DO TEENAGERS RESPOND TO HIV RISK INFORMATION? EVIDENCE FROM A FIELD EXPERIMENT IN KENYA

Pascaline Dupas

Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya
Pascaline Dupas
NBER Working Paper No. 14707
February 2009
JEL No. C93,I1,O12

<u>ABSTRACT</u>

I use a randomized experiment to test whether information can change sexual behavior among teenagers in Kenya. Providing information on the relative risk of HIV infection by partner's age led to a 28% decrease in teen pregnancy, an objective proxy for the incidence of unprotected sex. Self-reported sexual behavior data suggests substitution away from older (riskier) partners and towards protected sex with same-age partners. In contrast, the national abstinence-only HIV education curriculum had no impact on teen pregnancy. These results suggest that teenagers are responsive to risk information but their sexual behavior is more elastic on the intensive than on the extensive margin.

Pascaline Dupas
Department of Economics
UCLA
8283 Bunche Hall
Los Angeles, CA 90095
and NBER
pdupas@econ.ucla.edu

1   Introduction

Nearly 2 million people become infected with HIV/AIDS every year in sub-Saharan Africa, the great majority of them through sex, and a quarter of them before the age of 25. AIDS is incurable and no succesful AIDS vaccine has been developed yet. Thus ensuring the adoption of safer sexual behavior among youth remains critical to combating the disease. Now that the great majority of children in Africa acquire at least some primary education, some have argued that primary schools offer a unique opportunity to deliver HIV prevention education to youths before they become sexually active (Bundy, 2002). There is, however, considerable debate over whether scalable school-based HIV/AIDS education programs can be effective in limiting the spread of HIV/AIDS among youths, and over what should be the content of these programs. Many sub-Saharan African countries have incorporated HIV/AIDS education in their school curriculum, but the great majority of those curricula are limited to *risk avoidance* information: they aim at completely eliminating pre-marital sex, by promoting abstinence until marriage. They omit to provide *risk reduction* information, for example that condom use reduces the risk of HIV transmission.

Voluntarily limiting information so that youths are unaware of the "low risk" option (e.g. condom-protected sex) and only face the choice set {high risk; no risk} might be socially optimal since individuals do not internalize the epidemiological externalities of their own behavior, and therefore what might be optimal at the individual level (a low but non-zero risk level) might be suboptimal from a public health perspective. However, if sexual behavior is more elastic on the intensive margin (what type of sex to have and with whom) than on the extensive margin (whether to have sex or not), HIV education programs that focus only on abstinence may be ignoring an important margin along which youths could reduce their risk of infection.

This paper uses data from a field experiment to measure the responsiveness of teenagers to HIV information and compares their responses along both the risk avoidance and the risk reduction margins. The risk reduction margin I focus on is *partner selection.* Partner selection is an important risk reduction margin for teenagers in sub-Saharan Africa, where the prevalence of HIV is at least three times higher among teenage girls than among teenage

boys[1]. Multiple studies have suggested that this discrepancy is due, in part, to the high incidence of unsafe cross-generational sexual relationships — that is, unprotected sex between teenage girls and adult men more than five years their senior.[2,3] Men involved in these relationships, often called "sugar daddies", are more likely to be infected with HIV than teenage boys since they have been sexually active for longer. Thus, compared to relationships with teenage boys, cross-generational relationships pose a higher risk of HIV infection for teenage girls. On the other hand, older men, who typically have more income, are usually better able to provide for the teenage girl and the baby if the sexual relationship leads to pregnancy. Since the distribution of income is more readily observable than the distribution of HIV infection, adult men may have an advantage over teenage boys in negotiating for unprotected sex. Most HIV prevention campaigns may not reduce this advantage, since they only provide information on the average HIV risk (the overall prevalence) and their key message is that "Anyone can give you HIV." Though true in essence, this message obscures the fact that in sub-Saharan Africa 25-year-old men are more likely to have HIV than 16-year-old teenagers.

In this context, providing teenage girls with full information on HIV prevalence disaggregated by gender and age groups may reduce the incidence of unprotected cross-generational partnerships, and along with it new HIV infections of young women by older partners. The total amount of sexual activity might increase, however, if teenage girls who learn that sex with teenage boys is relatively safe increase their sexual activity with teenage boys. This might have negative public health consequences, both in terms of teen pregnancies and in terms of lifetime HIV risk and its epidemiological implications (Magruder, 2007). A rigorous test of the impact of risk reduction information and how it compares with the impact of risk avoidance information is thus needed.

I use data from a randomized field experiment involving 328 primary schools to compare the effects of providing abstinence-only versus detailed HIV risk information on teenage

---

[1] For example, in Kenya and Zambia, prevalence in the 15-19 age group has been found to be at least five times higher among girls than among boys (Glynn et al. 2001).

[2] Laga et al. (2001), Luke and Kurz (2002), Gregson et al. (2002), Kelly et al. (2003), Clark (2004).

[3] The prevalence gap between young women and young men is also due to the fact that risk of male-to-female HIV transmission is greater than the risk of female-to-male transmission (Peterman et al. 1988), but this biological factor, accounts for only a third of the gap observed (Gregson et al. 2002).

sexual behavior. Half of the schools, randomly selected, received teacher training on the national HIV/AIDS curriculum, which focuses on abstinence until marriage, but does not discuss risk reduction strategies (such as condom use or selection of safer partners). In 71 schools, randomly selected after stratifying by teacher training status, an information campaign provided teenagers with information on the prevalence of HIV disaggregated by age and gender group (the "relative risks information" campaign). The randomized design ensured that there would be no systematic difference in the prior information held by the students across groups at the onset of the programs. This ensures rigorous identification of the impact of each of the two programs, by comparing behaviors and outcomes across groups over time.

