogenous to the model, leading to biased results. In this study I provide a background of the established theories that examine education's relation with HIV status. Filling the void created by endogeneity issues that restrict past studies, I use an instrumental variable regression to model the relationship between education and HIV status in Zambia. Using the 2007 Zambian Demographic and Health Survey data, I utilize the exogenous shock to educational attainment created by the Zambian Basic Education Sub-Sector Investment Program's free primary education policy (2002), to instrument for education in my model. I find that within my total sample, education has a significant and negative effect on HIV status. However, when the instrumental variable regression is run by gender, it is only the males who display a significant and negative relationship, while the correlation between education and HIV is insignificant for females. Based on these findings, I suggest that HIV prevention policy may be best directed towards investing in universal primary education, initiatives that target women and girls, information campaigns and sexual health curricula in schools.

1. Introduction

"The HIV/AIDS epidemic kills millions of people, drains monetary resources and state budgets, cripples the economic and social well-being of countries, deters investment, depresses economic growth, increases poverty and state dependency, and erodes institutions of governance."

LaMontagne and Stockemer 2010

Every day approximately 6,800 people contract HIV and 5,700 people die of AIDS (LaMontagne and Stockemer 2010). As the fourth leading cause of death in the world Human Immunodeficiency Virus and Acquired Immune Deficiency Syndrome (HIV and AIDS) has become a health disaster. Approximately forty-two million individuals are estimated to be living with the virus, and in severely affected countries, life expectancies have plummeted (McIntosh and Thomas 2004; Baker et al. 2009). Not only is HIV and AIDS a pressing health issue, it is also severely impedes development for many nations, with the greatest impacts in sub-Saharan Africa where 76 percent of all AIDS deaths occur (Baker et al. 2009).

As is the case with the multitude of other population health concerns, those who study the HIV and AIDS pandemic look to the role that a formal education can serve in the prevention of the disease. The widely accepted argument regarding education and health is that an increase in schooling, through both heightened awareness and cognitive ability, begets a decrease in risky behavior and an increase in healthier lifestyles - causing education to be labeled a "social vaccine" (Baker et al. 2009). However, despite the public confidence in education's role as a "social vaccine" because of its negative correlation with health risks, within the literature on education and HIV and AIDS the correlation has become a point of contention (Baker et al. 2009).

A finding offered by Cogneau and Grimm (2006), Fortson (2008) and others states that education has a positive effect on HIV and AIDS prevalence rates because of the increased mobility and access to sexual partners that are indirect results of education. A second theory coincides with the general public health view of education and states that education's effect on HIV and AIDS is preventative due to the increased exposure to health information and increase in wealth that are correlates of schooling (De Walque 2009; LaMontagne and Stockemer 2010). Yet another theory is that both the positive and negative effects of education work in opposite directions, and therefore many populations exhibit a null association between education and HIV (De Walque and Kline 2010). A last explanation is that the role of education in the HIV and AIDS pandemic has been changing over time. During the initial stages of the pandemic it was the more educated and wealthy individuals who had the highest risk of contracting the disease, while the burden of HIV has now shifted to the least educated (Michelo et al. 2006; Hargreaves et al. 2008; Baker et al. 2009; De Walque et al. 2005).

While there is evidence of each hypothesis proving true in different contexts, in many studies it is often the case that education is endogenous to the econometric model used, meaning that its correlation with HIV is determined by other population characteristics also included in the model, leaving results subject to biases. It is crucial that we work towards developing a sound unbiased model identifying the relationship between education and HIV because "the identification of the socioeconomic characteristics of low and high risk groups seems indispensable to set up adequate AIDS prevention and therapy policies in developing countries" (Cogneau and Grimm 2006).

Using Zambia as a case study, I attempt to provide a solution to the endogeneity issues that restrict previous studies by using an instrumental variable regression. I use the shock to education caused by the Zambian Basic Education Sub-Sector Investment Program (BESSIP) – a policy that declared primary education free in Zambia in 2002 – as an instrument to isolate an exogenous variation in education. Regressing HIV data on this exogenous variation in education provides more accurate information regarding the relationship between education and HIV than the studies done previously that suffer from endogeneity biases. I find that, in Zambia, education is significantly negatively related to HIV. When the model is executed by gender, I find that the relationship is only significant for males.

In order to give the reader a holistic view of the discourse on the role of education in HIV and AIDS prevention, I first provide a literature review detailing past studies on each of the four prominent theories detailed above. However, I posit that, in reality, these four theories are not mutually exclusive but are all driven by common factors that increase with education – risky behaviors (such as an increase in sexual partners and greater mobility) and protective behaviors (such as increased financial stability for women and increased condom use). These behaviors work in opposite directions in forming the relationship between education and HIV, and the direction of the correlation is ultimately determined by an underlying country and time specific factor – the presence or the absence of HIV prevention information. Within this framework, I give a background of HIV in Zambia, prevention efforts in the country, as well as the Zambian Basic Education Sub-Sector Investment Program (BESSIP). Using the 2007 Zambian Demographic and Health Survey (ZDHS) data, I test for the role of education in preventing, or not preventing, HIV when a policy that provides an incentive for educational attainment is implemented. In order to assess the robustness of my model, I run the same regression using a smaller sample size. To gain a deeper understanding of the relationship between education and

HIV, I conduct gender specific regressions. Finally, I conclude with recommendations for further research and give policy proposals based on my findings.

2. Literature Review

In this section I provide a critical review of the existing scholarship that examines the relationship between education and HIV and AIDS. For the purposes of this paper, I only focus on studies done on the African continent, particularly in sub-Saharan Africa, because the circumstances of these populations are most relevant to the country I am studying directly – Zambia. I separately examine four theories – education as a positive, negative, null and changing correlate of HIV. In each section, I explain how both risky and preventative behaviors that arise with increased schooling are present and how the existence or absence of HIV prevention information is the determining factor behind the direction of education's correlation with HIV. I conclude by explaining the void that exists in the literature because of endgoeneity issues that bias many of the studies and how my analysis will bridge the gap.

In order to understand the role of information in each theory, it is necessary to provide a summary of the history of HIV prevention information and awareness efforts in sub-Saharan Africa. Foremost, HIV and AIDS spread rapidly throughout sub-Saharan Africa from 1970-1980 (Baker et al. 2009). At this time, public health information regarding HIV and AIDS was unavailable, leaving people unable to protect themselves from the disease. Exacerbating this issue, HIV was advertised as solely an issue faced by homosexual males in the United States and Europe. During this period of false branding, HIV was quickly spreading in sub-Saharan Africa through heterosexual means (Baker et al. 2009). Additionally, because of the extreme stigmatization of HIV, governments throughout sub-Saharan Africa refused to acknowledge that the disease existed amongst their citizens. Countries such as South Africa, which now suffers from one of the highest HIV and AIDS prevalence rates in the world, failed to fight and prevent HIV until the epidemic was heavily affecting their nation (Baker et al. 2009). Prevention information was finally disseminated to the public in the 1990's, but the factors above, combined with social norms that encouraged transactional sex and having multiple sex partners as a reflection of high social status, had already led to the rapid spread of HIV (Baker et al. 2009).

2.1 Risk Factor

Diverging from the common discourse surrounding the preventative role of education in disease control, a theory posited by many scholars asserts that education is a "risk factor" for HIV contraction, meaning that those who are more educated are more likely to have HIV (Cogneau and Grimm 2006; De Walque and Kline 2010; Baker et al. 2009). Scholars who have studied this phenomenon give reasons why the above counterintuitive relationship has proven true under many circumstances. Baker et al. (2009) and De Walque and Kline (2010) explain how increased educational attainment can increase one's vulnerability to contracting HIV because of the monetary and social resources that increase with education, particularly those resources that translate into risk. First, those who are more educated have a more expansive social network, leading them to have a similarly large sexual network. This greater exposure to sexual partners makes them increasingly susceptible to contracting HIV (De Walque and Kline 2010). Furthermore, greater education and greater socioeconomic status are often synonymous. With more wealth comes more disposable income that, for men, can mean more opportunities to engage in transactional sexual relationships. While more wealth may delay marriage for women, this delay may also mean they have more sexual partners before marriage (De Walque and Kline 2010). Lastly, the upward mobility and leisure time that increase with greater education and wealth can heighten one's chances for contracting HIV (Baker et al. 2009).

Indeed, an empirical study by Cogneau and Grimm (2006) finds that having more education increases the risk of HIV infection. Using 1994 Demographic and Health Survey data from Côte d'Ivoire, they construct a "susceptible-infected" model using information collected on attitudes towards and knowledge about HIV and AIDS as a proxy for becoming infected. The model relates the probability of contracting HIV to one's number of partners, their frequency of unprotected sex with each partner, and the average infection rate of partners. From there, they compute the risk factors for men and women by educational attainment. In their evaluation of the risks of education, Cogneau and Grimm (2006) find that, for both men and women, the risk of HIV infection is greater with greater educational attainment and that the differential increases with age. These results make sense in the context of their additional findings that establish that it is the more educated who are likely to have multiple sexual partners and exist in a more "risky environment." However, Cogneau and Grimm (2006) also find that the positive effects of education on HIV were slightly offset by a probability of condom use that also rises with schooling. While their study is imperfect because they do not have actual data on individuals' HIV status and are inferring risks based solely on reported behaviors, the conclusions highlight that both risky behaviors and preventative behaviors rise with education. However, in their sample the risky behaviors outweigh the protective, resulting in a positive correlation between education and HIV. I attribute this to the early date at which the data were collected (1994) and the lack of information available at the time, which did not allow the more educated to protect themselves from HIV.

