

How Does the Impact of an HIV/AIDS Information Campaign Vary with Educational Attainment? Evidence from Rural Uganda

Damien de Walque*
The World Bank, Development Research Group

Abstract

The responsiveness to information is thought to be one channel through which education affects health outcomes. This paper tests this hypothesis by examining the effectiveness of an information campaign that aims at preventing the HIV/AIDS epidemic in Uganda. Previous studies in the epidemiological literature have generally concluded that, in Africa, there was either a positive or no association between HIV infection and schooling levels. Using individual level data from a cohort study following the general population of a cluster of villages in rural Uganda over 12 years, this paper shows that, after more than a decade of prevention campaigns about the dangers of the epidemic, there has been a substantial evolution in the HIV/education gradient. Early in the epidemic, in 1990, there was no robust relation between HIV/AIDS and education. In 2000, among young individuals, in particular among females, education lowers the risk of being HIV positive. Results on HIV incidence in a duration framework confirm that finding by establishing that, for young individuals, education reduces the probability of seroconversion. These findings reveal that educated individuals have been more responsive to the HIV/AIDS information campaigns. The analysis of sexual behavior reinforces that conclusion: condom use is associated positively with schooling levels.

J.E.L. Classification: O12, I12, I2, J17.

*ddewalque@worldbank.org. I am very grateful to Dr. June Busingye, George Katongole, Jessica Nakiyingi, Julie Pickering, Anthony Ruberantwari and Professor James A.G. Whitworth from the Medical Research Council (UK)/Department for International Development (UK)/Uganda Virus Research Institute Programme on AIDS in Uganda for their hospitality in Uganda, for giving me access to the data set of the General Population Cohort and helping me with thoughtful advice and comments on the data sets and on the general situation of the HIV/AIDS epidemic in Uganda. I thank Alberto Abadie, Gary S. Becker, Hoyt Bleakley, Frikkie Booysen, Shiyao Chao, Raphael De Coninck, Mark Duggan, Michael Greenstone, Jonathan Guryan, Patrick Heuveline, Fabian Lange, Kenneth Leonard, Steven Levitt, Casey Mulligan, Kevin Murphy, Tomas Philipson, Chris Rohlfs, Desire Vencatachellum, James Whitworth and seminar participants at the University of Chicago, UQAM, Duke, Ohio State, University of Namur, Aarhus University, INSEAD, IIES at Stockholm University, University College Dublin, The World Bank, McGill, Boston University, Oxford, the NEUDC 2003, HEC Montreal and the CSAE 2004 for helpful comments and discussions. Financial support from the Flora and William Hewlett Foundation Fellowship and the Esther and T.W. Schultz Endowment Dissertation Fellowship is gratefully acknowledged. All errors are mine. The findings, interpretations, and conclusions expressed in this paper are those of the author and do not necessarily represent the views of the World Bank, its Executive Directors, or the governments they represent. Working papers describe research in progress by the author and are published to elicit comments and to further debate.

1 Introduction

The very strong correlation between education levels and health outcomes, even after controlling for income, has been recognized as a robust empirical observation in the social sciences and economic literature (Deaton and Paxson 1999; Fuchs 1982; Lleras-Muney 2001). However, the question of the mechanisms driving this correlation remains challenging. The ability to process information is thought to be one channel through which education affects health outcomes. This paper tests this hypothesis by examining the effectiveness of an information campaign that aims at preventing the spread of the HIV/AIDS epidemic in Uganda.

The HIV/AIDS epidemic is probably the greatest challenge facing Africa. According to UNAIDS, in 2003, between 25 and 28.2 million people were infected by HIV/AIDS in Sub-Saharan Africa (this represents 65 to 70 percent of the worldwide total), between 2.2 and 2.4 million died from the virus and between 3 and 3.4 million became newly infected. In Uganda, the percentage of infected adults in 2001 was 5 percent, down from 14 percent in the early 1990s (United States Government 2000). While the epidemic is still on the rise in most of Africa, especially in southern Africa, with a prevalence rate as high as 38.8 percent (UNAIDS 2002) in the adult population of Botswana, Uganda is one of the few countries that has been successful in fighting the epidemic. Therefore, there is a high potential to learn important lessons from the Ugandan experience, both in order to understand behavioral responses and from a policy perspective.

Unprotected sex was not initially perceived as dangerous in Uganda. At the end of the 1980s and early in the 1990s, as the Ugandan government and several national and international organizations embarked on large-scale prevention campaigns, information about the HIV/AIDS epidemic gradually became available and revealed the risk associated with multiple partners and unprotected sex. In addition to information about the transmission modes of the disease, the campaigns also promoted ways to avoid the infection, including condom use.

This paper uses a unique individual level data set from a longitudinal survey, the General Population Cohort (GPC) of the Medical Research Council (MRC) Programme on AIDS in Uganda, which follows the general population of a cluster of villages over 12 years. This cohort study started shortly after the initiation of the information campaign and is still ongoing, which makes it useful for assessing the impact of the prevention efforts. The longitudinal nature of the data allows us not only to observe trends in HIV prevalence, but also to analyze HIV incidence, the infection rate,

which is a more accurate measure of the epidemic. Another interesting feature of the data set is that it not only contains the results of yearly anonymous HIV testing in the population, but it also includes information about the mechanisms leading to HIV infection, in particular sexual behavior such as condom use and the number of partners.

One limitation of the analysis, however, is that there is no control group, as all surveyed villages were exposed to the information campaign. In order to contrast different levels of exposure to the information, this paper will rely instead on a distinction based on the birth cohort of the individuals and the period of observation. In particular, the paper compares those who started their sexual life before the arrival of the information and those who initiated their sexual activity afterward. The paper further compares HIV outcomes for the better and the less well-educated in the population.

The analysis indicates that, over the past decade, there has been a substantial evolution in the HIV prevalence/education gradient among the population in rural Uganda surveyed by the GPC. Early in the epidemic, in 1990, there is no significant relationship between education and HIV prevalence. In 2001, after more than a decade of information and prevention efforts about the dangers of the epidemic, for individuals who started their sexual life after the start of the prevention campaign, the higher the education level, the lower the risk of being HIV positive. This effect is mainly concentrated among females.

This paper seems to be one of the first to report robust evidence that for young cohorts in Africa there is now a negative gradient between education and HIV infection, suggesting that the effect of education only emerges over time.¹ As illustrated in figure 1, between 1990 and 2000, the HIV prevalence for young individuals with no education or primary education decreased by 6 percentage points, while it decreased by 12 percentage points for individuals who went to secondary school. Results from a study of HIV incidence in a duration framework confirm this finding by showing

¹Out of 27 studies reviewed by Hargreaves and Glynn (2002), only one, among sugar estate workers in Ethiopia, reported a significantly negative association between HIV infection and education. Most of these studies, however, are in urban settings and based on data collected in the beginning of the 1990s, at an earlier stage of the epidemic. Fylkesnes and others (2001) also has evidence of a negative relationship between HIV infection and education in urban Zambia. For other studies of the relationship between HIV prevalence and education see Blanc (2000); Gregson, Waddell, and Chamdiwana (2001); Kilian and others (1999); Vandemoortele and Delamonica (2001). The economic profession has mainly addressed the HIV/AIDS epidemic by looking at its demographic and economic consequences (Ainsworth, Beegle, and Koda 2000; Clark and Vencatachellum 2003; Garnett, Grossly and Gregson 2001; Lundberg and Over 2000; Over 1997), by investigating the compensating differential for condom use (Gertler, Shah and Bertozzi 2003; Rao and others 2003) or by studying the issue of the affordability of the treatment and the incentives to create a vaccine (Kremer 2000).

that, for young individuals, especially females, education reduces the probability of becoming HIV positive. Moreover, this paper documents some of the mechanisms leading to this result by showing that changes in behavior and sexual practices, in particular condom use and visits to "voluntary counselling and testing centers," have been more widespread among educated individuals. These findings reveal that more educated individuals have benefited more from the HIV/AIDS prevention campaign and the results of this paper are consistent with the hypothesis that education helps people to process health-related information.

However, the question of whether the effect of education in improving health is causal is important in order to determine whether these results justify larger public investments in education. One explanation (Becker 1993) for the education and health gradient stresses that more educated individuals are healthier because their investment in the future gives them stronger incentives to protect their health. Another theory (Grossman 1972; Kenkel 1991) emphasizes that education helps accessing and processing health-related information. A third view (Farell and Fuchs 1986; Fuchs 1982) claims that the observed correlation between health and education is mainly due to third unobservable factors, like the discount factor or ability. These three views are presented in a model in appendix 1.

As no plausibly exogenous variation in education levels could be found in the surveyed area, the evidence presented in this paper is not suitable to distinguish between these three channels and, therefore, to determine whether the effect of education in decreasing the risk of HIV infection is causal.

If at least some of the effect of education in reducing the risk of AIDS is causal, an important policy implication of this paper is that investing in education is not only beneficial for the labor market outcomes of individuals, but also helps them making decisions that improve their health and longevity. Rates of return to education, as traditionally calculated in the economics literature only account for labor market earnings. Given the large impact of HIV/AIDS on mortality, by showing that the interaction between educational achievement and information has an important role in the prevention of the epidemic, this paper suggests that traditional measures of the returns to education might be too low. In appendix 2, I estimate that, in the context of the HIV/AIDS epidemic in Uganda, by introducing the effect of education on health, an upper bound for the additional returns might be between 0.5 and 3.5 percentage points.

The paper is structured as follows: Section 2 gives a short background on the HIV/AIDS

epidemic in Africa. Section 3 describes the data set and the nature of the HIV/AIDS information campaign in the surveyed area. Section 4 exposes the empirical analysis of HIV prevalence and HIV incidence. Section 5 covers the analysis of sexual behavior. Section 6 discusses the issue of the causality between education and HIV and section 7 concludes.

2 Background on the HIV/AIDS Epidemic in Africa

The transmission of the HIV epidemic in Africa occurs overwhelmingly through heterosexual intercourse. There is a long interval between HIV infection—seroconversion—and the actual development of AIDS. Using the data from the Natural History Cohort, a nested cohort study within the GPC, it has been estimated that, for adults, the median time from seroconversion to AIDS was 9.4 years and 9.2 months from AIDS to death (Morgan and Whitworth 2001; Morgan and others 2002). This median survival time is similar to what was found in developed countries before the introduction of antiretroviral therapies. For the majority of HIV/AIDS patients in Africa, however, and certainly for the periods considered in this paper, antiretroviral therapies were neither available nor affordable.

The long interval between seroconversion and AIDS has several implications that should be kept in mind when studying the dynamics of the HIV/AIDS epidemic. The first implication is that most HIV positive individuals appear healthy. In an environment where voluntary testing is limited,² this means that most people ignore their own HIV status, as well as the status of their sexual partners. Another implication of this long interval is that, when the epidemic is in steady-state, any prevalence estimates reflect behavior which is on average approximately five years old. This is an important fact to take into account when trying to determine the extent to which some prevalence rates could have been influenced by information campaigns. For example, prevalence rates in 1989–1990 reflect behaviors that are on average five years old, while the first prevention campaigns were launched in 1986 at a national level and were, at a local level, reinforced by the establishment of the MRC Programme in November 1989 in the surveyed area. It is therefore reasonable to assume that a substantial component of the prevalence in 1989–1990 is the consequence of sexual behavior that predated the availability of the information about the epidemic.

²In the surveyed area, at round 11, in 2000, 8.9 percent of the individuals older than 17 had obtained the results of an HIV test.

Another implication of the long interval between seroconversion and AIDS is that the level of the epidemic can be seen as a stock/flow issue. The prevalence rate measures the stock of seropositive individuals which is affected both by the incidence—the flow of individuals who seroconverted—and the mortality rates.³

This paper analyzes both HIV prevalence and HIV incidence. The GPC is the first study to report a significant decline in adult HIV incidence in rural Africa (Mbulaiteye and others 2002a; Whitworth and others 2002), indicating that the decline in HIV prevalence observed in Uganda since 1994–1995 is not merely the consequence of the mortality of the people previously infected but can be attributed to behavioral changes.

It is also important to note that, in light of the fact that the median interval between infection and death is around 10 years, AIDS mortality among young adults is not widespread and therefore HIV prevalence at young ages is only marginally affected by the mortality rate. This alleviates the concern raised by the "survivor bias", the bias, present at later ages, induced by the fact the individuals most at risk already died and are not observed.