I find that the teacher training on the national HIV/AIDS curriculum had no effect on the likelihood that teenage girls had started childbearing within a year, suggesting no reduction in risky behavior. In contrast, the relative risks information campaign led to a 28% decrease in the likelihood that girls had started childbearing within a year, suggesting an important decrease in the incidence of unprotected sex among those girls. I provide evidence that the pregnancies averted by the relative risks information campaign would have been with partners more than five years older, suggesting that the reduction in the incidence of unsafe sex corresponds to a reduction in unsafe cross-generational sex. The relative risks information campaign also led to an increase in self-reported sexual activity among teenage boys, suggesting that girls substituted away from older partners and towards their agemates. But there was no increase in pregnancies among teenage couples, consistent with the fact that teenage girls report higher rates of condom use, presumably in order to avoid pregnancy with resource-constrained teenagers.

Taken together, these results suggest that the behavioral choices of teenagers are not responsive to risk avoidance messages but are responsive to information on the relative riskiness of potential partners. Overall, the relative risks information campaign led to an increase in reported sexual activity, but to a decrease in unsafe sex. This suggests that teenage sexual behavior is more elastic on the margin of what type of sex to engage in — the choice of partner and the choice of protection level — than on the margin of whether to engage in sex or not. These results suggest that, in the fight against HIV, risk reduction

messages might be more effective than risk avoidance messages. This is in line with the non-experimental evidence that has recently emerged from a longitudinal study in the US (Trenholm et al., 2007).[4]

Prior evidence on the effectiveness of sexual health education in Africa is almost non-existant. Gallant and Maticka-Tyndale (2004) review 11 studies of HIV education programs conducted in sub-Saharan Africa, and show that the mixed results generated by those studies are questionable, either because they do not have a convincing identification strategy, or because they rely on self-reported measures of risk-taking, which are likely to suffer from social desirability biases. The only prior randomized controlled experiment in Africa that studied biological outcomes lacked statistical power to detect small or moderate effects and thus can only reject very large effects (Hayes et al. 2005). This paper overcomes these shortcomings by using a randomized experiment that involved a large sample of primary schools in Kenya, combined with data on teenage pregnancy, arguably an objective measure of the incidence of unprotected sex.

Overall, the findings of this paper suggest that teenagers are responsive to risk information. This is in line with studies of youth behavior in other contexts, which have found that youths are responsive to information on the returns to education (Jensen, 2007), responsive to prices (Gruber and Zinman, 2001, Pacula et al., 2001) and responsive to regulatory incentives (Dee and Evans, 2001, Levitt and Lochner, 2001).

The remainder of the paper is organized as follows. Section 2 describes the Kenyan context and the experimental design. Section 3 presents the theoretical motivation behind the experimental design. Section 4 presents the data and outlines the evaluation strategy. Section 5 presents the results, and Section 6 concludes.

---

[4]Previous studies in the US have yielded mixed evidence. Douglas Kirby (2008) reviews 56 US-based studies and finds that most abstinence programs do not delay initiation of sex, while two thirds of the more comprehensive programs (that include boh abstinence and risk reduction information) seem to have an impact on both margins (delay of initiation and incresed condom use).

2   Kenya Background and Experimental Design

The experiment took place in two rural districts of Western Kenya, and involved a total of 328 primary schools. Among them, 163 were randomly chosen for Treatment 1, the Teacher Training on the national HIV/AIDS curriculum for primary school, which focuses on abstinence-only messages. In addition, 71 schools were sampled to receive Treatment 2, the Relative Risks Information Campaign, that provided 8th graders with information on HIV prevalence by sex and age. Before describing these two treatments in detail, I present some background summary statistics on HIV and sexual behavior among youths in Kenya.

2.1   Kenya Background

2.1.1   HIV prevalence

The principal mode of transmission of HIV in Kenya is heterosexual contact (Baltazar et al. 2001). The 2003 Kenya Demographic and Health Survey estimated that 7% of Kenyan adults are infected with HIV (Central Bureau of Statistics, Kenya 2004). The breakdown by age and gender group is presented in Figure 1. The highest infection level is for women in the 25-29 age group (12.9%). Levels of infection among young women rise quickly (3% in the 15-19 age group and 9% in the 20-24 age group). In contrast, prevalence rises gradually with age among men, starting at 0.4% in the 15-19 age-group, rising to 2.4% in the 20-24 age-group, and reaching its peak (8.8%) in the 40-44 age group.

2.1.2   HIV-related knowledge and sexual behavior among teenagers

Table 1, Panel A presents summary statistics on HIV-related knowledge and behavior collected at baseline among a subset of students sampled for the experiment. At the time they filled the survey, the students were enrolled in 8th grade (the last grade of primary school). Due to high levels of repetition through primary school, the average age of students in 8th grade is much greater than in the US, at 15.1 for girls and 15.5 for boys.

Overall, 21% of girls and 48% of boys sampled for the survey reported they ever had sex. Knowledge of risk reduction strategies among these students was relatively limited. While 46% of girls and 71% of boys thought that condoms could protect from HIV infection, only

29% of girls and 25% of boys knew that older men were more likely to be HIV positive than teenage boys.

Panel B of Table 1 presents data collected at the end of the study among girls in the control group who had started childbearing within a year. The data suggests that the great majority (87%) of teen pregnancies are unplanned. The older the male partner, the more likely pregnancy triggers marriage. While the rate of marriage is 42% if the teenage girl gets pregnant with a partner less than 5 years older, it rises to 63% if the age difference between partners is 5 to 10 years, and to 79% if the age difference is greater than 10 years. In 73% of teen pregnancies, the partnership involves regular cash transfers from the male to the female partner prior to the pregnancy.

Overall, 45% of teenage pregnancies observed in the control group were with partners more than 5 years older, and 11% with partners more than 10 years older. The share of sexually active teenage girls engaging in sexual partnerships with men 5 or more years older might have been lower than that, however, since cross-generational partnerships may be more likely to result in a pregnancy than within-cohort partnerships. In a study conducted in western Kenya, Luke (2003) finds evidence that a larger age difference between partners is associated with both a lower probability of condom use and higher transfers from the male to the female partner.