2.2 Social Vaccine

When the preventative behaviors that increase with education, such as condom use found by Cogneau and Grimm (2006), are stronger than the risky behaviors, education becomes negatively correlated with HIV status, causing education to be labeled a "social vaccine" (Baker et al. 2009). A quote by Baker et al. (2009) describes this "social vaccine" view of education that holds in public health discourse: “perhaps nothing illustrates the positive impact formal education can have on the health of whole populations as its association with the much-noted demographic transition of modern society, consisting of reduced mortality and fertility, leading to longer life spans.” Indeed, health and education can be said to mutually enhance the other as health begets more education in a population, which then can lead to greater healthy practices. They both become significant features of human capital (Baker et al. 2009).

In the context of HIV and AIDS, general education, (not just HIV-specific education) can positively affect many behaviors connected to HIV infection (De Walque and Kline 2010). Primarily, those who are more educated may have greater exposure to prevention information and may better comprehend the repercussions of their actions with regards to HIV (Hargreaves et al. 2008).

It is often argued that education can have the greatest effects on the female population. Particularly for women, those who are better educated are more likely to marry later in life, have fewer children and earn higher incomes than their less educated peers – all factors which reduce vulnerability to HIV infection (Pettifor et al. 2008). Indeed, in De Walque’s (2009) study of five African countries that I will summarize below, he finds that education is significantly correlated with a later sexual debut for women in each country. Furthermore, more educated women may also feel that they have greater control over their behavior than their less education peers, resulting in situations where educated women feel more confident negotiating safer sex (De Walque and Kline 2010).

However, in many regions where HIV prevalence is high, such as sub-Saharan Africa, people have not enjoyed the same increases in mass education as more economically developed areas. Despite this disparity, Baker et al. (2009) explain that provision of basic education has grown amply enough in these areas to have possibly cultivated the conditions for education to serve as a social vaccine.

In his empirical study, De Walque (2009) identifies this social vaccine relationship between education and HIV by testing the prominent socioeconomic determinants of HIV and AIDS. Using 2003-2004 DHS data from five African countries – Burkina Faso, Cameroon, Ghana, Kenya and Tanzania, De Walque (2009) conducts both pooled and country-specific cross sectional multivariate regressions holding age, location, marital status, wealth, education, religion, and ethnicity constant. His results are interesting because initially education does not appear to be significantly associated with HIV status; however, when the regressions are restricted to urban areas only, education is negatively correlated with infection and statistically significant. De Walque (2009) concludes that in urban areas education is not positively associated with HIV infection, as previously thought, but that education is a reliable predictor of preventative behaviors. However, I argue that education is a reliable predicator of both risky and preventative behaviors, but in De Walque's (2009) study the preventative behaviors were stronger. Connecting this to the presence of information, De Walque (2009) hypothesizes that the reason for the differences in results between urban and rural areas is because "…the negative relationship between education and HIV infection takes time to develop and is found earlier in cities, where information spreads faster and HIV prevalence is generally higher." The interaction between information and education allowed for the urban individuals in De Walque's (2009) study to adopt preventative behaviors against HIV.

Finally, De Walque's (2009) analysis, particularly with regard to education, suffers from endogeneity bias. He explicitly expresses these concerns in his paper – "most of the individual characteristics used as regressors…cannot be defined as completely exogenous variables…[and]

coefficients should therefore be interpreted with caution….” Because of these issues, causality between education and HIV cannot be proven in De Walque’s (2009) study.

2.3 No Effect

While the study by Cogneau and Grimm (2006) detailed above finds a positive relationship between education and risk of HIV, they also state that the positive effects of education on HIV were slightly counteracted by a probability of condom use that also rises with education. Similarly, other scholars, such a De Walque and Kline (2010), find that this increased use of condoms and other protective practices completely offset the effects of the risky behaviors that may rise with education, leading them to conclude that education has no bearing on one’s chances of contracting HIV. Using Demographic and Health Surveys from six African countries – Ethiopia (2005), Guinea (2005), Ivory Coast (1998/1999), Malawi (2004), Rwanda (2005) and Zimbabwe (2005/2006) - De Walque and Kline (2010) regress HIV status on primary and secondary education, holding marital status and wealth constant, to find no proof that there is an identifiable relationship between education and HIV. However, in further exploration of this relationship, De Walque and Kline (2010) run additional tests and find that both condom use (preventative behavior) and extramarital sex (risky behavior) increase with education. They attribute the observed null relationship between HIV and education to the fact that the above behaviors work in opposite direction, eliminating any positive or negative effects that education may have on HIV status. Concluding their study, De Walque and Kline (2010) emphasize the need for future research to be directed towards randomized experiments in order to uncover any causal effects that schooling may have on HIV prevention or the lack thereof.

Additionally, like De Walque's (2009) study that finds education working as a social vaccine in urban areas, De Walque and Kline's (2010) study may be biased due to the endogeneity of education to their model.

2.4 Changing Association

De Walque's (2009) above hypothesis, which emphasizes that the ultimate relationship between education and HIV takes time to develop, serves as a precursor to the literature that examines the associations between education and HIV over many years, rather than relying on a cross-section as was done in the aforementioned studies. Within this wider time-frame, many scholars have found that in the beginning of the HIV pandemic it was the more educated individuals who were likely to contract HIV, but as the pandemic has progressed, the burden of HIV has switched to those who are least educated (Hargreaves et al. 2008; Baker et al. 2009; De Walque et al. 2005; Michelo et al. 2006; Fyklenes et al. 2001). This understanding of the relationship between education and HIV ameliorates, to a large extent, the discrepancies between the three above theories by combining them together into a larger cohesive story. Because HIV testing information is rare and was not commonly done on a mass-scale in the beginning of the pandemic, supporters of the changing association theory have come to this conclusion via three methods of research: (1) reviewing studies over time and detecting trends in the studies (Hargreaves et al. 2008) (2) turning a cross-sectional dataset into a "time-series" dataset by fragmenting it by age group (Baker et al. 2009) and (3) using sporadic testing information from a single population (De Walque et al. 2005; Michelo et al. 2006).[1]

In Hargreaves et al.'s (2008) "Systematic review exploring time trends in the relation between educational attainment and risk of HIV infection in sub-Saharan Africa," the authors

[1] Baker et al. (2009) explain that because "there are no longitudinal data over the course of the pandemic on multiple waves of individuals becoming sexually active in SSA [sub-Saharan Africa] nations, one must approximate an answer to whether or not the relationship between education and HIV infection is shifting."

produce a review of articles that focus on the association between education and HIV prevalence, with the goal of testing the hypothesis that studies done in the late 1990's and after are more likely to report a negative association between education and HIV, while studies done in the early 1990's and previously are more likely to report a positive association. 36 articles were used in their final study, incorporating data from 72 discrete populations collected from 1987-2003. 11 countries and over 200,000 individuals are represented.

Hargreaves et al.'s (2008) findings uphold the veracity of their changing association hypothesis. Of 32 articles that used cross-sectional data collected before 1996, 15 of the populations were more likely to show a higher risk of HIV infection in the more educated individuals, while only 5 out of 40 populations with data collected post-1996 displayed this phenomenon.[2] However, only one pre-1996 study as opposed to seven from post-1996 reported a lower risk of HIV among the most educated. Second, in populations where data was available for different time periods (13 populations from 5 countries: Malawi, Tanzania, Uganda, Zambia and Zimbabwe) a common trend was observed. In each population from 1996 to the present, changes in an adjusted odds ratio for HIV status that compared groups from high and low educational attainment levels were directed towards no association or towards a smaller risk among the most educated. Lastly, where time series data were available, positive associations between education and HIV were generally replaced by negative associations.

Baker et al. (2009) find these same results via an alternative means of analysis – they create three age-cohorts from a cross-sectional dataset, with each age-cohort reflecting a "period effect" between the time the individuals in the cohort came to sexual maturity and historical points in the HIV pandemic. Using 2004 Demographic and Health Survey data from 11 African countries,

[2] Hargreaves et al. (2008) use 1996 as the cut-off because a previous review by Hargreaves used only studies from pre-1996. Hargreaves added more recent studies to the older review to create his 2008 study. Approximately half the studies included in his 2008 review use data from pre-1996 and half use data from post-1996.

Baker et al.'s (2009) analysis rests on three assumptions – (1) that the oldest cohort (over 34 years of age) became sexually active at a time in the pandemic when accurate information about HIV was scarce and so there should be evidence in this cohort of education as a risk factor (2) that the youngest cohort in the data (15-24 years of age) became sexually active when information about risks was more widely available and so among this cohort there should be evidence of education as a social vaccine and (3) that in the middle-aged cohort (25-34) there should be no relationship between education and HIV infection because in this generation both the risk factor and social vaccine processes are working in opposite directions.[3] After running multivariate regressions for each age-cohort separately, Baker et al.'s (2009) results generally fit their predictions, leading them to the conclusion that within the HIV pandemic, education is finally beginning to take its usual role in health and serve as a preventative measure against infection.

While Baker et al.'s (2009) use of "period effects" by comparing age-cohorts is a creative solution to solving issues resulting from a lack of time-series HIV data, his model suffers not only from endogeneity of education but also from the fact that he does not account for other "period effects" that may influence his results, such as differences in the political environment and policies under which the oldest cohort matured versus those under which the youngest cohort matured.

Differing from the two approaches above, both De Walque et al. (2005) and Michelo et al. (2006) use time-series data to compare age-adjusted odds-ratios of HIV infection by education categories constructed from the beginning of the data versus odds-ratios constructed from the final years of collection. By comparing the ratios, they can assess how the epidemiology of the

[3] Baker et al. (2009) use data from Rwanda, Ethiopia, Guinea, Malawi, Senegal, Burkina Faso, Cameroon, Ghana, Kenya, Tanzania and Lesotho.