The first AIDS case in Uganda was diagnosed in 1982. The HIV virus was probably introduced late in the 1970s or early in the 1980s (Republic of Uganda 2000; 1995). The Ugandan government was the first in Africa to recognize the extent of the epidemic (Morgan and Whitworth 2001). As early as 1986, the government set up a national AIDS Control Programme which has been in charge of disseminating the information about the HIV/AIDS epidemic. From a peak of 14 percent in the early 1990s, the prevalence rate in the adult population has now been brought down to an estimated 5 percent (UNAIDS 2003). Several studies have attributed this decline to a change in behavior associated with strong prevention campaigns (Kilian and others 1999; Republic of Uganda 2000; Whitworth and others 2002).

Uganda has a population of 25 million of which 15 percent is urban. In 2002, Uganda had a GDP per capita of US \$ 1,354 US in purchasing power parity (World Bank 2003).

³In the absence of migration, the relation between the prevalence of HIV, the incidence and the mortality can be summarized as such:

$$\dot{H} = (1 - H)\lambda_H - H\lambda_D$$

\dot{H} is the change in the prevalence rate at each period, typically one year. H is the HIV prevalence rate, λ_H is the incidence rate, i.e. the rate at which HIV negative individuals become HIV positive and λ_D is the mortality rate from AIDS for HIV positive individuals.

3 Data Description and Information Campaign

The analysis in this paper uses the data from the General Population Cohort study (GPC) of the Medical Research Council Programme on AIDS in Uganda. Since November 1989, this study surveys annually 15 villages in the Masaka District in Uganda.⁴ In 1999–2000, for the eleventh round, 10 villages have been added to the cohort. The villages are located in a rural area in southern Uganda where the main cash crops are bananas and coffee.

The GPC is an open cohort: every individual who reaches age 13 is enrolled, as well as the persons migrating in the survey area. Every year, the participants in the cohort answer a sociodemographic and a medical questionnaire. The medical questionnaire includes anonymous HIV testing.⁵

The coverage and compliance in the GPC have been very high. For example, there were 9,611 individuals in the resident adult population at round 12 (11/2000–09/2001): among them 6,594 (69 percent) were interviewed and 6,312 (65 percent) were bled and tested for HIV.

At least since round 5 (11/1993 – 9/1994), there has been a general decline in HIV prevalence among individuals older than 17.⁶

Questions about the educational achievement of the participants were included in rounds 1, 7, 8, 9 and 11. As evidenced in the summary statistics presented in table 1, there is a general trend toward higher educational achievement over the period, corresponding to the openings of several schools in the area in the 1990s. The proportion of individuals with no formal education has decreased substantially while the proportion with secondary education increased dramatically. It should be noted that since 1996, the Ugandan Government has introduced the Universal Primary Education (UPE) which pays for public schools at the primary level. However, tuition is charged in secondary schools.

The variables used in the analysis, for which summary statistics are reported in table 1, in addition to education level and HIV status, include gender, marital status, religion and a measure for wealth. The proxy for wealth included is the type of materials used to construct the house.⁷

⁴See Mulder and others (1994); Mulder and others (1995); Nunn and others (1994); Whitworth and others (2002); Nunn and others (1997).

⁵The MRC program also offers voluntary testing and counselling services organized independently of the GPC.

⁶HIV prevalence is calculated taking each round as a separate cross-section and does not involve interpolations.

⁷The distinction is made between houses made of soft materials (mud or wattle with grass/papyrus roof), a

That information was available at rounds 1, 8 and 12.

Questions about sexual behaviors and AIDS related practices have been included in the medical questionnaires since round 4 (11/1992 – 9/1993) and have been repeated in several subsequent rounds. They generally include information about the number of sexual partners, the age at first sexual intercourse, condom use and visits to the voluntary counselling and testing centers.

To understand the evolution of the HIV prevalence over time and across education and age groups, it is important to have in mind the history of the HIV/AIDS epidemic in Uganda. The first case of AIDS was diagnosed in 1982. In 1986, the first prevention efforts on a national level were launched. The health education messages in Uganda have been mostly carried by governmental and nongovernmental sources and in the media (private local radio station and newspapers) (Kamali and others 2003). The arrival of the MRC project in the surveyed area in November 1989 might have increased the AIDS awareness in the population followed in the General Population Cohort.⁸ However, except for the distribution of condoms and the provision of voluntary counselling and testing for HIV, the GPC is an observational cohort, with no specific interventions. A comparison of continuously (from 1989) and newly (from 2001) surveyed villages has demonstrated the generalizability of the results and the absence of an Hawthorne effect (Mbulaiteye and others 2002b).

One logical question is whether the intensity of the exposure to the information campaigns varies with schooling. HIV/AIDS specific education programs in schools were only introduced in 1996, as a pilot program for a later nationwide application (Kinsman and others 2001). Given that the analysis in this paper is based on individuals aged 18 and above, only few individuals could have been in school when these sex education programs were in place and be included in the different samples. This fact ensures that there is no systematic differences in the treatment received by educated and less educated individuals.⁹ Actually, the existence of a yearly survey on AIDS,

combination of both (generally mud or wattle wall with iron sheets roof) or hard materials only (walls in bricks and iron sheet roof).

⁸Another channel through which individuals become aware of the epidemic might simply be that, as mortality starts to take its toll, people see their friends and relatives die. However, to obtain the specific knowledge about the transmission modes and the way to avoid AIDS, information campaigns are crucial.

⁹This point is further reinforced by the fact that during the pilot study, from 1996 to 1998, 95 percent of the schools taught the module “basic information about HIV/AIDS”, 83 percent of schools taught the module “responsible behavior: delaying sex”, but only 33 percent of schools taught “responsible behavior: protected sex” (Kinsman and others 2001). Teachers were reluctant to teach specifics about condom use because they were afraid to be seen as “immoral”.

including HIV testing, is a factor that suggests a widespread dissemination of the information in the surveyed area.

As all surveyed villages were exposed to the information campaigns, one limitation of the analysis is that there is no control group that allows a comparison between treated and untreated individuals. In order to contrast different levels of exposure to the information, this paper will rely instead on a distinction, based on the birth cohort of the individuals and the period of observation, between those who started their sexual life before the arrival of the information and those who initiated their sexual activity afterward. Further, the data are cut by education categories.

4 The Empirical Analysis of HIV Prevalence and HIV Incidence

The empirical analysis in this paper will take the following steps. After a short overview of the trends in the association between education levels and HIV prevalence, I will compare HIV prevalence at rounds 1 and 11—the last one to include both HIV tests results and the education variable—in order to assess whether the reaction to the information campaigns differed by education groups. Given the existence of a "survivor bias" (the people most likely to be at risk might have already died) among older individuals, this comparison will focus on young individuals.

The next step is the analysis of HIV incidence, a more accurate measure of the current situation of the epidemic, in a duration framework, using a proportional hazard model. Finally, sexual behavior and AIDS related practices will be investigated in order to understand the mechanisms by which the information campaigns affected HIV status.

The analysis is limited to individuals aged 18 and above, so that their education level is relatively stable.¹⁰ Individuals are first grouped in 3 categories: no formal education, at least some primary education and at least some secondary education or more. Figure 2 illustrates the trends in HIV prevalence among individuals aged 18 to 29 for these three categories from rounds 1 to 12. It is apparent that, initially (round 1), there is no particular gradient between education and HIV prevalence for younger individuals, but that later, education is associated with a lower risk of infection.

¹⁰Given that the education variable is only available at rounds 1, 7, 8, 9 and 11, for the other rounds, the measure of education is derived from adjacent rounds. I am therefore restricting the analysis to individuals aged 18 and above so that the education level is more likely to be constant over time.

4.1 Comparing HIV Prevalence at Rounds 1 and 11

The data set of the General Population Cohort, which has collected data over 12 years, allows to make comparisons to estimate the impact of the information about the epidemic on HIV prevalence. The question that is central to this paper is how this impact might have differed with the level of education. In order to answer this question, this paper will start by comparing HIV prevalence among young individuals at rounds 1 and 11.

The intuition behind this comparison is that it allows a contrast, at identical ages, between individuals who have been exposed to the information campaigns and individuals who have not. Using age 15 as the potential starting point of the sexual activity,¹¹ individuals born before 1971 started their sexual life before the prevention campaigns. On the other hand, individuals born after 1971 had more access to information about the epidemic when they initiated their sexual activities. This separation in two distinct age groups will be used extensively in the next steps of the analysis. At round 1, from November 1989 to September 1990, except for a few individuals aged 18, everybody belongs to the group of individuals who have started their sexual life before the launching of the prevention campaigns. At round 11, from November 1999 to September 2000, the age group 18-29 is composed of people who have started their sexual life after the prevention campaigns. Moreover, because the average period between HIV infection and death is around 10 years (Morgan and others 2002), HIV prevalence reflects behavior that is on average five years old. It is, therefore, reasonable to think that the prevalence at round 1 for young individuals reflects preinformation behavior, whereas HIV prevalence for the same age group at round 11 corresponds to postinformation behavior. In addition, this comparison holds the age group of the individuals constant, eliminating potential age profile effects and is only marginally affected by mortality and the "survivor" bias. The assumption behind the strategy of comparing individuals who have been exposed to the information campaigns and individuals who have not is that, except for the HIV/AIDS information campaigns, there is no other period specific effect.

The comparison between the two young age groups at rounds 1 and 11 is illustrated in figure 1: compared with individuals of the same age group at round 1, at round 11, the prevalence for individuals with no education decreased by 6 percentage points, by 5.9 percentage points for those

¹¹Average age of first sexual activity is between 16 and 17. However, many individuals started earlier. But less than 10 percent started before age 15. I therefore chose 15 as the "potential" starting point of the sexual life.

with primary education and by 12 percentage points for those with secondary education¹². This estimation gives a measure of the interaction between the effect of the arrival of the information about the HIV/AIDS epidemic and the effect of education, showing that the more educated the individuals—in particular, for individuals with some secondary education—the more strongly their prevalence rate decreased after the arrival of that information.

Table 1, by analyzing the relation between education and HIV prevalence in a regression framework confirms the point made in figure 1. The sample has been divided in two age groups: 18–29 and 30 and above. As explained, at round 11,¹³ this distinction can be interpreted as, on the one hand, the group that started their sexual life after the beginning of the prevention campaigns, and, on the other hand, the group that initiated their sexual activity before the information was available.

At round 1, there is no significant association between HIV and schooling.¹⁴ For the younger group, a higher educational achievement leads to a decrease in the probability to be infected by HIV. The point estimates for the marginal effects imply that, compared to the baseline of no education, having some primary education decreases the risk by 5.1 percentage points and some secondary education, by 8.8 percentage points. The statistical significance of the coefficient for primary education is, however, not very high. For the group of individuals older than 29, the sign on the coefficients on the education variables is positive (even though significance is not very high), in contrast with the results for the younger group. It should be kept in mind, however, that any measure of HIV prevalence for individuals older than 30 is bound to be affected by the "survivor"

¹²The differences across education categories among individuals aged 18 to 29 are not significant at round 1, while they are at round 11: HIV prevalence is 10.81 percent (confidence interval 5.71-18.12) for individuals with no education, 7.81 percent (CI 6.25-9.61) for individuals with at least primary education and 3.23 percent (CI 1.82 - 5.28) for individuals with at least some secondary education. Another cut of the data at round 11 shows that the completion of primary education seems important: no education, 10.81 percent (CI 5.71-18.12), some primary education but not completed, 9.4 percent (CI 7.33-11.83), completed primary, 4.62 percent (CI 2.66-7.40), at least some secondary 3.33 percent (CI 1.81-5.71) and above secondary 2.53 percent (CI 0.30-8.84). Table 2 proposes different ways to measure schooling levels.

¹³I have chosen to present the analysis at round 11 rather than round 12, because the education variable was not measured at round 12.