2.1.3 National Policy on HIV Education

The Kenya Ministry of Education, Science and Technology (MoEST) integrated HIV/AIDS education into the primary school curriculum in 2001. The national HIV/AIDS curriculum includes information on the biology of HIV/AIDS, how it is transmitted, how to care for people living with AIDS, and the consequences of the HIV/AIDS epidemic on families, schools, and the nation at large. It also includes a prevention section which emphasizes moral values, refusal skills, and abstinence until marriage. The curriculum does not mention condoms and provides only limited scope for teachers to discuss protected sex in response to students' questions. It does not cover partner selection, and though it covers love relationships between same-age boys and girls, the official textbooks do not mention cross-generational relationships (and their associated risk). The proposed strategies to avoid infection are to "Avoid

Sex" and to "Say NO to sex before marriage".[5] All sexual activity outside of marriage, irrespective of the age of the partner, is thus considered equally risky.

While the HIV/AIDS curriculum was introduced in 2001, HIV/AIDS education was in effect largely absent from Kenyan primary schools by early 2003. Only few schools had included HIV/AIDS in their timetable, and when asked why they did not teach the mandated curriculum, schools and their teachers would often say they were not familiar with its content or did not know how to teach the topic. In response, the Kenya Institute of Education (KIE) and the MoEST trained a number of trainers to provide in-service training for teachers on HIV/AIDS education methodology. The training was being phased-in over a large period of time (starting in 2003), which allowed randomization.

2.2   Treatment 1: The Teacher Training Program (TT)

In 2003, 328 primary schools in Western Kenya were sampled for a randomized evaluation designed to test the effectiveness of the national HIV prevention curriculum for primary schools. After stratifying the 328 schools by location, test scores, and gender ratio, half of the schools were randomly chosen to receive Teacher Training (henceforth TT) on the HIV/AIDS curriculum in 2003.[6] The training was conducted jointly by one facilitator from the AIDS Control Unit of the Ministry of Education (ACU-MOEST), two facilitators from the Kenya Institute of Education (KIE), and one trained staff member from the non-governmental organization ICS. Three teachers by primary school were trained during multiple week-long in-service training sessions.

During the training, teachers discussed the material in the official HIV/AIDS curriculum and learned how to discuss HIV/AIDS issues in class. At the end of the training, teachers were asked to prepare an "action plan" for HIV/AIDS education in their school, including how they would reach out to the other teachers in the school and integrate HIV/AIDS into the timetable.

In addition to delivering the classroom-based activities, trained teachers were advised to

---

[5] These are quotes from the official textbook "Let's Talk About It", book 3 (grades 6, 7, 8), p. 26 and 19. (Textbook produced by the Kenya Institute of Education with the financial support of UNICEF).

[6] To my knowledge, the training has not yet been phased into all the control schools.

set up health clubs to encourage HIV avoidance through active learning activities such as role plays. A year after the training, 86% of the schools whose teachers had been trained had established health clubs.

2.3   Treatment 2: The Relative Risks Information Campaign (RR)

Information on the distribution of HIV infections by age and gender is typically not given to adolescents by their teachers, even if their teachers were trained, because it is not included in the official curriculum.[7] To test the effect of this information on teenage behavior, the "Relative Risk Information Campaign" (henceforth RR), was conducted in 2004 in 71 schools selected among the 328 schools involved in the TT program. The RR campaign was conducted by the non-governmental organization ICS. A trained project officer visited each of those 71 schools and, with the authorization of the teachers, spoke to Grade 8 students for a 40-minute period. At the start of the period, the students were asked to complete an anonymous, self-administered survey to determine how much they knew about the distribution of HIV in the Kenyan population.[8] After the survey, students were shown a 10-minute educational video on "sugar daddies".[9] The video screening was followed by an open discussion about cross-generational sex. During the discussion, the project officer shared the results of studies conducted in Kenya and Zambia (Glynn et al. 2001) and Zimbabwe (Gregson et al. 2002) on the role of cross-generational sex in the spread of HIV. In particular, the project officer wrote on the blackboard the detailed prevalence rates of HIV, disaggregated by gender and age group, in the nearby city of Kisumu, a place familiar to the students. The prevalence rates shared with students had been published by the WHO in 1997 and were reported in the Kenyan Government's brochure *AIDS in Kenya* of 2001 (Ministry of Health, 2001).[10] In accordance with the Kenyan government policy, the project officers conducting

---

[7] A few education programs specifically addressing the risks associated with "sugar daddy relationships" started after the onset of this study. See for example Population Services International (2005).

[8] The results of this survey were discussed in Section 2.1.2 and are presented in Table 1, Panel A.

[9] The animated movie, "Sara: the Trap," was produced by ACE communications, 2000, for UNICEF.

[10] The statistics provided to the students by the NGO were as follows:

Age	15-19	20-24	25-29	30-39
Female	22%	36%	35%	32%
Male	4%	13%	28%	32%

the RR program did not volunteer information on condoms nor demonstrate how to use condoms, but scientifically answered students' questions about condoms.

The 71 schools chosen for the RR program were selected randomly after stratifying by participation in the TT program of 2003. Thus, the RR program was implemented both in schools where teachers had been trained and in schools where teachers had not been trained.

3   Theoretical Motivation

This section provides a simple theoretical framework that incorporates the key facts observed in section 2.1.2, and highlights the outcomes that the experiment needs to identify in order to determine what information set (TT or RR) is socially optimal.

Consider that a teenage girl's utility depends on her consumption of purchased goods, C, as well as her health level, H, the number of sexual partnerships she is in, S, and whether she has a child, F:

$$U = U(C, H, S, F)$$

with $\partial U/\partial C \geq 0, \partial U/\partial H \geq 0$. U could be non-linear in S, with, for example, a maximum at $S = 1$. Having a child could be either desirable or undesirable.

Engaging in a sexual partnership entails both a risk of HIV infection and a risk of pregnancy. The health level H corresponds to the girl's HIV infection status, which itself depends on whether she engages in sexual partnerships, and on the characteristics of the partners she chooses: $H = H(S, age_{partner})$. Whether a teenage girl gets pregnant and has a child depends on whether she engages in sexual partnerships: $F = F(S)$.