HIV epidemic has changed over time. Michelo et al. (2006) rely on data collected from urban and rural communities in Zambia in the years 1995, 1999 and 2003, while De Walque et al. (2005) use data from a rural population in Uganda between the years of 1989/1990 and 1999/2000. I find De Walque et al.'s (2005) and Michelo et al.'s (2006) approach most compelling because they utilize actual data collected over many years from a single population, as opposed to Baker et al.'s (2009) study that relies on the creation of "period effects" from cross-sectional data and Hargreaves et al.'s (2008) study that compares populations from various countries.

De Walque et al. (2005) find that at round one of testing (1989/1990) the age-adjusted odds-ratios show that there is no significant trend between education and risk of HIV infection; however, at round 11 (1999/2000) age-adjusted odds-ratios show that higher schooling is associated with lower HIV prevalence for young females. Michelo et al. (2006) find similar results. In their sample, HIV prevalence declined among those who were higher educated. The most marked decline was seen in the population of urban men with greater than 11 years of education; in this group, HIV prevalence declined from 30.2 percent to 11.7 percent from the years of 1995-2003. Declines were also seen in urban women, and rural women and men, with greater than 11 years of education. However, in less educated groups, prevalence remained stable or increased during the same time period.

As expected, the discussion surrounding this phenomenon has focused on the reasons for which we see this change in the epidemiology of the HIV and AIDS pandemic and its implications. In response to the revelation of the above trends, Michelo et al. (2006) explain their significance: "this is of interest because knowledge, behavior and behavioral change may be linked to educational level in that ability to understand and act on health promotional messages as well as attitudes that can influence one's health are increased." Like Michelo et al. (2006), most hypotheses attribute the change in association to the heightened presence of information that occurred in the 1990's, which allowed those who were more educated to have greater access to and greater cognition of prevention information (Hargreaves et al. 2008; De Walque et al. 2005; Baker et al. 2009; Michelo et al. 2006; Gregson et al. 2001). If this trend is to be recognized, the challenge facing policy-makers is the creation of preventive programs combined with greater access to education that target hard-to-reach populations, particularly the rural poor (Michelo et al. 2006).

To conclude this literature review, it is important to note that there is no unanimous consensus regarding education's role in the HIV pandemic; we must therefore continue to deepen our understanding of how education can serve as a steward towards an AIDS free society (Fortson 2008). However, I argue that the theoretical discrepancies that exist are largely ameliorated by a recognition of the fact that each phenomenon (education's positive, negative, null and changing relationship with HIV) can be explained by identifying the risky and preventative behaviors that increase with education and exploring reasons why one set of behaviors may outweigh the others, which I posit is the presence of information. With this hypothesis in mind, the studies that examine the relationship between education and HIV over time, and conclude that the correlation has changed from positive to negative, are extremely convincing. Nevertheless, gaps remain in the research, namely a study that isolates the exogenous variation in education to identify an unbiased correlation (or lack thereof) between education and HIV status. I later explain how I plan to fill this void.

3. HIV and Public Policy in Zambia

I use Zambia as a case study to test the relationship between education and HIV status. In this section I provide a background on HIV in Zambia, prevention efforts in the country, as well as the Basic Education Sub-Sector Investment Program (BESSIP) that declared Zambian primary education free in 2002. The discontinuity created in educational attainment arising from this policy serves as a central component in my empirical analysis.

3.1 Background

In 2010 Zambia's population was estimated to be thirteen million (UNICEF 2003). The GNI per capita stands at US$1070 per person per year with 60.5% of Zambians living below the poverty line (UNICEF 2003; World Bank 2013). Additionally, Zambia has the fourth highest adult HIV prevalence rate in the world (Kandala et al. 2008). In 2011 this rate was 14% with 226 new infections occurring each day, 25 of which are among children (World Health Organization 2011). Over 40% of the population resides in the capital city of Lusaka and in the towns of the Copperbelt and it is in these urban areas that HIV prevalence in highest (54% of all adults in these areas are living with HIV or AIDS) (Kandala et al. 2008). Women are the most vulnerable group to HIV, with 54% of all Zambian HIV cases estimated to be women (World Health Organization 2005). Amplifying this problem is the fact that common occupations in Zambia, such as seasonal laborers, fishermen/women, and truck drivers, are characterized by spatial mobility. This mobility has spread HIV throughout the migrant population including those with whom the migrants have sexual relations (Kandala et al. 2008).

3.2 Prevention Efforts in Zambia

The history of HIV and AIDS in Zambia dates back to 1984 when the first person to be diagnosed with AIDS was reported (World Health Organization 2005). The government quickly responded and within two years formed the National AIDS Surveillance Committee (NASC) and the National AIDS Prevention and Control Program (NAPCP) to combat the spread of HIV and AIDS. However, despite the government's seemingly aggressive response, officials denied the harsh reality of AIDS through much of the 1990's. AIDS was not declared a national emergency until 2004 under the administration of President Mwanawasa (Avert 2013). Indeed, the turn of the millennium brought a change in priorities and in 2002 the government implemented a comprehensive antiretroviral treatment program in the hopes of containing the spread of the epidemic. Since then, HIV and AIDS cases have declined, but prevalence remains extremely high (Kandala et al. 2008).

Awareness raising, particularly amongst youth, has become a pillar of the government's response to the HIV epidemic (Underwood et al. 2006). Approximately one-half of Zambians are age 0-14, and of this age group relatively few have HIV, causing the Zambian youth to serve as a "window of hope" for a future free from the burden of AIDS (Avert 2013). A major way the government has sought to reach the minds of these youth is via awareness campaigns, such as the Helping Each Other Act Responsibly (HEART) Campaign, which launched in 1999 and was designed by the Government of Zambia and the United States Agency for International Development (USAID) (Underwood et al. 2006). The HEART Campaign uses both formal and informal means of providing HIV education, such as televised public service announcements, radio advertisements and grassroots community-based efforts to disseminate STI and HIV prevention messages (Underwood et al. 2006). These efforts by the Zambian government are extremely crucial in fighting the spread of HIV and AIDS because, as was concluded from the literature review, the presence of this information is what may allow education to act as a preventative measure against HIV. For example, those who reported that they viewed multiple campaign ads were shown to have more knowledge of HIV, particularly regarding methods to protect themselves from contracting HIV, than non-viewers (Underwood et al. 2006). While this increased informal awareness raising can have substantial effects on individuals with different levels of formal schooling, it is likely, based on the hypotheses above, that the presence of information will have the greatest effects on the most educated.

Despite the prevention efforts by the Zambian government, misconceptions about HIV and AIDS among Zambian youth are still extremely common. In 2007, two-thirds of young people aged 15-24 could neither identify misinformation about HIV nor identify major ways of preventing the transmission of HIV, showing that there is still progress to be made (Avert 2013).

3.3 The Basic Education Sub-Sector Investment Program

Because formal education, combined with informal education efforts, holds such a large role in shaping HIV dynamics, it is also crucial to examine the past and present state of education in Zambia. In addition to the HIV and AIDS epidemic, a historically troubled education sector has also plagued Zambia. Prior to 2000, schooling in Zambia was characterized by low enrollment, below standard learning outcomes, and high school fees that prevented children from poor families from attending school. Additionally, poor learning conditions such as low quality classrooms and shortages of teaching materials exacerbated these issues (World Bank 2007).

In light of the priorities set by the Millennium Development Goals, the government of Zambia began to focus on education reform and in 1999 created plans for the Basic Education Sub-Sector Investment Program (BESSIP). Funded largely by the World Bank, the primary objectives of BESSIP were to "(1) provide relevant education for all children in relevant age range and particularly for disadvantaged groups, with gross enrollment rates increase to 100% and net enrollment rates to increase to 90%; (2) improve learning outcomes for all children…"

(World Bank 2007). Various reforms were implemented under BESSIP, including the declaration of free primary education in 2002 (grades 1-7) and infrastructure development to improve learning outcomes (Hamusunga 2012).[4] BESSIP was extremely successful and exponentially stimulated a demand for education, particularly amongst the poor. Main measures of BESSIP's achievement include the increase in primary school enrollment, which rose from 2.9 million children in 2004 to 3.6 million children in 2009, and the Net Enrollment ratio, which rose from 93 percent in 2005 to 97 percent in 2009 (Hamusunga 2012).[5] Additionally, grade 9 completion rates increased from 43 percent in 2005 to 52 percent in 2009 (Hamusunga 2012).[6] The reforms under BESSIP are particularly critical in halting the spread of HIV in Zambia because of education's ability to affect how an individual responds to awareness raising efforts – in this case, the Zambian government's information campaigns.

---

[4] BESSIP reforms were extremely comprehensive, but for the purposes of this study I only go into detail about the declaration of free primary education.

[5] Net Enrollment Ratio is defined by the United Nations Statistics Division as the number of children enrolled in primary, secondary, or tertiary school as a percentage of the total children of the "official school age population" (United Nations Statistics Division 2013).

[6] The statistic given here is measured at the grade 9 level rather than grade 7 because at the time BESSIP was evaluated primary school was in transition to include grades 8-9. Therefore, much of the BESSIP assessment was done at the grade 9 level. However, the statistic is still informative because it implies that more individuals were able to achieve educational levels *above* grade 7 due to the incentive provided by the declaration of free primary education.

4. Analysis

The data I use in my analysis is from the 2007 Zambian Demographic and Health Survey (Central Statistics Office and Macro International Inc. 2007). The Zambian Demographic and Health Survey (ZDHS) is a nationally representative survey of 7,146 women age 15-49 and 6,500 men age 15-59 from randomly selected households throughout Zambia and includes a household survey, a women's survey, and a men's survey – all of which I use in my analysis. The 2007 ZDHS is only the second survey in Zambia to provide population-based prevalence estimates for HIV.