¹⁴A recent review of the epidemiological literature (Hargreaves and Glynn 2002) concluded that, in Africa, several studies had shown that higher education levels were associated with a greater risk of infection. In the data set of the GPC, there is a positive correlation between raw means at round 1. However, this is explained by an age or cohort effect: young individuals are more likely to have some secondary education and, at the same time, they are more likely to be HIV positive. With the introduction of age dummies in the regressions at round 1, the positive association between HIV infection and schooling disappears.

bias.¹⁵

Table 2 presents a stronger test that there has been a reversal of the education/HIV infection gradient between the two age groups. The regressions are based on the entire sample of individuals aged 18 and above. Interaction terms between the two cohorts and the education categories have been included instead of the education dummies. All the other covariates (gender, marital status, religion and wealth) are also interacted with the two age groups.¹⁶ The results confirm that there is no association between HIV prevalence and education at round 1 (see first column) while at round 11 (see second column) the coefficients on the education variables are of opposite signs between the two age groups, negative (and significantly so for secondary education) for the individuals younger than 30 at round 11 and positive for older individuals. The breakdown by gender is done, for round 11, in columns 3 and 4, and shows clearly that the negative gradient between education and HIV is only present among females.

An important issue that needs to be addressed in the analysis of the evolution across age groups of the education/HIV infection gradient is that the distribution of education levels between age groups varies. The summary statistics in table 1 show how, between the two cohorts considered, there are important differences in the distribution of the education categories.

One way to address this question is to consider the educational achievement of each individual relative to his gender and his birth cohort, rather than the absolute education level. Another alternative is to simply use the number of years of education as the measure of schooling achievement and enter it linearly, assuming that one additional year of education brings the same benefit at each point in the education distribution. The regressions in the last two columns of table 2 implement these alternatives and compare them with the use of absolute education levels categories. The fifth column uses the years of education linearly, whereas the last column decomposes the sample in three thirds depending on the place that each individual takes in the education distribution pertaining to his birth cohort (grouped in five years) and his gender. Otherwise, the specification is identical

¹⁵The results reported in table 1 document other interesting changes, beyond the scope of this paper, that took place in the studied population over the period considered: over time, young Muslims, and especially young Protestants have become less at risk of HIV infection, as compared to Catholics. Also, wealthier people, as measured by the quality of their housing, have become less likely to be HIV positive.

¹⁶This was necessary since the results in table 1 indicate that the coefficients on some covariates switched sign between the two age groups. In particular, the comparison between the regressions for the two different age groups reveals that males are less at risk than women at young ages but more at risk after age 29. This can be explained both by physiological and behavioral factors: at each exposure to the HIV virus, women are more likely to be infected; women start their sexual activity earlier, but men continue later.

across all columns. The conclusion is very similar in the three alternatives: when interacted with the dummy for young individuals, the higher the education level, the lower the risk. For example, in the fifth column, the coefficient indicates that for individuals aged 18 to 29, one additional year of education decreases HIV prevalence by 0.5 percentage point.

4.2 HIV Incidence and Survival Analysis

So far, this paper has focused on HIV prevalence, the "stock" of individuals infected at each round. However, HIV prevalence is affected by the incidence rate, the mortality rate, as well as by migration or attrition from the sample. As explained previously, it also reflects, given the 10 years period on average between HIV infection and death, sexual behavior that is on average five years old.

HIV incidence is a better measure of the evolution of the epidemic. An incident case is a individual that was first tested as HIV negative and later turned HIV positive. The General Population Cohort, by its design, allows to measure incidence, since the same individuals are followed over time. The GPC is the first study to have reported a significant decline in adult HIV incidence in rural Africa (Mbulaiteye and others 2002a; Whitworth and others 2002). A difficulty with the study of HIV incidence is that the number of incident cases is generally very low: given the 10 year span between infection and death, incidence rates are on average ten times lower than prevalence rates. Between round 1 and round 12, 280 incident cases can be detected in the GPC. There is a substantial attrition rate in the cohort, as transition rates from one round to the next one ranged between 66.7 percent (between rounds 3 and 4) and 79.2 percent (between rounds 8 and 9). The causes for attrition might be death, migration or noncompliance. It is also worth remembering that individuals enter in the cohort as they turn 13 or if they are new to the study area. At round 11, 10 new villages were added.

Given that the education variable is only available at rounds 1, 7, 8, 9 and 11, I am restricting the survival analysis to individuals aged 18 and above so that the education level is more likely to be constant over time.¹⁷

¹⁷When there was discordance across rounds in the educational level reported, I adopted the following rules: I first chose the level reported after age 18, if there was still discordance, I chose the level most often reported. Most of the discordances were for years of education inside the same education category (no education, primary and secondary). I found 107 observations (1 percent of the sample) with seriously discordant levels of education reported : in that case, I assigned the first level reported after age 18. I have run the analysis with and without these doubtful observations with no differences in the conclusion.

Following Meyer (1990) and Finkelstein and Poterba (2000), I use a semiparametric proportional hazard model based on Cox's partial likelihood technique (Cox 1972; 1975). The duration measure is the length of time between entry in the cohort (or age 18 if the individual entered the cohort before age 18) and exit for individuals who remained HIV negative or incidence in case of seroconversions. The failure event is seroconversion from HIV negative to HIV positive. The timing of incidence is assumed to be at mid-point between the last HIV negative and the first HIV positive test result available.

The hazard for individual i at time t is assumed to be:

$$h_i(t) = h_0(t) \exp(x_i' \beta)$$

where $h_0(t)$ is the baseline hazard at time t , which is left unestimated, x_i is the vector of explanatory variables for individual i and β is the vector of parameters to be estimated. The implicit assumption behind this proportional hazard model is that the baseline hazard has the same shape for everybody and that covariates like education, gender and age would affect it proportionally.

It is possible, however, to allow for the baseline hazard to take a different shape for different groups of individuals, by stratifying the estimation. In that case, the hazard for individual i at time t takes the form:

$$h_i(t) = h_{01}(t) \exp(x_i' \beta) \quad \text{if individual } i \text{ is in group 1}$$

$$h_i(t) = h_{02}(t) \exp(x_i' \beta) \quad \text{if individual } i \text{ is in group 2}$$

Table 3 presents results where the data have been stratified by gender, by age and also by round of entry in the cohort and by length of the observation period (follow-up as defined by the span between the last HIV test and the first one).¹⁸

The reason for stratifying by length of the follow-up is that there is some evidence of nonrandom attrition in the cohort: controlling for their age, their gender and their point of entry in the cohort, more educated individuals are more likely to stay longer under observation. Given that individuals enter or exit the cohort at different rounds, the model allows for left and right truncation.

¹⁸I have estimated the model with and without stratifying for the length of the observation period. The results were very similar under both specifications.

Table 3 indicates that if we consider all individuals, the level of education has no effect on HIV incidence (the hazard ratio is equal to 1). However, even if the data are stratified by age, gender, round of entry and length of follow-up to allow for different shapes of the baseline hazard, one of the features of the proportional hazard model is to assume that the effect of education is the same for both genders, at all ages and for any interval in the study period.

The remainder of table 3 addresses these issues by comparing estimations made separately for males and females, for old and younger individuals and for the first and last rounds of the survey. Significant results are obtained.¹⁹ Education decreases the risk for young people: one additional year of education decreases the hazard by 6.7 percent (see first column). But schooling increases the hazard for old people. This finding is consistent with the results obtained earlier in analyzing HIV prevalence. It is also important to realize that it is with young people that concerns about the "survivor bias" are minimized, since among older people, it might be the case that the individuals the most at risk already died. The second and the third columns indicates that a large component of the difference between young and old individuals, is actually due to a difference between old and young females. The contrast between the fourth and the fifth columns indicates that it is during the 1996–2001 period that the negative association between HIV incidence and education became significant.

Although, at first sight the survival analysis seems to indicate that there is no effect of education on HIV incidence, for young individuals education does decrease the risk of becoming HIV positive, especially for females and for the 1996–2001 period. Given that HIV incidence measures current behavior, one might expect that, after the arrival of the information, the effect of education would be the same for everybody. However, the theoretical model presented in appendix 1 is consistent with the fact that education has more effect for young individuals, since they face a longer time-span during which they can enjoy the returns of their human capital investment. Another way to explain the differences between young and old individuals is that older individuals have formed sexual habits before the HIV/AIDS epidemic and that it is more difficult for them to change their behavior. A model including habit formation and assuming that it takes time for individuals to modify their practices would also explain why most of the effect of education is concentrated in the 1996–2001 period. The next section that analyzes sexual practices will attempt to investigate how

¹⁹It should be reminded that the small number of incident cases makes it difficult to achieve statistical significance.

the effect of education is concentrated among young individuals.

The stronger and more significant effect of education on HIV incidence for females is reinforcing the similar finding obtained for HIV prevalence. An interesting direction to explore to investigate this differential might be the fact that in addition to the other channels through which education can affect the risk of getting HIV/AIDS, education might increase the bargaining position of the women in the sexual partnership. It has been documented that, generally, the great difficulty for wives to impose the use of the condom to their husband is a major obstacle in the prevention efforts (Republic of Uganda 1995/1996).

5 Can the Difference in HIV Outcomes across Education Levels Be Explained by Different Sexual Behaviors?

The empirical analysis developed in this paper shows that more educated individuals, once the information about the epidemic becomes available, are less at risk of being infected by the HIV virus. This section describes the mechanisms through which they reacted to the arrival of the information about the epidemic and how they did change their behavior, in particular their sexual behavior, to decrease their exposure to the virus.

From round 4 onward, the medical questionnaires have included several questions about sexual behaviors and attitudes toward HIV/AIDS. Figure 3 describes, for each round at which the question was included, the fraction of individuals who used a condom during their last sexual intercourse, giving a good measure of the regularity of condom use. Clearly, individuals with secondary education are more likely to use a condom on a regular basis.

Table 4 verifies that the education gradient reported in figure 3 is robust to the inclusion of controls for gender, age, village, marital status, religion and wealth. The analysis is performed at round 11.

The first row of results indicates that the finding that educated individuals are more likely to have ever used a condom and to have used a condom during their last sexual intercourse is robust and significant. Other behaviors and attitudes are also investigated in table 4. More educated individuals are more likely to have visited an AIDS counseling center and to have obtained the results from an HIV test. They are not likely to have significantly more or fewer sexual partners than others. However, if they have multiple partners, they are more likely to use a condom. They

are also more likely to start their sexual life at a later age. However, this differential pre-existed the HIV/AIDS epidemic, since it can be found among individuals born before 1971 (older than 15 in 1986, when the information campaigns where started).²⁰

Since the analysis of HIV incidence has shown that the effect of education in reducing the risk of becoming HIV positive was only present among young individuals, and in particular among females, the two last rows of table 4, as well as table 5, further scrutinize sexual behaviors by birth cohort and gender. For both age groups (born after 1970 and born before 1971), there is a positive and significant relation between education and the probability to have ever used a condom and the likelihood to have used a condom during the last sexual intercourse. However this positive gradient is substantially steeper for individuals born after 1970. The levels of condom use are also much higher among young individuals. These results are informative in order to understand why it is only among young individuals that education has an effect in reducing HIV incidence.

For both genders, condom use is positively related to educational levels. However, there is a difference in the sign of the education gradient for the number of partners. Among females, education reduces the number of partners, while more educated men tend to have more sexual partners, in particular when they are older than 29. This difference might explain that the effect of education in reducing HIV incidence is stronger for females.

Looking at sexual behavior confirms that more educated individuals have been more responsive to the threat of the HIV/AIDS epidemic. Condom use seems to be a mechanism where the differences by education levels are very pronounced. In addition, for females, the number of partners is negatively associated with schooling. The combination of higher use of condom and a reduced number of partners might explain why the effect of education in reducing HIV prevalence and incidence is concentrated among females.

The fact that this analysis is based on self-reporting might cast some doubts since more educated persons are more likely to know the "good answer". However, since their answers are confirmed by the lower HIV prevalence and incidence associated with education described previously, this gives credence to the analysis. The studies of the HIV outcomes and of the sexual behaviors are mutually reinforcing in establishing that more educated individuals have been more responsive to the HIV/AIDS campaigns and in documenting the behavioral mechanisms involved.

²⁰In general, it does not seem that the initiation of the sexual activity has been modified with the arrival of the epidemic: younger individuals started slightly earlier, on average, than older ones.

6 Why Is Education Associated with a Lower Risk of HIV Infection after the Diffusion of the Information?

This paper has shown that, once the information about the existence and the transmission modes of the HIV/AIDS epidemic was available, more educated individuals have reacted more to that information and are less likely to be infected by the HIV virus. The theoretical appendix 1 proposes three potential explanations for this finding: education is an investment for the future and therefore gives stronger incentives to individuals to protect their health, education enters as a factor in the health production function by giving better access to information and by helping to process that information and, finally, health and education are both caused by a third underlying variable, an unobservable like ability or the discount factor.