Raising a child is costly. In other words, consumption of purchased goods is constrained by a budget constraint:

$$pC + fE(F) = W + Transf$$

where p is the price of purchased goods; f is the cost of rearing a child; W is the girl's

---

The city of Kisumu is the capital of Nyanza province, which is predominantly Luo, while this study took place in Western Province, which is predominantly Luhya. While prevalence in Kisumu is higher than in the rest of Kenya, the ratios between male and female by age-groups and the ratios between age-groups by gender are similar.

initial cash endowment; and $Transf$ is the net amount she gets in transfers from her sexual partner(s), itself the outcome of bargaining between the teenage girl and her partner(s). Finally, I consider that the consumption of purchased goods is subject to a survival constraint: $C \geq \underline{C}$.

Teenage girls maximize their expected utility based on their beliefs about the risks of HIV infection and pregnancy occuring, and how those risks vary with condom use and partners' characteristics. For example, the case where girls have no knowledge of the risk of HIV and assume zero HIV risk can be represented with the assumption $\Delta H/\Delta S = 0$: girls consider their health stock independent of their sexual behavior when choosing their utility-maximizing sexual behavior. The case where girls are aware that there is a risk of HIV infection corresponds to the assumption $\Delta H/\Delta S < 0$. If girls are unaware that the risk of infection increases with age among men, they will assume $\partial H/\partial age_{partner} = 0$. If they know how prevalence evolves with men's age, they assume $\partial H/\partial age_{partner} < 0$.

Note that in the case where girls derive no utility from being in a sexual partnership and no utility from children, the only reason why they might want to engage in a sexual partnership is if their initial endowment W is so low that in the absence of any transfers from partners, their survival constraint is binding. This subcase of the model can be called the "poverty-driven prostitution" case.[11]

Comparative Static 1: impact of the Teacher Training program

The impact of introducing general information about HIV (through the TT program) on teenage girls' demand for sexual partnerships (S) depends on two things: prior beliefs about HIV risk ($\Delta H/\Delta S$), and men's ability to pay (the distribution of the upper bound on $Transf$ among men).

- The effect of the information is zero for girls who already know about HIV at the time the information is provided.

- For girls who are unaware or underestimate the risk of HIV at baseline and update their

---

[11]This subcase of the model seems to be what most aid agencies and practitioners in Africa have in mind. A number of calls have been made for programs offering young women "economic alternatives to trading their bodies for material and financial gain" (Population Services International, 2005).

beliefs about $\Delta H/\Delta S$ when exposed to the TT program, the impact is ambiguous:

– Girls' reservation transfer goes up, and therefore their demand for sexual partnerships, holding the level of partners' transfers constant, decreases.

– But men may have the resources to increase the amount they transfer. If they can match girls' new reservation transfer, the demand for sexual partnerships among teenage girls will remain unchanged.

Comparative Static 2: impact of the Relative Risks Information campaign

- If all men have the same ability to transfer resources to their sexual partners ($Transf$ is independent of partners' age), information on the distribution of prevalence among men unambiguously leads teenage girls to move towards lower risk partners (in our case, teenage boys) and thus reduce the rate of cross-generational transmission of HIV.

- However, in a setting where teenage boys have lower ability to transfer resources, the impact of the RR information on the transmission rate is ambiguous, for two reasons.

– First, the adjustement might occur on the transfer size margin rather than on the partner selection margin: older men could simply compensate girls for the greater HIV risk they pose by transferring more resources to them (e.g. marrying them).

– Second, even if the adjustment happens on the partner type rather than on the transfer size margin, girls who face a binding survival constraint might need to increase the number of men they engage in a partnership with: if younger men can transfer only half of what older men can transfer, teenage girls in the "poverty-driven prostitution" case may need to have two simultaneous teenage boyfriends in order to meet their survival constraint. This would increase the incidence of unprotected sex between teenage girls and teenage boys.

- Finally, if information about the average prevalence of HIV in the area is already known at the time teenagers receive the RR information, providing RR information may trigger entry into sexual activity by previously abstaining teenage girls: they

learn that the infection rate among teenage boys is below the average infection rate, and thus revise downwards their beliefs about the risk of engaging in sex with teenage boys, which can affect the decision to enter sexual activity for girls at the margin.

Overall, the theoretical predictions of both programs are unclear. The TT program could either leave the level of sexual activity unchanged or decrease it. The RR information campaign could either not change anything, or reduce cross-generational relationships with a one-for-one substitution towards same-age relationships, or reduce cross-generational relationships with a more than one-for-one increase in same-age relationships. The potential positive effect on the amount of within-cohort sexual activity might have negative health consequences. While individuals fully internalize the cost of getting infected with HIV themselves, they might not fully internalize the fact that they might transmit the disease to others. Because of this externality, a public health maximizing policy maker may not always want to provide information that may increase the overall amount of sexual activity in the population. An empirical test is thus needed to resolve the type of information to provide in order to maximize public health benefits.

4   Data and Evaluation Strategy

4.1   The Sample

The Teacher Training program (TT) was phased into 164 schools over 4 months, from February to May 2003. The Relative Risks Information Campaign (RR) was phased into 71 schools over 4 months, from July to October 2004.[12] In November 2004, students in Grade 8 took the Kenya Certificate of Primary Education (KCPE) exam, the gateway exam to secondary school. Those students who had performed well on the exam and whose family could afford the tuition fees began secondary school in February 2005.

The TT program affected all grades in a school, whereas the RR program affected only Grade 8 students. Hereafter I consider all students who were enrolled in Grade 8 at the time

---

[12]The school year in Kenya starts in January and ends at the end of November.

of the RR campaign (2004) as part of the "study cohort". All students who were either a year ahead or a year behind this group can be used as control cohorts for the RR campaign.

Students in the study cohort were followed-up in 2005. Even though most of them had left primary school at the time of the follow-up, information on their whereabouts could still be collected at their primary school of origin. Appendix Table 1 presents summary statistics on their status in July 2005, averaged by school and broken down by gender and treatment status. Attrition is relatively low, below 2%, and cannot be distinguished across groups.