The ultimate question I want to explore in my analysis is – given the information above detailing the transformation of the association between education and HIV infection, at what stage in the transformation is Zambia? Because the ZDHS data were collected in 2007, following attempts by the Zambian government to spread awareness about the risks of HIV and simultaneously to expand education, I hypothesize that education will be working as a social vaccine. As mentioned previously, a limitation of the previous studies is endogeneity. My analysis improves upon previous studies by comparing the individuals in the dataset who have been affected by Zambia's free education policy reform in 2002 (a treatment group) versus those who have not (a control group) and thus isolating the exogenous variation in education.

To form the treatment and control groups, I create two sub-samples from the population: a control group consisting of those who were above primary school age at the time of the policy implementation and a treatment group of those who were able to enjoy the benefits of free primary education. In Zambia, primary school runs from grades 1-7 and students attend primary school from age 7- 14 (World Bank 2007).[7] Because the DHS dataset was collected in 2007, the individuals who were of primary school age in 2002 were ages 15-18 at the time of data collection.[8] The 15-18 age bracket constitutes my treatment group. In 2007, the individuals who were age 19-23 at the time of data collection were above primary school age in 2002 and therefore are used as my control group. See Figure 1 for a visual representation of this timeline.

Furthermore, looking at the cumulative frequency of educational attainment by age (see Table 1), it is clear that at seven years of education, which is the maximum number of years that all individuals in the sample could have achieved, the number of respondents who completed at least seven years of education rises with age, and then falls after reaching a maximum of 73.77 percent at age 19. There is a discontinuity in the percentage of individuals who completed grade seven or above that begins at a point in the 18-19 age range. I attribute this divide to the BESSIP free primary education policy implementation taking effect.

Figure 2 summarizes the relationship between schooling and age in my sample. Using the discontinuity in educational attainment at age 18 a line of best fit relating education and age displays a positive relationship for individuals ages 15-18 but a negative relationship for individuals ages 19-23. This phenomenon supports my choice to use ages 15-18 as a treatment group and ages 19-23 as a control group. Also shown in Figure 2 is a line for a quadratic relationship between education and age. While there does not appear to be strong evidence of a quadratic relationship between the two variables, a discontinuity in schooling at age 18 is still apparent.

---

[7] At the time of the World Bank's evaluation of the "Free Primary Education Policy" in 2007, primary school was in transition from Grades 1-7 to Grades 1-9 (World Bank 9). However, because the data I will be using are from 2007, this transition will not affect my results.

[8] Those who were of primary school age in 2002 were ages 11-18 in 2007. However, the DHS data only include information on individuals who are 15 or older, which is why my treatment group begins at age 15.

My choice to select age 18 as the upper boundary of the treatment group is solely data-driven. I understand that selecting treatment and control groups using these methods presents issues, namely the fact that education in Zambia may not be an entirely linear process. For example, an individual who was above primary school age before the free education policy was implemented may have decided to return to school once fees were removed. Even though this individual technically belongs to the treatment group, using my methods he or she would be placed in the control. To account for these issues, as well as the ambiguity of the discontinuity shown in Table 1, I will also conduct my analyses using ages 15-19 as the treatment group and ages 20-23 as the control.

4.1 Data Description

My total sample of individuals age 15-23 is composed of 3,916 Zambians, 45.6 percent of whom are male and 54.32 percent of whom are female. In the total sample, 6.08 percent of individuals are HIV positive. The variables I use in my subsequent analyses are HIV status, education, urban, wealth and sex.[9] Later I will include a variable for the policy implementation.

The means of both the dependant and independent variables - HIV status, education, urban, wealth and sex - by treatment and control group are given in Table 2. Accompanied t-tests for differences in means show that the average HIV status for males of both groups is statistically equal with a mean of 0.03 (p = 0.68) but statistically greater for females in the control group

[9] I include variables for wealth, sex, and urban status because these variables are all strong correlates of both education and HIV status and thus must be controlled for.

a. HIV status is a dummy variable that takes the value 0 if the individual is HIV negative and 1 if they are HIV positive. HIV status was not self-reported but was found from testing.

b. Education is measured in years and is self-reported.

c. Urban is a dummy variable that takes the value 1 if the respondent lives in an urban area and 0 otherwise.

d. Wealth is measured in quintiles and is recorded as 1 for poorest, 2 for poorer, 3 for middle, 4 for richer and 5 for richest.

e. Sex is a binary variable with a value of 1 for male and 2 for female.

f. The policy variable has a value of 0 for control group and 1 for treatment group.

(control = 0.11, treatment = 0.06; p= 0.00). The mean years of education is significantly larger for both males in the control group (control: 8.17, treatment: 6.76; p= 0.00) and females in the control group (control: 7.14, treatment: 6.84; p= 0.03) – a difference that is anticipated because of the older individuals in the control group who, because of their age, have had the opportunity to achieve high educational attainment. Additionally, we see that in the treatment group the gender disparity in educational achievement is significantly smaller than in the control. A t-test for the differences in mean educational attainment by gender shows that for the treatment group the means are statistically zero (p= 0.49), while in the control group the difference is statistically different than zero (p= 0.00). This trend makes sense in the context of BESSIP because an additional achievement of the policy was that it narrowed the gender disparity in primary school education (Hamusunga 2012). In both the treatment and control group the same percentage of individuals live in urban areas (control: 0.49 for male and 0.50 female, treatment: 0.47 for male and 0.51 for female; p = 0.79). Lastly, the mean wealth among all groups in the sample ranges from 3.34 to 3.44 – a mean that lies in the middle wealth quintile – and is not statistically different between the treatment and control groups (p = 0.13). In general, other than the mean of characteristics affected by the BESSIP policy, the mean statistics of the treatment and control group are not statistically different, which will strengthen my analysis when comparing groups by educational attainment.

4.2 Methodology

An initial step to analyze the association between education and HIV status in the whole sample is by conducting an ordinary least squares (OLS) regression with the equation:

$$HIV = \beta_0 + \beta_1 Education + \beta_2 Wealth + \beta_3 Sex + \beta_4 Urban + \beta_5 Age + \beta_6 Age^2 + \beta_7 Education^2 + \varepsilon$$

I control for both age-squared and education-squared to attempt to capture variations in age and education that may not be controlled for in a linear model (in Figure 2 it appears that there may be some curvature to the relation, although not strongly).

In order for OLS estimates to be valid, key ordinary least squares assumptions must be satisfied. The assumptions for multivariate regressions given by Dougherty (2011) are:

1. Linearity: Y_i is a linear function of the x_is, plus the error term ε_i. This includes the model having no omitted variables and the correct functional form.
2. Each observation has a random error with a mean of zero: $E(\varepsilon_i) = 0$.
3. Exogeneity: All explanatory variables are uncorrelated with the error term, so $E(\varepsilon_i|X_{1i},X_{2i},\ldots, X_{Ki},) = 0$. This specification also includes no omitted variables.
4. No Serial Correlation: The errors terms across observations cannot be correlated with each other.
5. Homoskedasticity: The errors term has a constant variance.
6. No perfect multicollinearity: Independent variables cannot be a perfect linear combination of other regressors.

However, the results of the OLS regression above are likely biased because key OLS assumptions are not satisfied, namely the exogeneity requirement. Because age, education, wealth, and HIV status are all components of socioeconomic status, they are endogenous to the model (causing u_i to be correlated with the regressors). This means that they "are variables whose values are determined by the interaction of the relationships in the model" and thus the results given from OLS are not accurate (Dougherty 2011). The two main sources of endogeneity in the OLS model are reverse, or simultaneous, causality and omitted variable bias. Reverse causality exists because, while education may either positively or negatively affect one's HIV status, one's HIV status also contributes to their ability to pursue or not pursue an education; it is therefore difficult to obtain a coefficient on education that is not influenced by HIV status using an OLS regression. Furthermore, because of potential omitted variables that may not be included in the model, the coefficient on education also incorporates the effects of the missing variables.

Therefore, because of these sources of bias, the results from the OLS regression must be interpreted with caution.

Overall, the OLS regression is an inadequate approach to model the relationship between education and HIV. To correct for the issues described above and construct a more robust model of the relationship between education and HIV, I use an instrumental variable regression. An instrumental variable regression solves issues of endogeneity by utilizing a variable (instrument) that is itself exogenous (uncorrelated with the error term) but also correlated with the endogenous variable, X_i, of interest. By using this instrument to isolate the exogenous variation in X_i we can then conduct a regression using the exogenous variation in X_i as a new independent variable. In my study, I use the exogenous shock to education caused by the BESSIP policy as an instrument for education in an instrumental variable regression relating education to HIV status. By using an instrumental variable regression we can account for the omitted variable bias and possible reverse causality that were restricting my analysis above in order to possibly reveal a causal relationship between education and HIV.

To use Zambia's free education policy as an instrument that isolates an exogenous variation in education I construct a dummy variable that equals 1 if the individual was in primary school at the time of the policy implementation (age 15-18) and equals 0 if the individual was not. If valid, the instrument separates the variable for education into two parts: the variation that is correlated with u_i and a part that is exogenous to the model. HIV status can then be regressed on this exogenous variation for a more robust representation of the relationship between education and HIV status. The first-stage of the IV regression is:

$$Education = \pi_0 + \pi_1 Policy + \pi_2 Age + \pi_3 Policy \bullet Age + \pi_4 Age^2 + \pi_5 Policy \bullet Age^2 + \pi_6 Wealth + \pi_7 Urban + \pi_8 Sex + v_i,$$

where $Policy \bullet Age$ and $Policy \bullet Age^2$ are instruments for education. Stage two is:

$$HIV = \beta_0 + \beta_1 Educ\hat{a}tion + \beta_2 Age + \beta_3 Age^2 + \beta_4 Wealth + \beta_5 Urban + \beta_6 Sex + u_i$$

$Policy \bullet Age$ and $Policy \bullet Age^2$ are valid instruments because they fulfill the relevance requirement [corr $(Z_i, X_i) \neq 0$] and the exogeneity requirement [corr $(Z_i, u_i) = 0$] (Dougherty 2011).