The evidence presented in this paper does not allow to distinguish between these three channels and therefore to determine whether the effect of education in decreasing the risk of HIV infection is causal. However, some findings suggest that education affects the risk of HIV through more than one channel.

In order to distinguish between the first and the second explanations, I have introduced in the regressions a variable describing the type of housing of the family as a proxy for wealth. The idea is to control for the effect of the permanent income, generally associated with the education level. The fact that, for individuals who have been exposed to the information, the negative gradient between HIV infection and education is robust to the inclusion of this proxy for wealth suggests that the role of education in helping to process the information is an important element of the explanation. The advantage that education gives in facilitating the access to and the understanding of the information is not negligible in communities where illiteracy is still sizeable and where access to the general media is not widespread.

However, the fact that, in the regressions, the coefficients on the indicators for the types of housing associated with higher wealth are negative and significant for young individuals at round 11 (see table 1) suggests that the explanation based on the complementarity between investments in health and education is also relevant. Education can be seen as an investment in the future that will deliver higher wealth. Individuals who have made this investment are willing to protect it by taking greater care of their health and changing their behavior in order to avoid to be infected by HIV/AIDS.

I have been looking for a way to separate the two first theoretical explanations of the link between education and HIV/AIDS from the third explanation that states that the correlation between education and health is not causal but due to common unobservables like the discount factor or the ability. To distinguish this explanation from the two other ones, one would need some exogenous variation in the level of education. If choices about education are not entirely endogenous but are also determined by some exogenous factor, like the structure of the school system, then the contribution of education in the health decision can be estimated without the selection bias due to unobservables. The facts that there was no secondary school in the surveyed area before 1982, that there was only one between 1982 and the 1990s and there are now five secondary schools all located close to the center of the area suggests that an identification strategy based on the availability of secondary schools, given the birth cohort and the location of the household, might be feasible. I tried to implement this approach, but the first stage regressions were not significant, maybe due to the small number of observations, or due to the fact that all the surveyed villages are in a three miles radius of the center implying that differences in distance to schools are not very important.

The question of whether education is causal in reducing the risk of being infected by HIV/AIDS is an important one for policy purposes. This paper does not ignore this issue but does not provide a definitive answer. Elsewhere, I have addressed this question in studying the relationship between smoking behaviors and education (De Walque 2003; 2004). Using data from the U.S. and the exposure to the Vietnam draft as an instrument for college education, I suggest, in an instrumental variable strategy, that at least part of the effect of education on the smoking decision is causal. Similarly, Currie and Moretti (2003) use the openings of two and four year colleges in the United States to establish that higher maternal education improves infant health and in particular reduces smoking during the pregnancy.

7 Conclusions and Orientations for Further Research

This paper exploits the arrival of information about the existence and transmission modes of the HIV/AIDS epidemic and investigates how different education groups reacted to this new information.

Using individual data from the General Population Cohort of the Medical Research Council Programme on AIDS in Uganda, this paper analyzes how the gradient between education and HIV

infection has been evolving over time.

At the initial round in 1989–1990, there is no robust or significant association between HIV prevalence and educational attainment. This is in sharp contrast with the situation 10 years later. For young individuals aged 18–29 at round 11—individuals who started their sexual life after the start of the prevention campaigns—there is a significant and robust negative gradient between education level and the risk of being HIV positive. This effect is concentrated among females. This finding suggests that more educated individuals reacted more to the arrival of information about the dangers associated with HIV/AIDS. A comparison between young individuals early in the 1990s whose sexual behaviors were not affected by the information about the epidemic and young individuals at the end of the same decade whose sexual life started after the arrival of this information, reinforces this conclusion by revealing that, over the decade, the HIV prevalence has decreased twice as much among more educated individuals, indicating a stronger reaction since the diffusion of the information.

Results from a study of HIV incidence in a duration framework confirm this finding by showing that, for young individuals, education reduces the probability of seroconversion. Moreover, this paper documents some of the mechanisms leading to this result by showing that changes in behavior and sexual practices, in particular condom use and visits to voluntary counselling and testing centers, have been more widespread among educated individuals.

The evidence from the variations in HIV prevalence and in HIV incidence and from the analysis of sexual behavior converges to establish that more educated individuals, especially among females, have been more responsive to the dissemination of the information about the HIV/AIDS epidemic. The responsiveness to the information is thought to be one of the channels through which education affects health. This paper can be read as a test of the importance of this channel.

In further research, it would be interesting to try to explain why the effect of education in reducing the risk of HIV infection is stronger among females than males. In this regard, the hypothesis that education increases the bargaining position of the women in the household deserves to be tested.

The results of this paper also have policy implications for future HIV prevention efforts. By showing that the very individuals that economic theory would predict to display the strongest behavioral response to the HIV/AIDS epidemic—educated individuals—are the ones who actually did, the analysis also reinforces the point that the declines in HIV prevalence and incidence observed

in Uganda at the end of the 1990s are due to behavioral changes and therefore that information campaigns can be effective. On the other hand, the conclusion of this paper that the information campaigns were more effective among the educated population raises some redistributive issues and indicates that other methods might have to be designed to reach the less educated group of the population.

One of the questions that also motivates this paper is whether the effect of education on health adds substantially to the standard estimates of the returns to education, which are based only on labor market outcomes. The impact of education on health can increase the returns to schooling by extending longevity and therefore the length of the period during which the benefits of the human capital investment are enjoyed. The large impact of AIDS on mortality and life-expectancy, combined with the effect of education in decreasing HIV prevalence, suggests, as discussed in appendix 2, that returns to education as traditionally calculated might be substantially underestimated in the context of a country confronted with a generalized HIV/AIDS epidemic like Uganda. However, if the effect of education in reducing the risk of HIV/AIDS is not entirely causal, the addition to the traditional returns to schooling is overestimated and should be considered as an upper bound.

Whether or not education is causal in reducing the risk of getting AIDS, this paper demonstrates that educated individuals have been more responsive to information campaigns about the epidemic. In another study, I have shown how the current negative gradient between education and smoking prevalence only appeared in the United States after the information about the dangers associated with smoking was disseminated in the 1950s and 1960s (De Walque 2003; 2004). These two studies display strong complementarities and indicate that the effect of education in improving health operates, in a substantial way, through the access to information and the responsiveness to it.

APPENDIX 1. A THEORETICAL MODEL OF SEXUAL BEHAVIOR IN THE PRESENCE OF THE HIV/AIDS EPIDEMIC

The question of whether the effect of education is causal is central in the economics literature on the interaction between health and education. Theoretical explanations for this correlation can be classified into three broad categories. One explanation emphasizes that education is an investment (Becker 1993). Education will deliver an higher income, an higher consumption level in the future, raising the value of staying alive. More educated individuals are healthier because their investment in the future gives them stronger incentives to protect their health. Another explanation (Grossman 1972; Kenkel 1991), based on education entering as a factor in the health production function, stresses that education improves the access to health related information and helps processing it. A third theory (Farell and Fuchs 1986; Fuchs 1982) states that the observed correlation between health and education is essentially due to third unobservable factors, like the discount factor or the ability that causes the same individuals both to study longer and to take greater care of their health. This appendix presents three variations, based on the three theoretical explanations described above, of a model that sketches choices in sexual behavior during the HIV/AIDS epidemic. These three channels are not mutually exclusive, however, and it is very likely that they all contribute to the explanation of the positive correlation between health and education. However, for the purpose of clarity, I have developed my argumentation in three steps.

1. Baseline model: in the absence of information about the HIV/AIDS epidemic

The model presented here is very simple but could be easily extended. It is a two periods model.

At each period, individuals derive utility from consumption goods (c_t) and the number of sexual partners (n_t) they have in each period, with $t = 1, 2$. For convenience, I make the following assumptions, which could be relaxed without changing the main results : the agents choose only the number of their sexual partners and they have the same number of sexual encounters with each partner.

The period utility function is separable in consumption and the number of sexual partners.

$$U(c, n) = u(c) + v(n) \tag{1}$$

where $u(\cdot)$ and $v(\cdot)$ are both increasing and concave in their respective arguments.

Agents live two periods provided they survive from the first to the second. Their survival is determined by a survival probability Q where $0 \leq Q \leq 1$. In the baseline model, I assume that agents are not aware of the risks associated with the HIV/AIDS epidemic. Therefore, they take \bar{Q} , the average survival probability in the population, as given and do not consider the effects of HIV exposure on survival when deciding about their sexual behavior.

In order to introduce variation in the level of education, I introduce two type of agents: these with low human capital K_L and those with high human capital K_H .

The wage $W(K_i)$ is an increasing function of the level of human capital K_i with $i = L, H$.

The price of consumption goods is normalized to one in both periods. Individuals have the opportunity to protect their sexual intercourse, for example, by using condoms. π_1 denotes the proportion of protected sexual intercourse in the first period. The price of protection (price of condoms, of HIV testing, control of partner's fidelity) of sexual intercourse is defined as p_π . The "price", p_n (market price in case of prostitution, gifts, dowry or shadow cost in case of non commercial sex) of a partner is identical for each of them.

The asset level of the individual is defined by A . I assume a perfect annuity market so that assets A in period 1 carried to the next period become $\frac{A(1+r)}{Q}$, where r is the interest rate. For convenience, I will let the discount factor, $\beta = \frac{1}{1+r}$.

Agents choose for each period their level of consumption (c_1, c_2), their number of sexual partners (n_1, n_2) and the proportions (π_1, π_2) of protected sexual intercourse.

The intertemporal problem they are facing is therefore characterized by:

$$MAX_{c_1, c_2, n_1, n_2, \pi_1, \pi_2} \quad u(c_1) + v(n_1) + \beta Q[u(c_2) + v(n_2)] \quad (2)$$

subject to the following budget constraint with $i = L, H$: $[\lambda]$

$$c_1 + (p_n + p_\pi \pi_1)n_1 + \frac{Q}{1+r}[c_2 + (p_n + p_\pi \pi_2)n_2] = W(K_i)(1 + \frac{Q}{1+r}) \quad (3)$$

and subject to

$$[\theta] \quad n_1 \geq 0; \quad [\psi] \quad \pi_1 \geq 0; \quad [\phi] \quad \pi_1 \leq 1 \quad (4)$$

The first order conditions for a maximum for this problem are very standard and include:

$$[c_1] : \quad u'(c_1) = \lambda \quad (5)$$

$$[\pi_1] : \psi = \phi + \lambda p_\pi n_1 \quad (6)$$

$$[n_1] : v'(n_1) + \theta = \lambda(p_n + p_\pi \pi_1) \quad (7)$$

From equation (6), it follows that, since $\lambda p_\pi n_1 > 0$ (unless $n_1 = 0$, in which case π_1 is also zero), for the condition to hold, we have $\psi > 0$ and $\phi = 0$, i.e. $\pi_1 = 0$, no protection is used for any sexual intercourse. This is true both for low (K_L) and high (K_H) human capital individuals and is expected, since none are aware of the dangers of the epidemic.

Combining (7) with (5) yields, (assuming at least one sexual partner, i.e. $n_1 > 0$, $\theta = 0$ and remembering that, as shown above, $\pi_1 = 0$):

$$v'(n_1) = u'(c_1)p_n \quad (8)$$

If we now introduce the difference between high and low human capital individuals (K_H, K_L), using that:

$$W(K_H) > W(K_L) \quad (9)$$

and the normality of consumption goods, we have that:

$$c_1^H > c_1^L \quad (10)$$

This simply states that more educated people consume more goods because they earn higher wages. Then, using the concavity of both terms of the utility function $u(\cdot)$ and $v(\cdot)$, yields

$$u'(c_1^H) < u'(c_1^L) \quad (11)$$

which using (8) implies:

$$v'(n_1^H) < v'(n_1^L) \quad (12)$$

From the concavity of $v(\cdot)$, it then follows that, in the absence of the HIV/AIDS epidemic or if individuals are not aware of its existence, "sexual adventure" is a normal good, i.e. more educated people, because they earn higher wages, have more sexual partners:

$$n_1^H > n_1^L \quad (13)$$

2. Model 2: Introducing the exposure to HIV virus in the survival probability

Once individuals have access to the information about the existence and the dangers of the HIV/AIDS epidemic, their sexual behavior takes into account the effect of their possible contact with the HIV virus on their survival probability. I introduce this in the baseline model by letting the surviving probability depend negatively on the exposure of the individual to the HIV virus. The exposure to the virus is determined, in a multiplicative way, by the number of sexual partners, the proportion of infected people in the sexually active population and the fraction of the sexual encounters that were protected.

$$Q(\gamma_1 n_1 (1 - \pi_1)) \quad (14)$$

where $Q(\cdot)$, the survival probability, is decreasing, n_1 is, as described previously, the number of sexual partners in period 1, γ_1 is the proportion of infected people in the sexually active population during the first period and π_1 is the proportion of protected sexual intercourse.