4.2   Data

I focus on two key outcomes: the incidence of unprotected sex between teenage girls and male partners five or more years older; and the incidence of unprotected sex between teenage girls and teenage boys. My main measure of the incidence of unprotected sex is the incidence of childbearing. To completement the childbearing data, I report evidence from self-reported sexual behavior data.[13]

4.2.1   Childbearing Data

Childbearing data was collected in two steps. First, information on childbearing and marital status was obtained for all girls in the sample during school visits conducted at regular intervals in the 12-month period of study. At each visit, the list of all students on the 2004 enrollment form was read aloud to pupils enrolled in upper grades in 2005, and for each of the students on the list, the following questions were asked: Is X still in school? If yes, in what grade? In what school? Does she still live in the area? Is she married? Does she have any children? If so, how many? How old is her first born? Is she pregnant?[14]

---

[13]Data on the incidence of HIV and other sexually transmitted infections in the sample is not available. A pilot biomarker follow-up conducted in 2007 indicates that the incidence of HIV and other sexually transmitted infections in this age group is not high enough for this study to have statistical power to detect significant differences in HIV prevalence between the treatment and the control groups.

[14]This technique of collecting childbearing and marital outcomes generates accurate data. Among a subsample of 282 teenage girls that were tracked at their home and interviewed, 88% of those who were reported as having started childbearing by their former schoolmates had indeed started childbearing, and 92% of those who were reported as not having started childbearing had indeed not started. The accuracy rates were similar across groups.

Second, enumerators conducted a home follow-up visit with all female students who had been reported to have started childbearing by July 2005. This follow-up included questions on the child's father (his age, marital status, and the transfers, if any, he had made before and after the pregnancy) in order to identify pregnancies that resulted from a cross-generational relationship. When the teenage girl herself could not be found (for example, because she had moved with her husband to another district), her mother was interviewed.[15]

Childbearing is not a perfect proxy for the incidence of risky sex, for various reasons. First, adolescent girls who are in a long-term relationship with one partner are more likely to get pregnant than those who have several short-term relationships.[16] Second, since pregnancy by a teenage boy is unlikely to result in marriage or child support (see Section 2.1.2), teenage girls who get pregnant might be more likely to abort if the father of the child is another teenager. Or teenage girls might start engaging in non-vaginal sex with teenage partners. Anal sex is of particular concern since it is highly risky in terms of HIV transmission. All in all, the RR information campaign could have increased the incidence of unprotected sex with teenage partners in the treatment group, but also increased the incidence of abortion and/or the incidence of anal intercourse, therefore generating no increase in pregnancy rates among same-age partners. Given this, comparing the incidence of childbearing with same-age partners across groups could underestimate the impact of the RR information campaign on the incidence of risky unprotected sex.

It is difficult to estimate the importance of these effects, since neither data on anal sex nor data on abortion is available. However, as a rough test of differential incidence of abortions across groups, we can compare mortality rates among girls. Because abortion is illegal in

---

[15]In the presence of concurrent partnerships, it is possible that a girl might not know who is the true biological father of her child. However, as long as a childbearing girl is married or supported by an older partner, it is safe to assume that, even if she had a concomittant teenage partner, she *must have had* unprotected sex with that older partner, in order to be able to convince him that he is the father. In addition, given that older partners are typically wealthier, we can expect that in the presence ambiguity about the biological father, the teenage girl (and possibly her parents) would choose to "blame" the pregnancy on the older (wealthier) partner.

[16]In particular, partners willing to have a child might decide to undergo HIV testing to ensure they are both negative before having unprocted sex. This is unlikely for the age group in this sample since 87% of pregnancies by teenage girls the area of study are declared unplanned, as presented in section 2.1.2. In addition, voluntary HIV counseling and testing services are not offered to minors in Kenya, unless they are considered "mature". The definition of "mature" is unclear.

Kenya, those that do happen tend to be unsafe and often result in maternal death.[17] The results are in Appendix Table A. I find that the mortality rate among girls between July 2004 and July 2005 was low (less than 0.2%) and similar across groups, providing some suggestive evidence that the incidence of abortion was not greater in the RR treatment group than in the RR comparison group. Overall, the incidence of childbearing in the sample seems a reasonable, though imperfect, proxy for the incidence of unprotected sex.

4.2.2 Self-Reported Sexual Behavior

Since childbearing data is not a perfect proxy for unprotected sex, it can be insightful to complement it with self-reported sexual behavior data. However, as most students in the study sample left primary school shortly after the RR information program was implemented, conducting a follow-up survey with everyone was not feasible.

Instead, a follow–up survey was administered to students enrolled in a secondary school in the study area between May and July 2005, about 7 to 9 months after the RR program. The survey included questions on sexual activity, characteristics of sexual partners, condom use, and primary school of origin, to identify the treatment status of each student. Overall, 55% of the secondary school students who completed the follow–up survey came from one of the 328 primary school that participated in the study.

As shown in Appendix Table 1, the likelihood that sampled teenagers enrolled in secondary school in 2005 is balanced across RR treatment and RR control schools, suggesting that the students who filled the follow-up behavioral survey were not differentially selected across groups for the RR program.[18] This is not the case for the TT program, as a higher fraction of students in the TT control group enrolled in secondary school than in the TT

---

[17]Unsafe abortion is a leading cause of maternal deaths in developing countries (Grimes, 2003). Up to 50% of maternal deaths in sub-Saharan Africa are due to induced abortion (Rogo, 1993). However, because abortions are illegal, getting estimates of the incidence of abortion among teenage girls in Kenya proves difficult.