Because it can also be argued that the BESSIP policy takes effect at age 19 (see Table 1), I give regression results for a policy break at 19 alongside my results for the break at 18. Additionally, in order to test the robustness of my model I narrow my sample size from ages 15-23 to ages 17-20. Lastly, I run the instrumental variable regression by gender to assess the heterogeneity of education's effects on HIV status for males versus females.

In all my regressions I utilize the population weights provided by the Zambian Demographic and Health Survey in order to account for the fact that the distribution of characteristics in the sample may not be representative of the larger Zambian population. This can arise in survey data collection because of sampling biases such as non-response of individuals and households as well as under- or over-representation of certain groups arising from sample selection issues. DHS produces the sample weights by calculating the sample selection probabilities of each household, as well as response rates for both households and individuals. These weights are then standardized so that the sum of the standardized weights is equal to the sum of the cases from the whole sample (Demographic and Health Surveys and ORC Macro 2006).

Additionally, because the 3,916 individuals in my sample come from only 2,688 households, correlations in responses between individuals within a household may result in inconsistent standard errors, particularly underestimated standard errors. This issue, called an intraclass correlation, can affect significance tests. To correct for this issue as well as shared unobservable household characteristics, I cluster standard errors by household (UCLA: Statistical Consulting Group 2013).

5. Results

5.1 OLS Regression Results

The results from the OLS regression are given in Table 3. When HIV is regressed on education alone, the result is not significant. The coefficient on education remains insignificant once the additional controls, wealth, sex, urban, age, age-squared, as well as education-squared are added.

Sex ($p<0.01$), urban ($p<0.01$), age-squared ($p<0.1$) and education-squared ($p<0.1$) are significantly correlated with HIV status. According to the OLS model, females are 3.8 percent more likely to contract HIV than males and those who live in urban areas are 2.6 percent more likely to contract HIV than their rural counterparts. Additionally, because the coefficient on age is negative and the coefficient on age-squared is positive (coefficient of 0.001), it can be concluded that as an individual ages, he or she is less likely to contract HIV at an increasing rate. Furthermore, because the coefficient on education is positive, but the coefficient on education-squared is negative, as an individual attains an additional year of education, he or she is more likely to be HIV positive, but at a decreasing rate. Lastly, when all independent variables are included in my model the value for r-squared is 2.0 %. From this model, we cannot conclude that education has either a negative or positive correlation with HIV status. However, the results of the OLS model are to be interpreted with caution.

5.2 Instrumental Variable Results

The first-stage and second-stage results of the instrumental variable regressions are given in Table 4. The results for the regression using a schooling discontinuity at 18 and a discontinuity at 19 are both included in the table. The first stage of the instrumental variable regression regresses education on the dummy variable for the free education policy, age, the free education policy interacted with age, sex, urban, and wealth followed by an additional regression that includes the free education policy interacted with age-squared and age-squared. Importantly, the variables for the free education policy ($p<0.01$), the free education policy interacted with age ($p<0.01$), and the free education policy interacted with age-squared ($p<0.05$) are all significant in every model, except model 4 in which age-squared is included in the regression using the policy break at age 19. This signifies that there is a definite discontinuity in educational attainment that occurred once the BESSIP free education policy was implemented in 2002 and thus validates my use of the policy as an instrument for education. Furthermore, the F-statistic for instrument significance in the first stage of the instrumental variable regression is 68.76 ($p = 0.00$) for model 1 (policy break at 18, age-squared not included), 43.94 ($p = 0.00$) for model 2 (policy break at 18, age-squared included), 65.65 ($p = 0.00$) for model 3 (policy break at 19, age-squared not included) and 42.81 ($p = 0.00$) for model 4 (policy break at 19, age-squared included). Because the F-statistics are significant and greater than 10, my choice of using the BESSIP policy as an instrument for education and using ages 18 and 19 as the time of the policy implementation is further validated.

In general, the treatment group is correlated with having less education ($p<0.01$), which is anticipated because the individuals in the group are younger than those in the control and thus have been unable to reach the same educational attainment as some of their older counterparts. Additionally, the coefficient for the free education dummy variable interacted with age is significant ($p<0.01$) and positive for every model, except model 4 explained above. For those individuals in the treatment group, a one-year increase in age is correlated with a 0.414 year increase in educational attainment for model 1 (policy break at age 18, age-squared not included) and a .384 increase in educational attainment for model 3 (policy break at age 19, age-squared not included). Once age-squared is included in the regression using the policy break at age 18, the dummy variable for free education interacted with age is also significant and positive. Lastly, the variable for the free education policy interacted with age-squared is significant ($p<0.05$) only when the policy break occurs at age 18 and has a negative value of -0.164, implying that while education increases with age for those in the treatment group, it is at a decreasing rate.

Other significant correlates in the first-stage regression are sex ($p<0.01$), urban ($p<0.01$), and wealth ($p<0.01$). The coefficient for sex is negative (-0.493 in model 1) meaning that, as expected, the males in the sample have greater access to education than females. Indeed, the subsequent models give similar results and the coefficient for sex is -0.490 in models 2 and 4 and -0.493 in model 3. The coefficient for urban is positive, with a value of 0.6 in model 1, implying that living in an urban area is correlated with greater educational attainment. The coefficients for urban in the additional models give the same results with values of 0.601, 0.596 and 0.605. Lastly, in all models wealth is correlated with greater educational attainment, with coefficient values of 0.958, 0.959, 0.959 and 0.957.

Once education is instrumented for, and HIV is regressed on the instrumented education and the additional controls, the second stage of my instrumental variable regression shows significant results. In model 1 and 3 (policy breaks at both 18 and 19, not including age-squared in the regression) the coefficient on education is significant ($p<0.1$) and negative. Both models display a coefficient of -0.028, meaning that those who attain an additional year of schooling are 2.8 percent less likely to be HIV positive. Additionally, in both models 1 and 3, the coefficient on age is significant ($p<0.01$), positive, and has a value of 0.012. This implies that an individual's probability of being HIV positive rises by a proportion of 0.012 per year. Sex in models 1, 2, 3 and 4 is also significant ($p<0.05$ for model 1 and 3, $p<0.1$ for model 2 and 4) and positive in all four regressions, implying that females are significantly more likely than males to be HIV positive. In models 1 and 3, sex displays a coefficient of 0.024; females are 2.4 percent more likely than males to be HIV positive. This coefficient becomes 0.037 in model 2 and 0.076 in model 4. Urban is also significant ($p<0.05$) and positive in models 1 and 3, with a coefficient of 0.042 and subsequently, compared to those living in rural areas, urbanites are 4.2 percent more likely to be HIV positive. Lastly, wealth is significant ($p<0.1$) and positive in models 1 and 3, with a 3.2 percent increase in the likelihood of an individual being HIV positive as he or she moves up a wealth quintile.

An interesting finding to note is that age-squared is neither a significant correlate of education nor HIV status. While implementing additional controls for age is a necessity for a model like mine, the relationships between age and education and age and HIV status clearly do not display a quadratic form. Figure 2 also supports this conclusion in regard to age and education. Additionally, after age-squared is included in the model, education, age, urban and wealth are no longer significant. This may be due to the fact that because the data do not display a quadratic form, by including age-squared in the model I am including an unnecessary control.

Furthermore, I can deduce that my results are robust because the conclusions drawn from my model using the policy break at age 18 do not change when the policy break occurs at age 19. However, the instrumental variable regression using the policy break at age 18 has slightly stronger first-stage results in that the coefficients have smaller standard errors. Additionally, the coefficients in the first-stage remain significant even after a policy interaction term with age-squared and age-squared are added. As I move forward in my analysis I will continue to use the variable for a policy break at age 18 due to its slightly stronger first-stage results.

5.3 Robustness Check

In order to address concerns that my results may display bias due to the asymmetry of my age window, as well as the large size of my sample, I conduct a robustness check. To do this, I run the same instrumental variable regression as above, using only the individuals between the ages 17-20 (1,813 observations). Additionally, because the smaller age window significantly reduces the variation in possible educational attainment by removing the oldest and youngest individuals from the sample, I address concerns that nonlinear trends in schooling may skew my results. I use the variable for a policy implementation at age 18. The results of this regression are displayed in Table 5.

Because the sample size is reduced from 3,916 individuals to 1,813, the model loses power, which results in a large variation in standard errors and a loss of significance for many of the variables, including the instruments. Using this smaller sample, the F-statistic for instrument significance is only 5.87 ($p = 0.00$). Although the F-statistic is significant, it is less than 10, implying that the instruments are weak when a smaller age window is used. However, despite the reduction in power and loss of instrument strength, the results in both stage 1 and stage 2 remain consistent with the results from the larger sample, showing that my results above are indeed robust. In both stages of the robustness check, all of the explanatory variables retain the same direction in correlation as in Table 4, albeit with different magnitudes. Additionally, when HIV status is regressed on the instrumented education, the coefficient on education is negative with a value of -0.016.

5.4 Gender Effects

In Table 2, I give the mean HIV status by gender and by group. From this table, it appears as if the variation in females HIV rates would be the cause for a significant relationship between education and HIV status because of the significant difference in female HIV status between the treatment and control groups (0.11 for control; 0.06 for treatment). Furthermore, the literature supports this hypothesis in that many studies argue that education has greater effects on female HIV status than male HIV status, by significantly delaying their sexual debut and giving them more bargaining power with their male sexual partners. To test the dominant hypotheses provided by the literature, as well as the conclusions implied by Table 2, I run my model separately for females and males.