In accordance with the fact that there is a long period (10 years on average (Morgan and others 2002) between HIV infection and the apparition of the AIDS symptoms and that HIV testing is not widespread in Africa, I assume that it is not possible, both for the agent himself and for his potential sexual partners to know whether he is infected or not. The only information potentially available is the overall proportion of infected people in the population. Then, every agent acts as if his sexual partner has a probability to be infected equal to the proportion of HIV positive, i.e. the prevalence of HIV. In other words, γ_1 is taken as given. π_1 incorporates the possibility to avoid

HIV infection by using condoms or a combination of HIV testing and subsequent reciprocal fidelity.

The multiplicative feature of the function $\gamma_1 n_1 (1 - \pi_1)$ measuring the exposure to the HIV virus is important : if the individual abstains from sex ($n_1 = 0$), if nobody is infected by the HIV virus in the pool of potential sexual partners ($\gamma_1 = 0$) or if he has only protected sexual intercourse ($\pi_1 = 1$), then exposure to the HIV virus is zero.

The maximization problem faced by the agents is the same as in the baseline model with the only modification that \bar{Q} is replaced by $Q(\gamma_1 n_1 (1 - \pi_1))$

The first order conditions for c_1 , c_2 are similar to the ones in the baseline model. The first-order conditions for n_1 and π_1 are modified:

[n_1] :

$$v'(n_1) + \beta Q'(\gamma_1 n_1 (1 - \pi_1)) \{ \gamma_1 (1 - \pi_1) \} [u(c_2) + v(n_2) + \lambda \left(\frac{A(1+r)}{Q(\gamma_1 n_1 (1 - \pi_1))} \right)] - \lambda(p_n + p_\pi \pi_1) + \theta = 0 \quad (15)$$

[π_1] :

$$\psi - \phi - \beta Q'(\gamma_1 n_1 (1 - \pi_1)) \{ \gamma_1 n_1 \} [u(c_2) + v(n_2) + \lambda \left(\frac{A(1+r)}{Q(\gamma_1 n_1 (1 - \pi_1))} \right)] - \lambda p_\pi n_1 = 0 \quad (16)$$

It is worth mentioning in (15) and (16) that the term $\beta Q'(\gamma_1 n_1 (1 - \pi_1))$ is negative (the survival probability is decreasing in the exposure to the HIV virus).

The model now assumes that individuals are aware of the existence of the HIV/AIDS epidemic and that a positive proportion, $0 < \gamma_1 \leq 1$, of the pool of potential sexual partners is infected by the HIV virus. Under which conditions would the agents decide, for a given number of sexual partners n_1 , to have all their sexual intercourse protected, i.e. $\pi_1 = 1$?

Continuing with the analysis of the Kuhn-Tucker conditions, $\pi_1 = 1$ implies $\phi > 0$ and $\psi = 0$. Then, from (16), we have:

$$-\beta Q'(0) \{ \gamma_1 n_1 \} [u(c_2) + v(n_2) + \lambda \left(\frac{A(1+r)}{Q(0)} \right)] > \lambda p_\pi n_1 \quad (17)$$

Keeping in mind that $Q'(0)$ is negative and using again equation (5), this can be rewritten as:

$$\frac{[u(c_2) + v(n_2)]}{u'(c_1)} + \frac{A(1+r)}{Q(0)} < \frac{p_\pi}{-\beta Q'(0)\gamma_1} \quad (18)$$

For more educated individuals, we expect, by normality, c_2 , n_2^{21} , and A , the assets carried from one period to the other, to be larger. Using also equation (11), this gives us that the left-hand side of equation (18) is higher for educated people. For the inequality in (18) to hold, one, therefore, needs $\tilde{\gamma}_1$, defined as the cut-off level at which agents would decide to protect all their sexual intercourse, to be lower for educated individuals, i.e.:

$$\tilde{\gamma}_1^H < \tilde{\gamma}_1^L \quad (19)$$

This result indicates that, holding the number of sexual partners n_1 constant, more educated people will decide to protect all their sexual intercourse ($\pi_1 = 1$), at a lower level of HIV infection in the general sexually active population.

A similar analysis can be conducted on equation (15), holding π_1 and γ_1 constant. At an interior solution for n_1^{22} , equation (6) can be rewritten as:

$$\frac{v'(n_1)}{u'(c_1)} + \beta Q'(\gamma_1 n_1 (1 - \pi_1)) \{ \gamma_1 (1 - \pi_1) \} \left[\frac{[u(c_2) + v(n_2)]}{u'(c_1)} + \left(\frac{A(1+r)}{Q(\gamma_1 n_1 (1 - \pi_1))} \right) \right] = p_n + p_\pi \pi_1 \quad (20)$$

The right-hand side of equation (20) is constant. The terms $\frac{[u(c_2) + v(n_2)]}{u'(c_1)}$ and A are larger for more educated individuals as previously described. Keeping in mind that $Q'(\cdot)$ is negative and assuming for simplicity that the survival probability decreases linearly in the exposure to the HIV virus, i.e. $Q''(\cdot) = 0$, we need, in order to keep the equality with the constant term in the right-hand side, the term $\frac{v'(n_1)}{u'(c_1)}$ to be larger. The inequality (11) already contributes to this, however a lower number of sexual partners for more educated people would further help keeping equation (20) hold with equality. This would increase, because of the concavity, $v'(n_1)$ and would increase $Q(\gamma_1 n_1 (1 - \pi_1))$, therefore reducing, in absolute value, the second term of the left-hand side which

²¹Even in the presence of the HIV/AIDS epidemic, it is reasonable to expect $n_2^H > n_2^L$, i.e. that more educated individuals have more sexual partners in the second period, because this is the last period and there are no gains to be made in terms of survival.

²²The model could impose that $n_1 > 0$ by imposing an Inada condition of the type $v'(0) = \infty$, but this is not necessary for the analysis.

is negative. This gives the following result which says that, in the presence of the HIV/AIDS epidemic, at a constant level of protection π_1 and a constant level of infection in the population γ_1 , more educated people tend to have less sexual partners:

$$n_1^H < n_1^L \quad (21)$$

This is the opposite of the result obtained in the absence of the epidemic.

Of course those exercises in comparative statistics analyze extreme cases. It is more likely that the responses will be mixed : more educated individuals might reduce more the number of their sexual partners but also use greater protection. The results derived above in this simplified model show how, in the presence of the HIV/AIDS epidemic, more educated individuals will modify more strongly their behavior.

In the present model, the way by which higher education induces a stronger behavioral change in the presence of HIV/AIDS is twofold, illustrating two of the mechanisms that can explain that more educated people modify more strongly their behavior when confronted to the HIV/AIDS epidemic. The first channel operates through the fact that higher education leads to a higher level of utility in the second period giving more incentives to avoid a decrease in the survival probability. This is illustrated by the fact that the terms $\frac{u(c_2)+v(n_2)}{u'(c_1)}$ and A in equations (18) and (20) are larger for more educated individuals, increasing the shadow cost of unprotected sex with several partners. This points to the fact that more educated people take greater care of their health since they have invested more in their future.

The model also includes a second way through which the complementarity between health and education is operating: more educated people have an advantage in accessing or processing the information about the existence, the transmission modes and the ways to prevent the HIV/AIDS epidemic. Because more educated individuals were the first to understand the information about the epidemic, they were the first to optimize their behavior according to model 2, while less educated groups were still ignoring or not understanding the fact that unprotected sex was risky (baseline model).

3. Model 3: Different discount factors

In order to account for the view initiated by Farrell and Fuchs (Farrell and Fuchs 1986; Fuchs

1982) that the correlation between education and health is due to a third unobservable variable like the discount factor or ability, I introduce in model 2, two different levels of discount factors, β_L and β_H with $\beta_L < \beta_H$. The set up is the same as in model 2, except that β_i is now indexed by $i = L, H$. The first order conditions for a maximum are similar with the following differences for the condition for n_1 and π_1 :

[n_1] :

$$v'(n_1) + \beta_i Q'(\gamma_1 n_1 (1 - \pi_1)) \{ \gamma_1 (1 - \pi_1) \} [u(c_2) + v(n_2) + \lambda \left(\frac{A(1+r)}{Q(\gamma_1 n_1 (1 - \pi_1))} \right)] - \lambda(p_n + p_\pi \pi_1) + \theta = 0 \quad (22)$$

[π_1] :

$$\psi - \phi - \beta_i Q'(\gamma_1 n_1 (1 - \pi_1)) \{ \gamma_1 n_1 \} [u(c_2) + v(n_2) + \lambda \left(\frac{A(1+r)}{Q(\gamma_1 n_1 (1 - \pi_1))} \right)] - \lambda p_\pi n_1 = 0 \quad (23)$$

For individuals with an higher discount factor β_H , and keeping in mind that $Q'(\cdot)$ is negative, it follows that the shadow cost of unprotected sex in equations (18) and (20) will be higher for individuals with an higher discount factor. Individuals with a high discount factor β_H will therefore decide to protect their sexual intercourse or restrict the number of their partners sooner in the epidemic, i.e. at a lower prevalence rate in the population γ_1 . Both the second and the third models predict that more educated individuals will be more responsive to the information about the HIV/AIDS epidemic²³.

²³An interesting extension of the model would be to endogenize the discount factor and use the complementarity between time preference and future utility, following Becker and Mulligan (1997). That might be a way to reconcile the hypotheses of Grossman (1972) and Farrell and Fuchs (1986).

APPENDIX 2. ESTIMATION OF THE ADDITIONAL RETURNS TO EDUCATION

One of the questions that motivates this paper is whether the effect of education on health could add substantially to the standard estimates of the returns to education, which are based only on labor markets outcomes. The impact of education on health can increase the returns to schooling by extending longevity and therefore the length of the period during which the benefits of the human capital investment are enjoyed. This impact can be counted as an additional return to education, however, only in as far as it results from a causal effect of education on health. This appendix proposes a way to estimate how, including mortality differentials due to the different responses to the HIV/AIDS epidemic in Uganda, by individuals from different education levels, might increase the returns to education.

Even though as discussed in section 6, it is not possible to disentangle the three theories proposed behind the health/education gradient, in the following calculations, I will, for the purpose of this exercise, consider all of the effect of education on health as being causal. This will yield upper bounds of the additional returns to schooling due to the effect of education in reducing HIV/AIDS infection, since the effect is not necessarily causal. On the other hand, AIDS is the leading cause of premature death for adults in Africa, but it is not the only one. Thus, by concentrating on the mortality differentials due to HIV/AIDS, those estimates will only take into account part of the impact of education on mortality. Similarly, they only consider the impact on life-expectancy and do not account for differences in morbidity.

The estimation of the rate of return to schooling is based on the model of the internal rate of return:

$$\sum_{t=0}^{\infty} \left(\frac{1}{1+r}\right)^t Q_t^P W_t^P = \sum_{t=6}^{\infty} \left(\frac{1}{1+r}\right)^t Q_t^S W_t^S - \sum_{t=0}^6 \left(\frac{1}{1+r}\right)^t Q_t^S C_t$$

Q_t^P and Q_t^S are the respective survival probability for individuals with primary and secondary education at time t (i.e. at age $t + 13$, since $t = 0$ is age 13). W_t^P and W_t^S are the income of individuals with primary and secondary education at time t and C_t represents the costs of going to secondary school. The rate of return to education is the rate r that equalizes the left hand side and the right hand side of the equation of the internal rate of return, i.e. the present discounted values of earnings at age 13 for individuals who went to primary and secondary school respectively.