[18]The fact that girls in the RR treatment group are not more likely to enroll in secondary school than girls in the RR control group, despite the fact that I show below a decrease in pregnancy in the RR treatment group, is interesting in itself. It suggests that girls who know that they can go to secondary school (because their parents can afford the fees and because they are performing well enough in grade 8) are not the marginal girls on whom the RR information program had an effect. This is not surprising, since the opportunity cost of pregnancy for girls who can attend secondary school is very high (pregnancy is *de facto* incompatible with schooling).

treatment group. This selection issue should be kept in mind while analyzing the results of the behavioral survey.

Another important caveat to the behavioral survey is that it includes only teenagers who joined secondary school, and those are clearly not representative of all teenagers. In Kenya, only students with sufficient financial resources and with high enough scores at the primary school exit exam can go to secondary school. This means that the behavioral data at hand includes only the richer and smarter students in the sample.

Finally, self-reported data on sexual behavior may suffer from reporting biases, and has been found to be often inconsistent with biological outcomes (Gersovitz et al. 1998). For example, in a study conducted in Western Kenya, Glynn et al. (2001) found that 12% of women who reported being virgins were HIV-positive (and some had other sexually-transmitted infections, making it unlikely that they acquired HIV non-sexually). Likewise, self-reported sexual data are typically impossible to reconcile at the population level (e.g. men typically report much more sexual activity than women).

While overall the self-reported data in this paper seems consistent with the biological (childbearing) data, it is important to keep in mind that the magnitude of the effects observed in the self-reported data could be largely due to reporting biases. In particular, it is possible that students who participated in the RR information campaign recognized that the NGO collecting the data was the same NGO that had talked to them about sexual partnerships the year before, and they might have been less wary of truthfully reporting their sexual activity than students who were less familiar with the NGO.

4.3   Evaluation Strategy

The randomized design provides a straightforward source of identification. For both the TT and the RR programs, random assignment of schools to the treatment and comparison groups ensures that the schools in either group are similar in all other respects except in that treatment schools were exposed to the program. Table 2 shows the baseline school averages for a series of school and pupils outcomes, by treatment groups. Except for class size, which is lower on average in RR treatment schools, all other differences in pre-treatment school characteristics are small and insignificant. However, the sample is less balanced when it

comes to the long-term schooling status of teenagers. Schools in the RR treatment group are significantly less likely to allow their students to repeat Grade 8 compare to schools in the RR control group. What's more, schools in the TT treatment group are significantly less likely to see their students go on to secondary school than schools in the TT comparison group. To palliate these imbalances, I systematically include controls for schooling status in the regression analysis outlined below.

To estimate the impact of the intervention, I use simple reduced form regression specifications. Denote Y_{isc} the outcome of individual i formerly enrolled in primary school s in cohort c. RR_s is the RR treatment status of school s and $StudyCohort_c$ the dummy for being in the study cohort. TT_s is a dummy equal to 1 if school s received the Teacher Training on the HIV curriculum.

First, I estimate the simple difference (SD) in means by ordinary least squares with clustering at the school level. The model is a linear probability model :

$$Y_{is1} = \alpha_1 + \beta_1 \times RR_s + \chi_1 \times TT_s + I_i'\gamma_1 + \varepsilon_{is}$$

where I_i is a vector of controls for individual characteristics (age and schooling status). The average effect of coming from a RR treatment school ($RR_s = 1$) versus a RR comparison school ($RR_s = 0$) is captured by β_1. Since $RR_s = 1$ was randomly assigned, we should expect $E(\varepsilon_{is}|RR_s) = 0$ so that the estimator of β_1 is unbiased. The average effect of coming from a school that received the teacher training can be captured by χ_1. Since $TT_s = 1$ was also randomly assigned, we should expect $E(\varepsilon_{is}|TT_s) = 0$ so that the ordinary least squares estimator of χ_1 is also unbiased. By comparing β_1 and χ_1, we can thus compare the impact of the two information sets.[19] To increase the precision of the estimators, I control for the observable characteristics of the primary school of origin (for the childbearing data) and for the characteristics of the current secondary school (for the behavioral survey). Finally, when the outcome is binary, I also estimate the effect of the program on the probability that the outcome occurs using a Probit model.

---

[19]Considering the low incidence of childbearing, the sample size does not provide enough power to estimate the effect of the interaction between the two programs.

Second, I estimate the difference-in-differences (DD) when data on a control cohort is available:

$$
\begin{aligned}
Y_{isc} =\ & \alpha_2 + \beta_2 \times RR_s \times StudyCohort_c + \delta \times RR_s + \theta \times RR_s \\
& + \chi_2 \times TT_s + I_i' \gamma_2 + \omega_{isc}
\end{aligned}
$$

Comparing the single-difference to the difference-in-differences estimates is useful for two reasons. First, if the randomization of the RR program assignment was not perfect, the difference-in-differences will adjust for potential pre-existing random differences between RR treatment and RR comparison schools. Second, the difference-in-differences allows the inclusion of school fixed effects, which allows to control for unobservable school characteristics. However, the double-difference estimates could be biased in the presence of treatment spillovers across cohorts. I will come back to this issue when I discuss the results below.

5 Results

5.1 Impact on Incidence of Teen Childbearing

Table 3, Columns 1 to 4 show the estimates of the effects of each program on the incidence of childbearing with four different regression specifications: the simple difference with a linear probability model (OLS); the simple difference with a probit model (reporting marginal effects); the OLS estimate of the difference-in-differences; and the OLS estimate of the difference-in-differences with school fixed effects. The RR information campaign reduced the incidence of childbearing by 1.5 percentage points among treated girls relative to girls in the comparison group (Table 3, Column 1). The childbearing rate in the comparison group is 5.4%, and thus the RR treatment effect corresponds to a 28% decrease in the incidence of childbearing. The magnitude and significance of the RR effect are robust to all specifications. In the difference-in-differences without school fixed effects, the estimate of the coefficient for "RR Information" is close to zero, confirming the absence of ex-ante difference between treatment and comparison schools (Column 3, Row 1).

In contrast, the TT program had no impact on the incidence of childbearing (Row 3).