Table 6 displays the effects of education on HIV status by gender using the variable for the policy break at age 18. Importantly, when the instrumental variable regression is conducted by gender the instruments remain significant. In the first stage, the instruments (policy interacted with age and policy interacted with age-squared) are significant for females in both models ($p<0.01$ in model 3; $p<0.05$ in model 4), but only significant for males before the policy interacted with age-squared is used as an instrument. The F-statistics for instrument relevance are all significantly greater than 10, with a value of 74.44 ($p = 0.00$) for model 1, 47.17 ($p = 0.00$) for model 2, 16.39 ($p = 0.00$) for model 3 and 12.99 ($p = 0.00$) for model 4.

Additionally, many of the gender-specific results remain similar to the pooled regressions. For both genders, when education is regressed on the instruments and the controls in stage 1, urban is significant ($p<0.01$) and positively correlated with education with a value of 0.583 for males (model 1) and 0.557 for females (model 3). Models 2 and 4 give similar results, with coefficients for urban of 0.577 and 0.559. Wealth is also significant ($p<0.01$) and positively correlated with education, displaying a value of 0.774 for males (model 1) and 1.140 for females (model 3).

However, in stage 2 of the gender-specific regressions, the instrumented education is only significant ($p<0.01$) for males and only significant before education is instrumented with the policy interacted with age-squared. When HIV status is regressed on the instrumented education, the coefficient on education for the male sample is -0.045, implying that it is the males in the sample who are the cause for the significant and negative relation between education and HIV status shown in the pooled regression. This may be attributed to the fact that, despite the Zambian government's free education policy promoting gender equality in schooling, more males in the treatment group have achieved higher levels of schooling than females once population weights are used. Therefore, when gender-specific regressions are run, different ages must compose the control and treatment groups in order to account for a lag in female assimilation into the school system. Additionally, the UNAIDS Inter-Agency Task Team on Education (2008) finds that girls who complete secondary education have a significantly lower risk of being HIV positive and adopting protective sexual practices than their female counterparts who have only completed primary school. Because the females in my treatment group are only ages 15-18, many are not old enough to have completed secondary school. This may potentially affect the regression results and be the cause for education not yet acting as a preventative factor against HIV for females.

5.5 Limitations

This study has limitations, the first of which being that the ZDHS data do not include information on which individuals specifically were affected by the BESSIP policy. Without this information, I assume in the model that everyone age 15-18 received education under BESSIP, which is not a perfect assumption. Additionally, as is always the case with studies utilizing DHS data, many of the measures – such as wealth and educational attainment – are self-reported. Furthermore, as mentioned above, the individuals who were of age to enjoy the free primary education offered by BESSIP in 2002 were only ages 15-18 at the time of data collection in

2007. A survey taken at a later year, giving the individuals in the treatment group more opportunity to increase their level of education attainment as well as to mature sexually, would likely produce more robust results using my model.

An ideal dataset for measuring the relationship between HIV and education in Zambia using my current model would include information on which individuals received education directly under the BESSIP program. This would allow me to more precisely assign individuals to treatment and control groups and conduct a more robust instrumental variable regression. Additionally, it would be useful to obtain data on these individuals over time, since the youngest individuals in the 2007 dataset are only 15-18. This would also account for time-specific factors that may be contributing to the 2007 relationship. Finally, a different and more accurate approach to measuring the relationship between education and HIV would be to conduct a controlled experiment. The data from the experiment would include HIV data over time from a randomly selected treatment group of youth who receive incentives to attend primary school and a randomly selected control group who do not. We can then construct an ordinary least squares regression for each group, regressing HIV on educational attainment, to compare the coefficients on education. Any differences in the values of the coefficient on education and HIV outcomes between the groups can be attributed to differences in an exogenous variation in educational attainment.[10]

[10] The random assignment of individuals to treatment and control groups ensures that, in theory, the groups are similar in all characteristics except the fact that the treatment group received incentives to attend primary school. This can be proven by conducting a baseline survey. Therefore, any variations in education are exogenous to the model.

6. Discussion

"Education is the most powerful weapon you can use to change the world."
Nelson Mandela

"Universal primary education could save at least 7 million young people from contracting HIV over a decade. However, without dramatic increases in aid to education, Africa will not be able to get every child into school for another 150 years."
Global Campaign for Education 2004

6.1 Results

Both my theoretical and empirical analyses prove that education can work as a social vaccine against HIV if the proper mechanisms are in place, namely the presence of HIV prevention information. This study builds on the existing literature examining the relationship between education and HIV; however, by using an instrumental variable regression to model this relationship in Zambia, I fill a void in the discourse arising from a lack of studies conducted that isolate an exogenous variation in education.

While it is true that prior empirical studies of the correlation between education and HIV lack consensus, I conclude that this is due to actual heterogeneity in the relationship over time and location. The theories stating that education has a positive, negative, null and changing relationship with HIV are not mutually exclusive but, in reality, are each driven by a set of underlying mechanisms that form the relationship. The risk factors that accompany education - such as expansive sexual networks and greater disposable income - coupled with the preventative factors that also rise with education – including increased condom use and greater exposure to prevention messages – are both present in shaping the relationship (De Walque and Kline 2010; Baker et al. 2009; Cogneau and Grimm 2006; Hargreaves et al. 2008). While it may seem that these factors would work in opposite directions to create only a null correlation between education and HIV, one set of factors often outweigh the other, which has motivated the four theories above. The external dynamic that determines the direction of the correlation is the presence of HIV information campaigns in a country and this factor greatly explains why the association between education and HIV appears to have changed over time. Educational attainment plays an integral role in helping individuals to apply and comprehend messages about HIV and AIDS in order to protect themselves, and so before governments provided their citizens with important HIV information, education was unable to occupy a role as a preventative measure against the epidemic (Hargreaves and Boler 2006).

My empirical analysis of Zambia supports these conclusions. Because, the data I use is from the 2007 Zambian Demographic and Health Survey (Central Statistics Office and Macro International 2007) – a time in Zambia's history characterized by the long-time presence of HIV prevention campaigns - I hypothesize that education in Zambia should be working as a social vaccine. Using the Zambian free education policy as an instrument in an instrumental variable regression, I find that education has a significantly negative relationship with HIV. This relation holds true for both versions of my model – one using the free education policy implementation for 18 year olds and the other for 19 year olds. However, when separated by gender, this relationship only holds true for males. From this analysis it appears that education can work as a social vaccine against HIV if provided in unison with outside prevention information, and is a tool in which governments should invest in their pursuit against the epidemic.

6.2 Further Research

While countless scholars have studied the relationship between education and HIV, both across space and time, there is still further research to be done in order to gain a complete understanding of their connection. First, in the education sector, it is important to know whether primary school or secondary school has the greatest impact on one's HIV status. While investing in both levels of education is ideal, with this information, countries with limited resources can make more effective decisions regarding where to invest. Additionally, we need more research regarding young people and the underlying factors in their lives that form the relationship between education and HIV. For example, understanding whether economic status, educational attainment, or gender is the stronger determinant of HIV status may be important for directing initial prevention efforts. However, all of these factors must ultimately be addressed. Finally, future research needs to be directed towards studying the information that can best change the behaviors of young people that may put them at risk for HIV. Through what forms of media is HIV prevention information most likely to reach all youth? What are the most effective types of messages to include in prevention campaigns?

6.3 Prevention and Policy Recommendations

While theoretical research is integral in fighting HIV, there are also tangible steps that can be taken in pursuit of this goal. My theoretical conclusions and empirical results may give some insight as to how prevention efforts can be best directed, both in Zambia and in other countries suffering from high HIV prevalence rates:

1. First, prevention efforts need to be targeted towards women and girls. In Africa, there is said to be a "feminization" of AIDS, as 74 percent of young people living with HIV are women (Hargreaves and Boler 2006). Indeed, women bear the majority of the burden of HIV and AIDS in many countries, highlighted by the fact that my results do not show significant negative results for women. Instead, my results show that as education in Zambia increased under the BESSIP policy, this increase only had a significantly negative effect on male HIV rates, despite the fact that women's educational attainment also increased under the policy. Therefore, schooling may be a necessary, but not sufficient step, in preventing HIV in women. In the case of Zambia, women and girls need to be reached directly via outside information campaigns in order for education to become a preventative factor against HIV. While basic information campaigns exist, one way to improve upon the existing efforts is to create campaigns that are more purposeful and relevant to their population of interest. For example, Hargreaves and Boler (2006) explain that HIV prevention campaigns typically do not attempt to raise

awareness for the vulnerability of women. While campaigns may advocate for condom use, they do not first address the fact that women often lack the power to decide who to have sex with, let alone the power to discuss using protection during intercourse. Therefore, prevention messages need to also bring the subject of gendered power dynamics in sexual relationships into public conversation.

2. We also need universal primary (and secondary) education in every country. My analysis, as well as countless other studies, shows that education can work as a social vaccine against HIV. However, despite this body of evidence, the majority of children in Africa will not complete primary school. Additionally, most children in Africa have to pay a substantial fee to attend primary school, which results in their inability to enroll (Hargreaves and Boler 2006; Global Campaign for Education 2004). The government of Zambia has recognized this issue and, with its free primary education policy (2002) implemented under BESSIP, has taken important steps to combat the HIV epidemic.

3. As explained in the literature review, an expansion in education can increase in efficacy when coupled with outside information campaigns. The Global Campaign for Education (2004) argues that an increase in schooling should not be an alternative to HIV and AIDS treatment and prevention efforts (Global Campaign for Education 2004). It has been proven that this two-pronged approach can be extremely successful. In the case of Uganda, HIV prevalence rates fell from 15 percent in 1990 to five percent in 2000 because of the government declaration of free primary education as well as their use of aggressive and pointed information campaigns (Global Campaign for Education 2004).