As a benchmark, I evaluate the rate of return with the same mortality schedule for both categories ($Q_t^P = Q_t^S$). I have estimated the yearly income per worker by education level for males in rural Uganda, using the Uganda National Household Survey 1999/2000.²⁴ The costs of going to secondary school per pupil, including tuition, room and board and other supplies, have been estimated from the same survey. The age specific mortality rates, identical for both education levels, are based on estimates published in the reports of the DHS 1995 and 2000 (Republic of Uganda 1995 and 2000) for Uganda. The calculation yields a benchmark rate of return to education of 10.23 percent as reported in column A of table A1.

The next step is to introduce mortality differentials due to the fact that more educated individuals are less at risk of being infected by HIV/AIDS. Since I am looking at the prospect of a young individual in 1999/2000, I use as benchmark for the difference in HIV exposure the ratio of the HIV prevalence at ages 18–29 for individuals with primary and secondary education at round 11 (see figure 1). To estimate the timing of the death due to AIDS, I take the age-specific incidence rate estimated from the data from the GPC and add 10 years, the median survival between HIV infection and death (Morgan and others 2002). The difference in the mortality schedule for individuals with primary and secondary education obtained using this procedure is illustrated in figure A1 that depicts the survival curves from age 13 for both categories. Figure A1 shows that HIV/AIDS is a disease that affects individuals in their prime ages, since most of the effect of the education differential is concentrated between ages 30 and 60. The first line of column 2 of table A1 indicates that taking into account the mortality differential ($Q^P < Q^S$) across the two education groups due to their different HIV exposures increases the returns to education from 10.23 percent to 10.71 percent, an addition of 0.48 percentage point.

Those estimates, however, only account for the lost income stream due to HIV/AIDS mortality. Starting with Rosen (1988), the economic literature has stressed that the value of a statistical life exceeds the present discounted value of future income streams. The value of life should account for the utility derived from full income and full consumption, including the value of leisure. Different studies have estimated the value of a statistical life. I am using, as a starting point in the following calculations, two ranges of estimates for the United States: the value between 1 and 2 million dollars

²⁴In rural Uganda, most individuals do not earn wages as employees but work on the household farm. In such cases of household enterprises, I have used the family enterprise revenue divided by the number of full-time workers from the family in that enterprise as the measure of income per worker.

for the median driver/voter obtained by Ashenfelter and Greenstone (2004) using mandated speed limits and the median value of 5 million dollars used by Murphy and Topel (2001). From both estimates, I get the value of a life-year at each age (see figure 1a in Murphy and Topel 2001). Even though no estimates of the value of life in Uganda or in Africa are readily available, I have tried to account for the fact that the value of life is greater than the present discount value of income streams. In order to translate the U.S. estimates to Ugandan levels, I calculated the ratio between the value of a life-year and the income for both estimates in the U.S., and applied the same ratio to the Ugandan incomes.

The value of life in the model of the internal rate of return,²⁵ is introduced by adding to the earnings of high school and graduates at each age, the part of the value of a life-year that exceeds labor income. More precisely, I subtract from the value of a life-year at each age the average labor income at that age, to obtain an estimate:

$$\bar{\theta}_t = \text{full utility at age } t - \text{average labor income at age } t$$

For simplicity, and because it is difficult to estimate the skill premium outside the labor markets, I assume that $\bar{\theta}_t$ is the same for both education groups, and I add it, at each age, on both sides of equation of the internal rate of return.

The results, reported in table A1, vary depending on the estimates of the value of life. When including the value of life in the calculations, returns to education go from a benchmark of 10.23 percent to a range between 11.54 percent and 13.74 percent.

Counting only future income streams, the effect of education in reducing HIV/AIDS prevalence adds 0.48 percentage point to the standard estimates of the returns to education. When adding estimates of the value of life, the addition to the returns to schooling varies between 1.31 and 3.51 percentage points. These additions are substantial.²⁶ They indicate that, in the presence of an epidemic that kills early in life and in large proportions, the nonmarket, health related, returns to schooling can be important. However, it should be kept in mind that these calculations assume that the effect of education in reducing the risk of HIV infection is causal, a point that has not been conclusively established by this paper.

²⁵I have benefited from discussions with Fabian Lange on this subject.

²⁶They are orders of magnitude larger than what I calculated for the effect of education in reducing smoking prevalence in the U.S. In that case the addition was between 0.1 and 0.5 percentage point, See De Walque (2003)

An interesting issue is that the HIV/AIDS epidemic has two effects on the returns to education. The first is to lower drastically the overall life-expectancy with the consequences that incentives to invest in education should be reduced. But, the second effect of the epidemic is that more educated people react more to the information about the epidemic, and this tends to increase the returns to education. The calculations described above only considers that second effect, since it is the only one that is potentially causal. However, it would be interesting, in further research, to try to estimate and disentangle both effects.

References

The word “processed” describes informally reproduced works that may not be commonly available through library systems.

Ainsworth, Martha, Kathleen Beegle, and Godlike Koda. 2001. "The impact of adult mortality on primary school enrolment in Northwestern Tanzania." UNAIDS, Geneva. Processed.

Ashenfelter, Orley, and Michael Greenstone. 2004. "Using Mandated Speed Limits to Measure the Value of a Statistical Life." *The Journal of Political Economy* 112 (1,1): S226–68.

Becker, Gary S. 1993. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, 3rd Edition. Chicago, Ill: The University of Chicago Press.

Becker, Gary S., and Casey B. Mulligan. 1997. "The Endogenous Determination of Time Preference." *Quarterly Journal of Economics* 112(3): 729–58.

Blanc, Ann K. 2000. "The Relationship Between Sexual Behavior and Level of Education in Developing Countries." Report prepared for UNAIDS, Geneva. Processed.

Chamberlain, Gary. 1982. "Multivariate Regression Models for Panel Data." *Journal of Econometrics*. 18: 5–46.

Clark, Robert C., and Desire Vencatachellum. 2003. "Economic Development and HIV/AIDS Prevalence." Scientific Series 2003s–25, Centre interuniversitaire de recherche en analyse des organisations (Cirano), Montreal.

- Cox, D. R. 1972. "Regression Models and Life-Tables." *Journal of the Royal Statistical Society B* 34: 187–202.
- Cox, D.R. 1975. "Partial Likelihood." *Biometrika*, 62: 269–76.
- Currie, Janet, and Enrico Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings and Longitudinal Data." *Quarterly Journal of Economics* 118 (4): 1495–532.
- Deaton, Angus, and Christina Paxson. 2003. "Mortality, Income, and Income Inequality over time in Britain and the United States." In David A. Wise, ed., *Perspectives on The Economics of Aging*, vol. 8. Chicago. Ill.: University of Chicago Press for NBER.
- De Walque, Damien. 2003. "How Do Information and Education Affect Health Decisions? The Cases of HIV/AIDS and Smoking." Ph.D. Dissertation, University of Chicago. Chicago. Ill.
- De Walque, Damien. 2004. "How Do Education and Information Affect Smoking Decisions: Evidence from Smoking Histories (1940–2000) in the United States." Development Research Group, World Bank, Washington, D.C. Processed.
- Farell, Phillip, and Fuchs, Victor. 1986. "Schooling and Health: The Cigarette Connection." In Victor R. Fuchs, ed., *The Health Economy*. Cambridge, Mass., and London: Harvard University Press.
- Finkelstein, Amy, and James Poterba. 2004. "Adverse Selection in Insurance Markets: Policyholder evidence from the U.K. Annuity Market." *The Journal of Political Economy* 112 (1,1): 183–208.
- Fuchs, Victor. 1982. "Time Preference and Health: An Exploratory Study." In Victor Fuchs, ed., *Economic Aspects of Health*. Chicago, Ill.: University of Chicago Press.
- Fylkesnes, Knut, Rosemary M. Musonda, Moses Sichone, Zacchaeus Ndhlovu, Francis Tembo, and Mwaka Nonze. 2001. "Declining HIV prevalence and risk behaviours in Zambia: evidence from surveillance and population-based surveys." *AIDS* 15: 907–16.
- Fylkesnes, Knut, Rosemary Mubanga Musonda, Kelvin Kasumba, Zacchaeus Ndhlovu, Fred Mluanda, Lovemore Kaetano, and Chiluba C. Chipaila. 1997. "The HIV epidemic in Zambia: socio-

demographic prevalence patterns and indications of trends among childbearing women." *AIDS* 11: 339–45.

Garnett, Geoff P., Nicholas C. Grassly, and Simon Gregson. 2001. "AIDS: The Makings of A Development Disaster?" *Journal of International Development* 13: 391–409.

Gertler, Paul, Manisha Shah, and Stefano Bertozzi. 2003. "Sex Sells, But Risky Sex Sells for More." Processed.

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

Gregson, Simon, Nicola Terceira, Phyllis Mushati, Constance Nyamukapa, and Catherine Campbell. 2001. "School Education & Avoidance of Early HIV Infection: The Mediating Roles of Social Capital & Psychosocial Factors among Young Women in Rural Zimbabwe." Processed.

Grossman, Michael. 1972. "On the Concept of Health Capital and the Demand for Health." *The Journal of Political Economy* 80 (2): 223–55.

Hargreaves, James R., and Judith R. Glynn. 2002. "Educational attainment and HIV-1 infection in developing countries: a systematic review." *Tropical Medicine and International Health* 7 (6): 489–98.

Kamali, A., M. Quigley, J. Nakiyingi, J. Kinsman, J. Kengeya-Kayondo, R. Gopal, A. Ojwiya, P. Hughes, L. M. Carpenter, and J. Whitworth. 2003. "Syndromic management of sexually transmitted infections and behaviour change interventions on transmission of HIV-1 in rural Uganda: a community randomised trial." *The Lancet* 361: 645–52.

Kenkel, Donald S. 1991. "Health Behavior, Health Knowledge, and Schooling." *The Journal of Political Economy* 99 (2): 287–305.

Kiefer, Nicholas M. 1988. "Economic Duration Data and Hazard Functions." *Journal of Economic Literature* 26(2): 646–79.

Kilian, Albert H. D., Simon Gregson, Bannet Ndyabangi, Kenneth Walusaga, Walter Kipp, Gudrun Sahlmuller, Geoffrey P. Garnett, Godwill Asiime-Okiror, Geoffrey Kabagambe, Peter Weis,

- and Frank von Sonnenburg. 1999. "Reductions in risk behaviour provide the most consistent explanation for declining HIV-1 prevalence in Uganda." *AIDS* 13: 391–98.
- Kinsman J., J. Nakiyingi, A. Kamali, L. Carpenter, M. Quigley, R. Pool, and J. Whitworth J. 2001. "Evaluation of a comprehensive school-based AIDS education programme in rural Masaka, Uganda." *Health Education Research* 16 (1): 85–100.
- Kremer, Michael. 2001. "Creating Markets for New Vaccines, Part I: Rationale and Part II: Design Issues." In Adam B. Jaffe, Josh Lerner, and Scott Stern, eds., *Innovation Policy and the Economy*. Cambridge, Mass.: MIT Press, Volume 1.
- Lleras-Muney, Adriana. 2004. "The Relationship Between Education and Mortality in the United States." Forthcoming, *Review of Economic Studies*.
- Lundberg, Mattias, and Mead Over. 2000. "Transfers and Household Welfare in Kagera." Development Research Group, World Bank, Washington, D.C. Processed.
- Mbulaiteye, S. M., A. Ruberantwari, J. S.Nakiyingi, L. M. Carpenter, A. Kamali and J. A. G. Whitworth. 2000. "Alcohol and HIV: a study among sexually active adults in rural southwest Uganda." *International Journal of Epidemiology* 29: 911–15.
- Mbulaiteye S.M., C. Mahe, J.A.G. Whitworth, A. Ruberantwari, J.S. Nakiyingi, A. Ojwiya, and A. Kamali. 2002a. "Declining HIV-1 incidence and associated prevalence over a 10-year period in a rural population cohort in south-west Uganda." *The Lancet* 360: 41–46.
- Mbulaiteye S.M., C. Mahe, A. Ruberantwari, and J.A.G. Whitworth. 2002b. "Generalizability of population-based studies on AIDS: a comparison of newly and continuously surveyed villages in rural southwest Uganda." *International Journal of Epidemiology* 31: 961–67.
- Meyer, Bruce D. 1990. "Unemployment Insurance and Unemployment Spells." *Econometrica* 58 (4): 757–82.
- Morgan, Dilys, and Jimmy A. G. Whitworth. 2001. "The natural history of HIV-1 infection in Africa." *Nature Medicine* 7(2): 143–45.