This despite the fact that the training had a large impact on the amount of HIV education delivered in schools and increased scores of pupils on HIV knowledge tests (Duflo et al., 2006). This result may reflect the fact that the curriculum promotes abstinence until marriage as the only way to avoid HIV infection, and so would not deter teenagers from having children and marrying at a young age.

Columns 5 to 12 in Table 3 show estimates of the treatments on childbearing broken down by marital status. The bulk of the decrease in the incidence of childbearing in the RR treatment group corresponds to a decrease in childbearing outside of marriage, while the incidence of childbearing within marriage decreased only slightly and not significantly. This means that, among girls who started childbearing, the proportion of girls who are married is significantly larger in the treatment group than in the comparison group. Since women typically receive greater financial support from their partner when they are married than when they are not, these findings imply that, relative to girls in the comparison groups, girls who received RR information are more likely to refuse to enter into an unprotected sexual relationship with an adult man unless they get compensation commensurate with the higher risk involved. In other words, in the absence of the RR information program, girls at the margin (just above the threshold of engaging with an older partner) would have engaged in unprotected sex with an older partner with relatively limited resources, who would have turned out to not marry in case of a pregnancy.

5.2   Age of Childbearing Partner

To determine the extent to which the observed decrease in the incidence of childbearing in the RR treatment group corresponds to a decrease in the incidence of unprotected sex with older men, I look at the age differential between girls who have started childbearing and their partners. The data is available for two cohorts: the study cohort (Grade 8 of 2004) and one control cohort (Grade 7 of 2004).

Since the RR information campaign reduced the incidence of childbearing in the RR treatment group, the data is available for differentially selected subsamples of each group. Nevertheless, as long as the RR treatment did not affect the likelihood of getting pregnant conditional on having unprotected sex differentially across relationships' types (cross- or

intra-generational), we should expect the ratio of cross-generational to intra-generational pregnancies to be the same across subsamples, unless the RR treatment had an effect on age differentials between partners.

The dependent variable in Table 4, Columns 1 and 2, is the age difference between the respondent and her baby's father. The RR treatment effect is negative and significant: among girls who had begun childbearing, the average age gap with the baby's father is 1.7 years smaller for RR treated girls compared to girls who did not receive the RR treatment (Column 1). This difference is significant at the 99% confidence level. In Columns 3 to 5, the dependent variable is a dummy indicating whether the baby's father is more than 5 years older than the teenage girl. The coefficient of the treatment effect is negative and large (-22 percentage points off of a mean of 48% in the control group) and significant at the 95% confidence level (Column 3). In contrast, the TT program seems to have, if anything, increased the likelihood that teenage girls start childbearing with older partners, though none of the coefficients are significant (Row 4).

The difference-in-difference estimates of the effect of the RR program on the RR cohort are greater in magnitude than the simple difference results. This seems driven by the fact that the coefficients for being in an RR information school (but not in the RR cohort) are large and positive (Row 1, Columns 2, 5 and 9). This suggests that the RR program might have had negative spillovers onto non-treated students in the RR treatment schools. Indeed, the control cohort used in Table 5 is a *younger* cohort (the 7th graders of 2004). This cohort could have been indirectly and negatively affected by the RR information program if the sugar daddies newly turned down by informed 8th graders decided to try their luck with 7th graders instead. Alternatively, the 7th graders could have benefitted from positive information spillovers if the 8th graders shared the information with their younger schoolmates. To fully assess the epidemiological consequences of providing information on HIV risk by group, it would be necessary to look at such general equilibrium effects. In particular, one should address the question of what becomes of the older men who are turned down by informed teenage girls. On one hand, they might reduce their sexual activity. On the other hand, they might engage in unprotected sex with commercial sex workers, which could have negative epidemiological consequences. I do not address these issues in this paper, as I do not have

data on the behavior of older men in the study area.[20]

5.3 Overall impact of RR Information on pregnancies by partner's type

Table 5 combines the results of Tables 3 and 4 to compute the treatment effects of the RR and TT programs on the incidence of pregnancies with older partners and the incidence of pregnancies with teenage partners. I consider a normalized case in which 100 pregnancies occur in the RR comparison group. Of these, 47.6 are by men more than 5 years older (hereafter labelled "older men"). In the RR treatment group, we would observe 72 pregnancies, 18.4 of which are by older men. Thus, the RR program averts 29.2 pregnancies by older partners in the treatment group. This means that the incidence of cross-generational pregnancies declined by 61.3% in the RR treatment group relative to the comparison group, while intra-generational pregnancies remained stable.

These results suggest that providing teenagers with information on relative risks led to a large decrease in the incidence of unprotected sex between teenage girls and older men, but did not lead to an increase in the incidence of unprotected sex between teenage girls and teenage boys. This suggests that the RR program might have reduced teenagers' exposure to the risk of HIV infection.

In contrast, the TT program had, if anything, a positive impact on the incidence of childbearing by older partners, although the effects cannot be distinguished from zero (none of the coefficient estimates for the TT program in Tables 3 and 4 are statistically significant).

---

[20]It is possible that similar spillover effects could have been at play across primary schools. On one hand, information on relative risks could have spread to comparison primary schools that are near treatment schools, and girls in those comparison schools may also have avoided unprotected sex with adult partners. This would mean that the treatment effect on childbearing estimated above is an underestimation of the overall effect of the RR information campaign. On the other hand, if adult men responded to the change in the price charged by treated teenage girls by moving away from treatment schools and towards the surroundings of comparison schools when looking for sexual partners, the information campaign may have generated an increase in childbearing by adult men in the comparison schools, and consequently the comparison between treatment and comparison schools would be overestimating the treatment effect. However, since the treatment group is more than 4 times smaller than the comparison group, it is unlikely that this price effect could explain more than a fourth of the treatment effect found.

5.4 Mechanisms? Suggestive evidence from self-reported sexual behavior

Did the decrease in the incidence of childbearing by older partners in the RR treatment group come from an increase in condom use within cross-generational partnerships or from a decrease in the number of cross-generational partnerships? If teenage girls in the treatment group did not engage in partnerships with older men, did they substitute towards teenage boys (low-risk option) or towards abstinence (zero-risk option)? The general case of the model predicts a one-for-one substitution towards teenage partners, but the "poverty-driven prostitution" subcase (when the survival constraint is binding) implies that the substitution towards teenage boys should be more than one for one, since teenage boys have less resources than older men, and thus would predict an increase in the number of partners reported by teenage girls.