4. Lastly, schools themselves can take initiative in preventing HIV. One of the most notable advantages of schools in protecting their students from HIV is their ability to directly disseminate HIV information to their students. Therefore, sexual health education - including HIV awareness and family planning concepts - is a critical component of schooling. However, a worldwide study finds that 40 percent of countries have not attempted to provide HIV and AIDS education in their nation's classrooms (Global Education Campaign 2004).

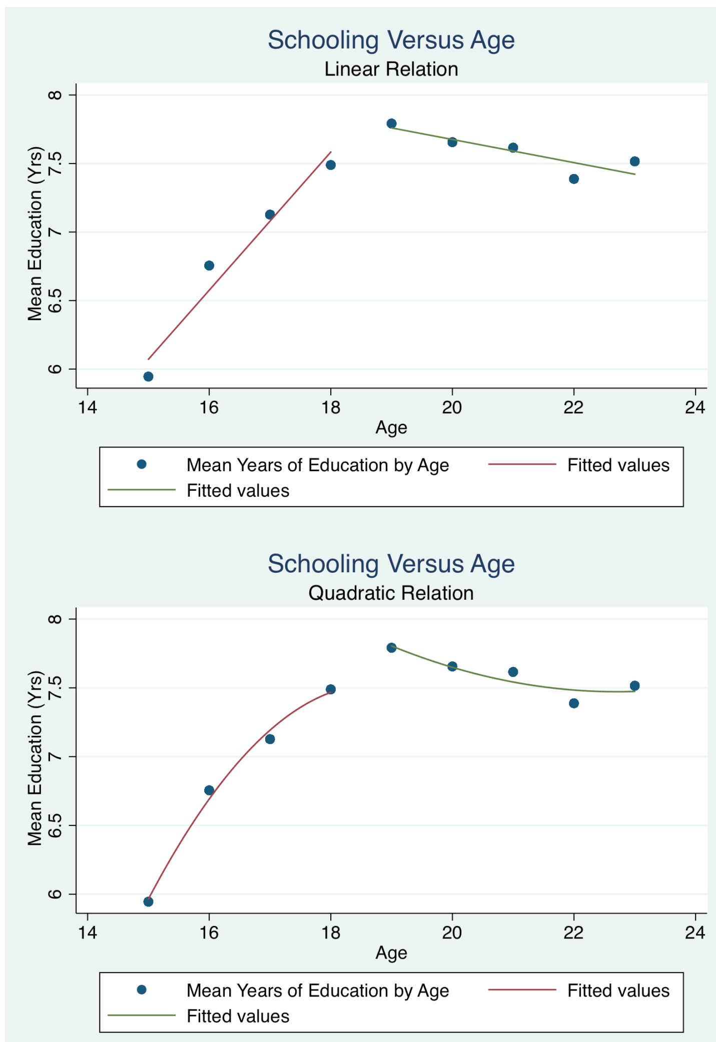
In light of the 25th anniversary of the first HIV diagnosis, in 2008 over one hundred countries promised to ensure universal provision of HIV and AIDS prevention information and care to their citizens (Hargreaves and Boler 2006). The above recommendations give crucial education-focused goals towards which these HIV prevention efforts can be directed. However, increased access to schooling and more relevant information campaigns are only a small part of a complex web of issues that affect the HIV and AIDS epidemic, such as poverty, limited access to anti-retroviral drugs and societal gender inequality. Therefore, while education is a powerful tool, a more holistic approach must be taken in working towards an AIDS free future.

7. Appendix 1

Figure 1: Timeline

** In 2002, the BESSIP policy was implemented and primary education became free in Zambia. Those who were of age to take advantage of free primary school make up my treatment group and are the individuals ages 15 - 18 in the 2007 DHS data.

Figure 2: Relation between schooling and age for treatment and control groups



Note: Both of these figures graph the relationship between age and the mean years of education by age in the sample. To assess if there is a discontinuity at age 18, as hypothesized, I fit a line relating education and age for individuals between the ages of 15-18 (treatment group) and a line for individuals between the ages 19-23. To assess the shape of the relation, I first fit a linear relation then a quadratic relation.
Source: Author's analysis based on data from Zambian Demographic and Health Survey (Central Statistics Office and Macro International Inc. 2007).

8. Appendix 2

Table 1: Cumulative frequency of education by age

	Age								
Education (Yrs)	15	16	17	18	19	20	21	22	23
0	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1	97.46	98.12	99.14	95.18	96.19	92.84	94.78	93.20	93.77
2	96.48	97.49	98.92	94.52	95.29	92.17	93.68	91.18	92.35
3	93.15	95.40	97.84	93.86	93.05	89.93	89.56	88.92	88.95
4	87.87	92.89	94.61	90.35	88.79	87.70	85.44	84.13	85.84
5	78.08	88.28	88.36	86.18	84.08	83.67	81.87	77.83	77.90
6	63.99	76.36	77.37	79.61	79.60	76.51	74.45	70.03	69.97
7	43.25	59.00	62.50	70.18	73.77	69.80	67.86	64.99	63.17
8	24.46	39.75	41.59	53.51	56.95	54.14	51.92	48.36	48.16
9	8.41	18.62	27.37	38.82	45.29	44.74	43.68	39.04	41.64
10	1.17	7.32	14.87	23.68	29.60	30.65	30.49	28.72	29.18
11	0.20	1.88	7.54	14.47	21.52	24.16	25.55	25.44	27.48
12	0.00	0.42	2.59	8.11	13.90	17.45	19.78	20.91	24.93
13	0.00	0.00	0.00	0.22	1.12	1.79	1.65	3.53	4.25
14	0.00	0.00	0.00	0.22	0.00	0.00	0.55	2.27	2.83
15	0.00	0.00	0.00	0.00	0.00	0.00	0.27	0.25	1.13

Source: Author's analysis based on data from Zambian Demographic and Health Survey (Central Statistics Office and Macro International Inc. 2007)

Table 2: Mean characteristics by group

	Control		Treatment	
Variable	Male	Female	Male	Female
HIV Status	0.03	0.11	0.03	0.06
	(0.17)	(0.31)	(0.18)	(0.24)
Education	8.17	7.14	6.76	6.84
	(3.07)	(3.74)	(2.25)	(2.59)
Urban	0.49	0.5	0.47	0.51
	(0.5)	(0.5)	(0.5)	(0.5)
Age	20.82	20.94	16.49	16.42
	(1.41)	(1.41)	(1.12)	(1.13)
Wealth	3.35	3.34	3.39	3.44
	(1.4)	(1.37)	(1.39)	(1.36)
N	895	1112	894	1015

Note: Numbers in parentheses are standard deviations.
Source: Author's analysis based on data from Zambian Demographic and Health Survey (Central Statistics Office and Macro International Inc. 2007).

Table 3: Education's association with HIV status (OLS Estimator)

	HIV Status					
	(1)	(2)	(3)	(4)	(5)	(6)
Education	0.002	-0.001	-0.000	-0.001	-0.002	0.007
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.005)
Wealth		0.012***	0.011***	0.005	0.006	0.007
		(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Sex			0.039***	0.039***	0.037***	0.038***
			(0.008)	(0.008)	(0.008)	(0.008)
Urban				0.026*	0.026*	0.026*
				(0.014)	(0.014)	(0.014)
Age					0.007***	-0.043
					(0.002)	(0.029)
Age-sq						0.001*
						(0.001)
Educ-sq						-0.001*
						(0.000)
Constant	0.044***	0.025**	-0.038**	-0.025	-0.147***	0.277
	(0.010)	(0.010)	(0.016)	(0.018)	(0.039)	(0.264)
R-squared	0.001	0.004	0.011	0.013	0.018	0.020
N	3916	3916	3916	3916	3916	3916

* p<0.10 ** p<0.05 *** p<0.01
Note: Numbers in parentheses are clustered standard errors. The data are weighted with the sample weights given by the data provider.
Source: Author's analysis based on data from Zambian Demographic and Health Survey (Central Statistics Office and Macro International Inc. 2007).

Table 4: Education's effect on HIV Status (IV)

First Stage

	Education			
	Policy at 18		Policy at 19	
	(1)	(2)	(3)	(4)
1. Policy	-7.764***	-55.000**	-7.197***	-35.341
	(1.275)	(23.276)	(1.587)	(38.987)
Age	-0.010	-0.748	0.004	-1.182
	(0.049)	(1.783)	(0.070)	(3.570)
1. Policy x Age	0.414***	5.973**	0.384***	3.403
	(0.068)	(2.466)	(0.077)	(3.693)
Sex	-0.493***	-0.490***	-0.493***	-0.490***
	(0.088)	(0.089)	(0.089)	(0.089)
Urban	0.600***	0.601***	0.596***	0.605***
	(0.132)	(0.148)	(0.148)	(0.148)
Wealth	0.958***	0.959***	0.959***	0.957***
	(0.047)	(0.053)	(0.053)	(0.053)
Age-squared		0.018		0.028
		(0.043)		(0.083)
1. Policy x Age-sq		-0.164**		-0.082
		(0.067)		(0.088)
Constant	5.067***	12.757	4.763***	17.453
	(1.033)	(18.609)	(1.502)	(38.197)
R-squared	0.324	0.325	0.324	0.324
N	3916.000	3916.000	3916.000	3916.000

* p<0.10   ** p<0.05   *** p<0.01

Second Stage

	HIV Status			
	Policy at 18		Policy at 19	
	(1)	(2)	(3)	(4)
Education	-0.028*	-0.002	-0.028*	0.078
	(0.017)	(0.041)	(0.017)	(0.087)
Age	0.012***	-0.045	0.012***	-0.196
	(0.004)	(0.088)	(0.004)	(0.176)
Sex	0.024**	0.037*	0.024**	0.076*
	(0.012)	(0.021)	(0.012)	(0.043)
Urban	0.042**	0.025	0.042**	-0.023
	(0.018)	(0.028)	(0.018)	(0.056)
Wealth	0.032*	0.006	0.032*	-0.070
	(0.017)	(0.040)	(0.017)	(0.084)
Age-squared		0.001		0.005
		(0.002)		(0.004)
Constant	-0.124***	0.329	-0.124***	1.526
	(0.040)	(0.715)	(0.040)	(1.409)
R-squared	.	0.019	.	.
N	3916.000	3916.000	3916.000	3916.000

* p<0.10   ** p<0.05   *** p<0.01

Note: Numbers in parentheses are clustered standard errors. The data are weighted with the sample weights given by the data provider.
Source: Author's analysis based on data from Zambian Demographic and Health Survey (Central Statistics Office and Macro International Inc. 2007.)