- Morgan, Dilys, Cedric Mahe, Martin J. Okongo, Billy Mayanja, and J. A.G. Whitworth. 2001. "Genital Ulceration in Rural Uganda: Sexual Activity, Treatment-Seeking Behavior, and the Implications for HIV Control." *Sexually Transmitted Diseases* 28(8): 431–36.
- Morgan, Dilys, Cedric Mahe, Billy Mayanja, Martin J. Okongo, Rosemary Lubega, and James A.G. Whitworth. 2002. "HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries?" *AIDS* 6: 597–603.
- Mulder, Daan, Andrew Nunn, Anatoli Kamali, and Jane Kengeya-Kayondo. 1995. "Decreasing HIV-1 seroprevalence in a rural Ugandan cohort." *British Medical Journal* 311: 833–36.
- Mulder, Daan W., Andrew J. Nunn, Hans-Ulrich Wagner, Anatoli Kamali, and Jane-F. Kengeya-Kayondo. 1994. "HIV-1 incidence and HIV-1-associated mortality in a rural Ugandan population cohort." *AIDS* 8: 87–92.
- Murphy, Kevin M., and Topel, Robert. 2001. "The Economic Value of Medical Knowledge." Forthcoming in Robert H. Topel and Kevin M. Murphy, eds., *Exceptional Returns: The Economic Value of Medical Research*. Chicago, Ill.:University of Chicago Press.
- Nunn, Andrew J., Jane F. Kengeya-Kayondo, Sam S. Malamba, Janet A. Seeley, and Daan W. Mulder. 1994. "Risk factors for HIV-1 infection in adults in a rural Ugandan community: a population study." *AIDS* 8: 81–86.
- Nunn, Andrew J., Daan W. Mulder, Anatoli Kamali, Anthony Ruberantwari, Jane-Frances Kengeya-Kayondo, and Jimmy Whitworth. 1997. "Mortality associated with HIV-1 infection over five years in a rural Ugandan population: cohort study." *British Medical Journal* 315: 767–71.
- Over, A. Mead. 1997. *Confronting AIDS. Public Priorities in a Global Epidemic*. World Bank Report. Oxford: Oxford University Press.
- Philipson, Tomas J., and Richard A. Posner. 1993. *Private Choices and Public Health. The AIDS Epidemic in an Economic Perspective*. Cambridge, Mass.: Harvard University Press.
- Rao Vijayendra, Indrani Gupta, Michael Lokshin, and Smarajit Jana. 2003. "Sex Workers and the Cost of Safe Sex: The Compensating Differential for Condom Use among Calcutta Prostitutes." *Journal of Development Economics* 71: 585–603.

- Republic of Uganda and Macro International. 1995. "Demographic and Health Survey (DHS III)." Statistics Department, Ministry of Finance and Economic Planning, Entebbe, Uganda.
- Republic of Uganda and Macro International. 2000/01. "Demographic and Health Survey (DHS III)." Statistics Department, Ministry of Finance and Economic Planning, Entebbe, Uganda.
- Republic of Uganda and Macro International. 1995/96. "Demographic and Health Survey (DHS III). In-Depth, Negotiating reproductive outcomes." Statistics Department, Ministry of Finance and Economic Planning, Entebbe, Uganda.
- Republic of Uganda. 1999/2000. "Uganda National Household Survey 1999/2000." Uganda Bureau of Statistics. Entebbe, Uganda.
- Rosen, Sherwin. 1988. "The Value of Changes in Life Expectancy." *Journal of Risk and Uncertainty* 1: 285–304.
- Seeley, Janet A., Sam S. Malamba, Andrew J. Nunn, Dann W. Mulder, Jane F. Kengeya-Kayondo, and Thomas G. Barton. 1994. "Socioeconomic Status, Gender, and Risk of HIV-1 Infection in a Rural Community in South West Uganda." *Medical Anthropology Quarterly* 8(1): 78–89.
- UNAIDS. 2003. "AIDS epidemic update: October, December 2003." Geneva. Processed.
- UNAIDS. 2003. "Epidemiological Fact Sheets (by Country) on HIV/AIDS and sexually transmitted infections, 1998, 2000 and 2003 Updates (revised)." Geneva. Processed.
- United States Government. 2000. "HIV/AIDS Surveillance Data Base, HIV/AIDS Profile: Uganda." Census Bureau, International Programs Center, Population Division. Washington, D.C.
- Vandemoortele, Jan, and Delamonica Enrique. 2000. "The 'education vaccine' against HIV." *Current Issues in Comparative Education* 3 (1). [Retrieved on March 19, 2004 from: <http://www.tc.columbia.edu/CICE/articles/jved131.htm>].
- World Bank. 2003. "Development Indicators." Washington, D.C.
- Whitworth, J. A.G., C. Mahe, S. M. Mbulaiteye, J. S. Nakiyingi, A. Ruberantwari, A. Ojwiya, and A. Kamali. 2002. "HIV-1 epidemic trends in rural south-west Uganda over a 10-year period." *Tropical Medicine and International Health* 7(12): 1047–52.

Table 1: Medical Research Council, General Population Cohort, Rural Uganda.

Regressions and summary statistics at Rounds 1 and 11

Logit estimation: Mean of Marginal Effects, observation by observation

Dependent variable: HIV positive =1, HIV negative = 0

	Round 1 (11/89-11/90)				Round 11 (11/99-9/00)			
	Age 18 -29		Age 30 and above		Age 18 -29		Age 30 and above	
	Logit	Means	Logit	Means	Logit	Means	Logit	Means
HIV positive	Dep. Var	0.144	Dep. Var	0.077	Dep. Var	0.075	Dep. Var	0.085
		{0.010}		{0.005}		{0.006}		{0.005}
no education	omitted	0.133	omitted	0.408	omitted	0.068	omitted	0.247
		{0.010}		{0.011}		{0.006}		{0.008}
primary	-0.021	0.742	-0.008	0.504	-0.051	0.643	0.049	0.599
	[-0.76]	{0.013}	[-0.60]	{0.011}	[-1.39]	{0.011}	[1.40]	{0.009}
secondary	-0.012	0.124	0.006	0.087	-0.088	0.287	0.035	0.0153
	[-0.34]	{0.009}	[0.33]	{0.006}	[-1.89]	{0.011}	[1.10]	{0.007}
male	-0.011	0.445	0.04	0.468	-0.035	0.439	0.037	0.457
	[-0.52]	{0.014}	[1.81]	{0.011}	[-1.49]	{0.011}	[1.60]	{0.009}
single	omitted	0.399	omitted	0.124	omitted	0.46	omitted	0.045
		{0.014}		{0.007}		{0.012}		{0.003}
married	0.035	0.542	-0.44	0.638	0.037	0.45	0.001	0.604
	[1.24]	{0.014}	[-1.55]	{0.010}	[1.14]	{0.012}	[0.04]	{0.009}
previously married	0.034	0.057	0.021	0.237	0.074	0.086	0.085	0.35
	[0.74]	{0.007}	[1.14]	{0.009}	[1.84]	{0.006}	[1.83]	{0.009}
catholic	omitted	0.581	omitted	0.623	omitted	0.62	omitted	0.64
		{0.014}		{0.011}		{0.011}		{0.009}
protestant	0.002	0.111	0.001	0.144	-0.102	0.131	0	0.129
	[0.10]	{0.009}	[0.09]	{0.008}	[-2.14]	{0.008}	[-0.00]	{0.006}
muslim	-0.036	0.304	-0.04	0.22	-0.064	0.246	-0.023	0.224
	[-1.26]	{0.013}	[-1.62]	{0.009}	[-1.98]	{0.010}	[-1.07]	{0.008}
house soft materials	omitted	0.431	omitted	0.366	omitted	0.146	omitted	0.159
		{0.014}		{0.015}		{0.09}		{0.007}
house mixed materials	-0.036	0.38	-0.001	0.478	-0.039	0.292	-0.019	0.334
	[-1.56]	{0.014}	[-0.15]	{0.011}	[-1.43]	{0.012}	[-0.94]	{0.009}
house hard materials	-0.01	0.187	-0.0009	0.154	-0.075	0.56	-0.009	0.506
	[-0.38]	{0.011}	[-0.05]	{0.008}	[-2.58]	{0.013}	[-0.57]	{0.010}
N =	1052		1549		1156		1611	
Pseudo R-2	0.069		0.12		0.153		0.149	

Note: Robust z-stat in square brackets (regressions), standard errors in curly brackets (means).

Age and village dummies included.

Primary is some primary education (grades 1 to 7), secondary is some secondary education (grades 8 to 13) or above.

The type of housing materials (soft roof and house, mixed or hard roof and house) serves as proxy for wealth.

Type of housing is measured at Round 12 in the case of Round 11.

Table 2: HIV prevalence and education. General Population Cohort (Medical Research Council), Rural Uganda.

Round 1 (11/89 & 9/90) & Round 11 (11/99-9/2000), Age 18 and above

Dependent variable: HIV positive =1, HIV negative = 0

Logit estimation: Mean of Marginal Effects, observation by observation

Round 1 By education categories Both genders	Round 11 By education categories Both genders	Round 11 By education categories Males	Round 11 By education categories Females	Round 11 By years of education Both genders	Round 11 By rank in the education distribution Both genders
age 18-29	0.116	-0.084	0.262	age 18-29	0.16
	[1.12]	[-0.54]	[1.77]		[1.38]
primary age >29	-0.014	0.088	0.36	years of education	0.001
	[-0.75]	[1.98]	[1.30]	age > 29	[0.37]
secondary age>29	0.008	0.028	0.051	years of education	-0.005
	[0.29]	[0.59]	[1.57]	age 18-29	[-1.84]
primary age 18-29	-0.015	0.062	-0.037		2nd third, age 18-29
	[-0.68]	[0.89]	[-1.30]		[-0.64]
secondary age 18-29	-0.005	0.043	-0.078		3rd third, age 18-29
	[-0.19]	[0.54]	[-1.88]		[-1.75]
N =	2601	1028	1361		2861
Observed Probability	0.102	0.083	0.08		0.081
Pseudo R-2	0.093	0.178	0.151		0.141

Note: Robust z-stat in square brackets.

Controls for gender, marital status, religion and wealth are included and are interacted, like the education variables, with the dummies for the two age groups.

Age and village dummies included. The rank in the education distribution is calculated by gender and 5 years age group.

Primary is some primary education (grades 1 to 7), secondary is some secondary education (grades 8 to 13) or above.

The type of housing materials (soft roof and house, mixed or hard roof and house) serves as a proxy for wealth and is measured at Round 12, in the case of Round 11.

Table 3: HIV incidence analysis. Medical Research Council, General Population Cohort. Individuals aged > 17. By age, gender and period. Proportional Hazard Model (Cox Model).

Stratified by gender, by year of birth, by round of entry and length of follow-up:

	All	Males	Females	Rounds 1-6	Rounds 6-12
All					
	Hazard Ratio				
years of education	1.0005	0.99	1.013	1.019	0.978
Z- Stat	[0.02]	[-0.32]	[0.35]	[0.6]	[-0.65]
N	6405	2959	3446	3852	5434
Failures	230	124	106	127	103
Time at risk	33793	15887	17906	11685	19406
Born after 1970					
	Hazard Ratio				
years of education	0.933	0.959	0.918	0.988	0.88
Z- Stat	[-1.9]	[-0.68]	[-1.9]	[-0.21]	[-2.71]
N	2503	1153	1350	992	2246
Failures	82	33	49	40	42
Time at risk	9467	4712	4755	2702	6765
Born before 1971					
	Hazard Ratio				
years of education	1.052	1.002	1.16	1.04	1.069
Z- Stat	[1.66]	[0.08]	[2.56]	[0.96]	[1.47]
N	3902	1806	2096	2860	3188
Failures	148	91	57	87	61
Time at risk	24326	11175	13151	11685	12641

Note: Robust z-stat in square brackets.

Based on individuals present at least 2 rounds and who were HIV negative at the first round.

Failure is seroconversion from HIV - to HIV +

Education is measured at age 18. In case of discordance, the most often cited grade is taken, otherwise the one mentioned just after age 18

Table 4: Sexual behavior and AIDS related practices.

Round 11 (11/99-9/2000), MRC, General Population Cohort. Individuals aged 18 and above: By age group.

Logit: Mean of Marginal Effects, observation by observation

Dependent variable:		Ever used a condom?	Used condom last time	Number partners last year if active (*)	Ever visited VCT	Ever got result test	Age at first sex (*)	Used condom last time if multiple partners partners
All	Years of education	0.021 [10.10]	0.012 [5.30]	0.002 [0.36]	0.008 [3.58]	0.006 [2.86]	0.17 [9.19]	0.024 [3.77]
	N =	3253	2719	2816	3811	3811	3040	336
	Mean	0.256	0.087	1.33	0.113	0.089	16.83	0.214
	R-square	0.22	0.25	0.12	0.06	0.07	0.18	0.24
Age 18-29	Years of education	0.026 [6.75]	0.016 [4.99]	0.004 [0.28]	0.01 [3.39]	0.007 [2.66]	0.11 [4.89]	0.025 [3.01]
	N =	1340	1293	1197	1505	1505	1218	237
	Mean	0.424	0.176	1.419	0.113	0.086	16.43	0.324
	R-square	0.153	0.24	0.13	0.09	0.1	0.11	0.24
Age 30 and above	Years of education	0.012 [4.5]	0.008 [1.84]	0.009 [1.58]	0.012 [2.40]	0.007 [1.71]	0.18 [6.87]	** **
	N =	1913	1288	1619	2272	2207	1822	**
	Mean	0.16	0.037	1.265	0.112	0.091	17.09	0.107
	R-square	0.21	0.15	0.14	0.06	0.07	0.22	**

Note: Age and village dummies, controls for gender, marital status, religion and wealth included.

Robust z-stat in parentheses. (*) = linear regression, ** sample too small for meaningful regression.

Age dummies reduce slightly the sample size, as at old ages, the outcome is predicted perfectly.

"Did you ever used a condom" was asked to all individuals who ever had sex,

while condom use at the last intercourse was asked to all individuals sexually active in the last year.

**Table 5: Sexual behavior and AIDS related practices, by gender and age group
Round 11 (11/99-9/2000), Medical Research Council, General Population Cohort.**

Individuals aged 18 and above.

Logit: Mean of Marginal Effects, observation by observation

Dependent variable:		Ever used a condom?	Used condom last intercourse?	Number partners last year if active (*)	Age at first sex (*)
Males	Years of education	0.023 [6.86]	0.013 [3.38]	0.022 [1.56]	0.09 [2.71]
	N =	1419	1095	1338	1337
	Mean	0.351	0.126	1.651	17.6
	R-square	0.19	0.21	0.11	0.2
Males Age 18-29	Years of education	0.028 [4.37]	0.015 [3.39]	0.012 [0.32]	0.064 [1.74]
	N =	578	551	494	529
	Mean	0.578	0.272	1.971	16.67
	R-square	0.09	0.19	0.11	0.18
Males Age 30 and above	Years of education	0.017 [2.97]	0.013 [1.25]	0.024 [2.24]	0.113 [2.41]
	N =	841	380	844	808
	Mean	0.225	0.048	1.457	18.23
	R-square	0.2	0.14	0.17	0.18
Females	Years of education	0.022 [5.61]	0.018 [4.02]	-0.006 [-2.29]	0.2 [9.79]
	N =	1598	1328	1478	1703
	Mean	0.182	0.057	1.04	16.22
	R-square	0.17	0.3	0.07	0.16
Females Age 18-29	Years of education	0.023 [4.18]	0.016 [3.00]	-0.009 [-1.62]	0.156 [5.53]
	N =	762	723	703	689
	Mean	0.313	0.109	1.054	16.26
	R-square	0.16	0.33	0.07	0.15
Females Age 30 and above	Years of education	0.014 [2.27]	0.006 [1.49]	-0.004 [-1.64]	0.219 [8.32]
	N =	809	464	775	1014
	Mean	0.098	0.025	1.028	16.19
	R-square	0.16	0.24	0.08	0.2

Note: Age and village dummies, controls for marital status, religion and wealth included.

Robust z-stat in parentheses. (*) = linear regression.

Age dummies reduce slightly the sample size, as at old ages, the outcome is predicted perfectly.

"Did you ever used a condom" was asked to all individuals who ever had sex,

while condom use at the last intercourse was asked to all individuals sexually active in the last year.

Table A1: Additional returns to schooling from the effect of education in reducing the risk of HIV infection for males in rural Uganda..

	Same mortality schedule for primary and secondary education	Lower mortality schedule for secondary education	Extra return to schooling
	A	B	= B-A
Future income stream only	10.23%	10.71%	0.48%
Add non-income part of the value of life			
Ashenfelter, Greenstone	10.23%	11.54%	1.31%
Murphy, Topel	10.23%	13.74%	3.51%

Based on:

- Income schedules by age and education, and cost of secondary education per pupil estimated from the Uganda National Household Survey, 1999/2000.
- HIV prevalence in 1999/2000 by education category and HIV incidence by age group estimated from the data of the General Population Cohort of the Medical Research Council Program on AIDS in Uganda.
- Age specific mortality rates for Uganda published in the final reports of the Demographic and Health Surveys 1995 and 2000 for Uganda.

Figure 1: HIV prevalence among young individuals (aged 18-29) at Rounds 1 and 11, by education category: General Population Cohort, MRC Programme, Rural Uganda.

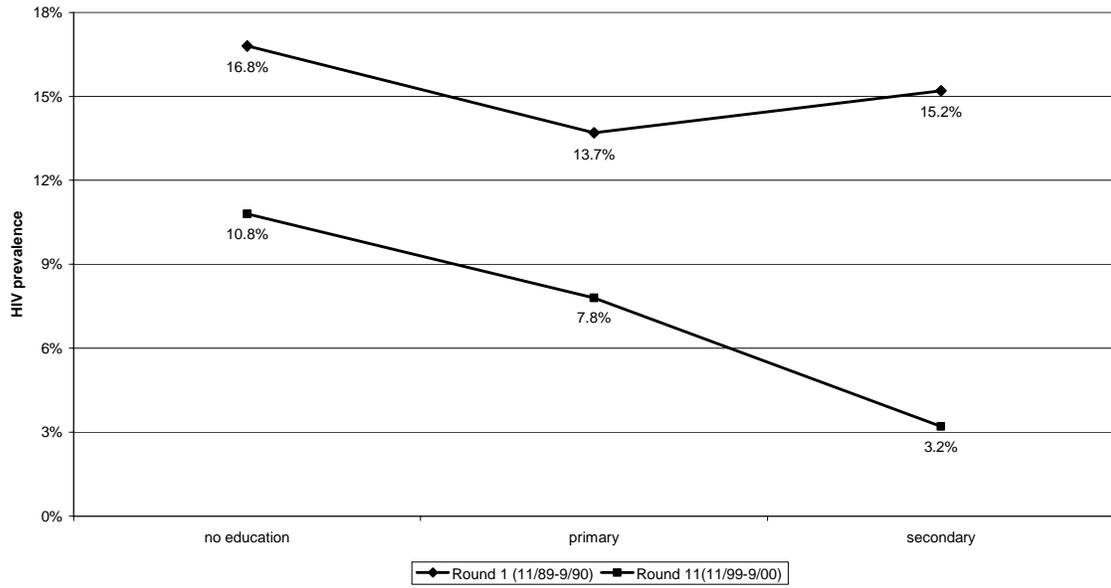


Figure 2: HIV prevalence by education category, MRC General Population Cohort, Rural Uganda, 1990-2001. Individuals aged 18-29

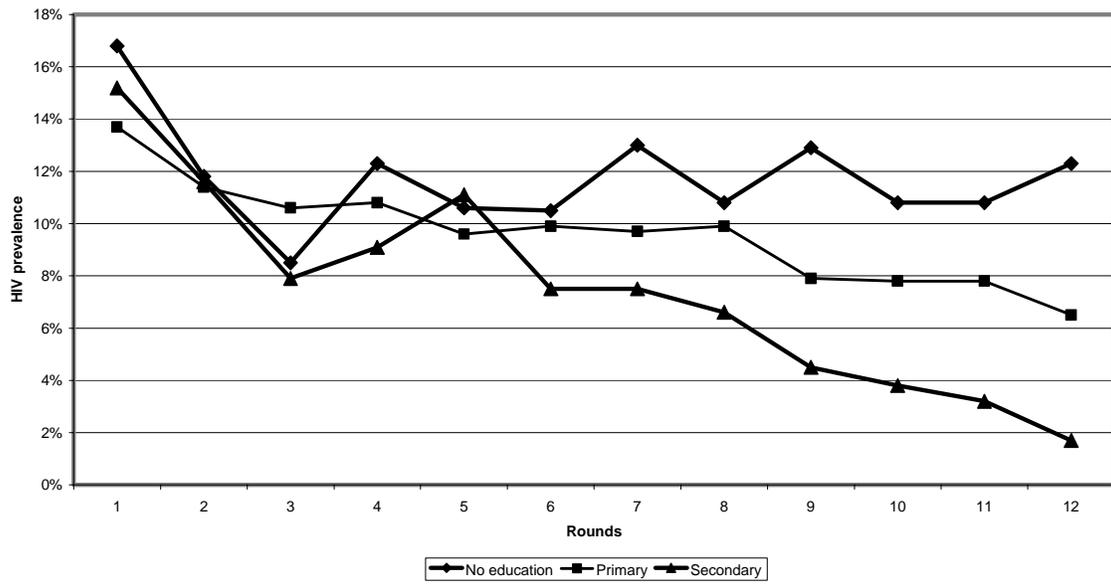


Figure 3: Percentage of individuals older than 17 who used a condom during their last sexual intercourse, by education category at several rounds of the General Population Cohort (MRC), Rural Uganda.

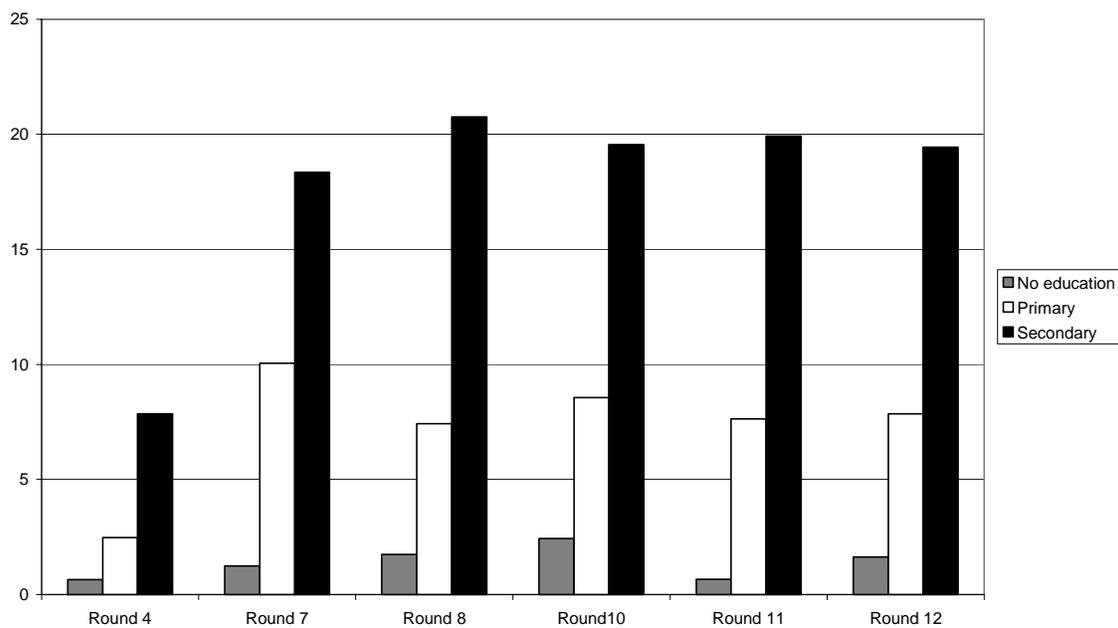


Figure A1: Rural Uganda, post AIDS epidemic, survival curve from age 13, by education category.
(Source: MRC General Population Cohort and Demographic and Health Survey, Uganda, 1995 and 2000)