.

Table 5: Robustness check

First Stage		Second Stage	
	Education 17-20		HIV Status 17-20
1. Policy	-3.839	Education	-0.016
	(4.683)		(0.053)
Age	-0.051	Age	-0.001
	(0.197)		(0.011)
1. Policy x Age	0.185	Sex	0.030
	(0.251)		(0.021)
Sex	-0.317**	Urban	0.019
	(0.124)		(0.048)
Urban	0.788***	Wealth	0.017
	(0.205)		(0.049)
Wealth	0.909***	Age-squared	
	(0.074)		
Age-squared		Constant	0.067
			(0.094)
1. Policy x Age-sq		R-squared	.
		N	1813.000
Constant	5.690		
	(3.870)		
R-squared	0.307		
N	1813.000		

* p<0.10   ** p<0.05   *** p<0.01

Note: Numbers in parentheses are clustered standard errors. The data are weighted with the sample weights given by the data provider.
Source: Author's analysis based on data from Zambian Demographic and Health Survey (Central Statistics Office and Macro International Inc. 2007).

Table 6: Heterogeneity of Education's Effects by Gender

First Stage

	Education			
	Male		Female	
	(1)	(2)	(3)	(4)
1. Policy	-7.447***	-46.496	-7.946***	-70.327**
	(1.746)	(32.646)	(1.775)	(31.834)
Age	0.046	-1.575	-0.044	-0.515
	(0.069)	(2.552)	(0.066)	(2.423)
1. Policy x Age	0.373***	4.693	0.450***	7.930**
	(0.093)	(3.455)	(0.096)	(3.396)
Urban	0.583***	0.577***	0.557***	0.559***
	(0.196)	(0.197)	(0.203)	(0.203)
Wealth	0.774***	0.775***	1.140***	1.141***
	(0.070)	(0.070)	(0.072)	(0.072)
Age-squared		0.039		0.011
		(0.061)		(0.058)
1. Policy x Age-sq		-0.121		-0.224**
		(0.094)		(0.093)
Constant	4.368***	21.267	3.946***	8.854
	(1.438)	(26.599)	(1.424)	(25.300)
R-squared	0.320	0.321	0.347	0.349
N	1789.000	1789.000	2127.000	2127.000
* p<0.10	** p<0.05	*** p<0.01		

Second Stage

	HIV Status			
	Male		Female	
	(1)	(2)	(3)	(4)
Education	-0.045**	-0.028	-0.007	0.017
	(0.021)	(0.031)	(0.022)	(0.034)
Age	0.016**	-0.019	0.011***	-0.057
	(0.008)	(0.080)	(0.004)	(0.079)
Urban	0.038	0.028	0.042*	0.029
	(0.025)	(0.026)	(0.025)	(0.029)
Wealth	0.039**	0.026	0.012	-0.015
	(0.017)	(0.024)	(0.027)	(0.039)
Age-squared		0.001		0.002
		(0.002)		(0.002)
Constant	-0.072	0.222	-0.146***	0.434
	(0.061)	(0.692)	(0.053)	(0.669)
R-squared	.	.	0.016	.
N	1789.000	1789.000	2127.000	2127.000
* p<0.10	** p<0.05	*** p<0.01		

Note: Numbers in parentheses are clustered standard errors. The data are weighted with the sample weights given by the data provider.
Source: Author's analysis based on data from Zambian Demographic and Health Survey (Central Statistics Office and Macro International Inc. 2007).

References

Avert. 2013. “HIV and AIDS in Zambia.” http://www.avert.org/aids-zambia.htm (accessed March 8, 2013).

Baker, David P., John M. Collins and Juan Leon. 2009. “Risk factor or social vaccine? The historical progress of the role of education in HIV and AIDS infection in Sub-Saharan Africa.” *Prospects*, 38: 467-86.

Central Statistical Office (Zambia), Macro International, Inc. 2007. “Zambia Demographic and Health Survey.” Calverton, United States: Macro International, Inc.

Cogneau, Dennis and Michael Grimm. 2006. “Socioeconomic status, sexual behavior, and differential AIDS mortality: evidence from Côte d’Ivoire.” *Population Research Policy Review*, 25: 393-407.

Demographic and Health Surveys and ORC Macro. 2006. “Guide to DHS Statistics: Demographic and Health survey Methodology.” Calverton, United States: Macro International, Inc.

De Walque, Damien. 2009. "Does Education Affect HIV Status? Evidence from Five African Countries." *The World Bank Economic Review*, 23(2): 209-33.

De Walque, Damien and Rachel Kline. 2010. "The Relationship Between HIV Infection and Education: An Analysis of Six Sub-Saharan African Countries." *The Socioeconomics of HIV/AIDS in Africa*. Ed. David E. Sahn: Cornell UP, 110-33.

De Walque, Damien, Jessica S. Nakiyingi-Miiro, June Busingye and Jimmy A. Whitworth. 2005. “Changing association between schooling levels and HIV-1 infection over 11 years in a rural population cohort in south-west Uganda.” *Tropical Medicine and International Health*, 10(10): 993-1001.

Dougherty, Christopher. 2011. *Introduction to Econometrics Fourth Edition*.: Oxford UP.

Fortson, Jane G. 2008. “The Gradient in Sub-Saharan Africa: Socioeconomic Status and HIV/AIDS.” *Demography*, 45(2): 303-22.

Global Campaign for Education. 2004. *Learning to survive: How education for all would save millions of young people from HIV/AIDS.* Global Campaign for Education, Brussels.

Gregson, Simon, Heather Waddell and Stephen Chandiwana. 2001. “School Education and HIV Control in Sub-Saharan Africa: From Discord to Harmony.” *Journal of International Development*, 13: 467-85.

Hargreaves, James and Tania Boler. 2006. *Girl Power: the impact of girls’ education on HIV and sexual behaviour.* ActionAid International.

Hargreaves, James R., Christopher P. Bonell, Tania Boler, Delia Boccia, Isolde Birdthistle, Adam Fletcher, Paul M. Pronyk and Judith R. Glynn. 2008. “Systematic review exploring time trends in the association between educational attainment and risk of HIV infection in sub-Saharan Africa.” *AIDS*, 22: 403-14.

Hamusunga, George. 2012. *An Analysis of the Performance of the Zambian Education System*. Zambia National Education Coalition.

Kandala, N-B, C. Ji, P.F. Cappuccio and R.W. Stones. 2008. “The epidemiology of HIV infection in Zambia.” *AIDS Care: Psychological and Socio-medical Aspects of AIDS/HIV*, 20(7): 812-19.

LaMontagne, Bernadette, and Daniel Stockemer. 2010. "Determinants of HIV Prevalence: a Global Perspective." *International Politics,* 47(6): 698-724.

McIntosh, William Alex, and John K. Thomas. 2004. "Economic and Other Societal Determinants of the Prevalence of HIV/AIDS: A Test of Competing Hypothesis." *The Sociological Quarterly,* 45(2): 303-24.

Michelo, Charles, Ingvild F. Sandoy and Knut Fylkesnes. 2006. "Marked HIV prevalence declines in higher educated young people: evidence from population-based surveys (1995-2003) in Zambia." *AIDS*, 20: 1030-38.

Pettifor, Audrey E., Brooke A. Levandowski, Catherine MacPhail, Nancy S. Padian, Myron S. Cohen and Helen V. Rees. 2008. "Keep them in school: the importance of education as a protective factor against HIV infection among you South African women." *International Journal of Epidemiology*, 37: 1266-73.

UCLA: Statistical Consulting Group. 2013. "STATA Library: Analyzing Correlated (Clustered) Data." http://statistics.ats.ucla.edu/stat/stata/library/cpsu.htm (accessed April 15, 2013).

Underwood, Carol, Holo Hachonda, Elizabeth Serlemitsos, and Uttara Bharath-Kumar. 2006. "Reducing the risk of HIV transmission among adolescents in Zambia: Psychosocial and behavioral correlates of viewing a risk-reduction media campaign." *Journal of Adolescent Health*, 38: 55.e1-55.e13.

UNAIDS Inter-Agency Task Team on Education. 2008. *Girls' Education and HIV Prevention*. UNESCO, Paris.

UNICEF. 2003. "Zambia – Statistics." http://www.unicef.org/infobycountry/zambia_statistics.html (accessed March 8, 2013).

United Nations Statistics Division. 2013. "Millennium Development Goals Indicators." Mdgs.un.org (accessed March 8, 2013).

World Bank. 2007. *Implementation Completion and Results Report on a Loan/Credit in the Amount of US $Million 35.59 to Zambia for Basic Education Subsector Investment Program*. World Bank, Maputo.

World Bank. 2013. "Zambia – Data." http://data.worldbank.org/country/zambia (accessed March 8, 2013).

World Health Organization. 2005. "Summary Country Profile for HIV/AIDS Treatment Scale-Up." http://www.who.int/hiv/HIVCP_ZMB.pdf (accessed April 17, 2013).

World Health Organization. 2011. *Annual Report of the WHO Country Office Zambia 2011*. World Health Organization, Lusaka.