To shed light on these issues, Table 6 presents self-reported data collected among teenagers who joined a secondary school in the study area. This subgroup is not representative of all teenagers in the sample.[21] Nevertheless, studying the impact of the RR program on the self-reported sexual behavior of secondary school students is interesting in itself and can help shed light on the mechanisms behind the observed decrease in childbearing with older partners.

For each outcome, Table 6 shows the results of two specifications: the OLS estimate of the simple difference (Column 1) and the probit estimate of the simple difference (Column 2). Panel A shows the coefficient estimates for girls and Panel B shows the estimates for boys.

Girls in the RR treatment group are not more likely to have had multiple partners than other girls, but conditional on being sexually active, they are 50% more likely to report having a regular partner. The likelihood that their partner is more than 5 years their senior is significantly lower, reduced to close to zero (against 8% in the RR control group). Girls

---

[21]Girls enrolled in secondary school have a higher incentive to avoid childbearing than girls who are out of school, since childbearing is often not compatible with schooling. As such, the sexual behavior of secondary school girls may differ substantially from that of out-of-school girls. Similarly, out-of-school boys are more likely to be working, and thus may have more income than secondary school boys, which may raise their ability to make transfers to girls in return for sex.

in the TT treatment group are also 50% more likely to declare having a regular partner, but the TT program had no effect on the likelihood that girls choose an older partner.

The results for boys are quite different. Boys in the RR treatment group are 10 percentage points more likely to report having had sex with multiple partners. This represents an almost 50% increase compared to the base rate of 23% in the control group.

The RR program also led to an increase in the likelihood that teenagers report ever having had sex (Columns 7-8). The difference between RR treatment and RR control girls is 8 percentage points, significant at 1%. This suggests that the RR information decreased the perceived riskiness of teenage boys, and therefore some girls who were abstinent before the RR information campaign decided to enter the market for sex with teenage boys.[22] Accordingly, boys coming from RR treatment schools are 12 percentage points more like to report ever having had sex. In contrast, the TT program, if anything, decreased the likelihood that teenagers report being sexually active, but the effect is small and not significant.

The increase in sexual activity for both boys and girls in the RR treatment group does not seem to correspond to an increase in unsafe sexual activity: the share of teenagers who ever had sex but never used a condom did not increase significantly (Columns 9-10). The share of girls who report having used a condom at their last sexual intercourse is 24% in the RR control group, and this increases by 11 to 12 percentage points (a 50% difference) among girls in the RR treatment group (Columns 11-12). However, there is no signicant change in condom use at last intercourse for boys in the RR treatment group compared to the comparison.

Overall, the self-reported sexual behavior data at hand, while imperfect, confirms the childbearing results, and suggests that, in response to the RR information, teenage girls substituted away from older partners and towards protected sex with teenage partners, but not more than one-for-one. In addition, the self-reported data suggests that the RR information triggered some teenage girls to enter sexual activity earlier, but with condoms, presumably to avoid pregnancy with a resource-constraint teenage boy.

---

[22] Alternatively, this result could be a pure reporting artefact as discussed in section 4.2.2.

5.5 Cost-Effectiveness of the RR Information Campaign

This section provides a rough calculation of the cost-effectiveness of providing RR information to teenagers through schools, and compares it to other HIV prevention programs that have been shown successful.

The information campaign reached about $1,300$ girls in 71 schools and cost just under US$2,000$. The campaign reduced the incidence of childbearing by 1.5 percentage points in the treatment group, which means that a total of 20 ($1,300 \times 0.015$) pregnancies were averted thanks to the program. All of the averted pregnancies would have resulted from a cross-generational partnership as computed in section 5.3. Thus, the overall cost per cross-generational pregnancy averted is just under US100.

To calculate the cost per HIV infection averted, I need an estimate of the ratio of the risk of HIV infection to the risk of cross-generational pregnancy, a ratio which is not available in the literature.[23] Instead, I compute cost-effectiveness estimates using three hypothetical ratios: $5/100$, $15/100$, and $25/100$.[24] Table 7 shows the cost per HIV infection using these ratios. For a ratio of $15/100$, US98 per cross-generational pregnancy averted corresponds to a cost of US653 per primary HIV infection averted among teenage girls (Scenario 2). It is important to note, however, that these estimates consider only primary cases of HIV transmission, and thus do not include averted secondary HIV infections (i.e. transmission to subsequent sex partners).

These rough cost-effectiveness estimates compare favorably with other HIV prevention programs, such as treating sexually transmitted infections other than HIV (estimated at US213 per HIV infection averted by Gilson et al. 1997); male circumcision (estimated at US$1269-3911$ per infection averted by Gray et al., 2007); or voluntary HIV testing (estimated at US537 per additional HIV positive person tested by Thornton, 2008).

---

[23]What makes this ratio particularly complicated to estimate is the fact that infectiousness of a person with HIV varies with her viral load, and the viral load follows a U-shape: it is very high during the first weeks or months after infection, then decrease substantially and peaks again (in the absence of ARV treatment) when the patient develop AIDS 8 or 9 years later (Magruder, 2007).

[24]The Kenya antenatal surveillance sites recorded HIV prevalence rates in pregnant women between 12 and 35 percent in 2002 (Baltazar et al. 2001).

6 Conclusion

This paper uses a randomized field experiment to study the change in the sexual behavior of Kenyan teenagers in response to information on HIV risk. I use the incidence of pregnancy as an objective proxy for the incidence of unprotected sex. I find that providing information on the relative risk of HIV infection by type of partner led to a 61% decrease in the incidence of pregnancies with older (riskier) partners among teenage girls, without increasing pregnancies with same-age partners. In contrast, the national HIV/AIDS education program, which provided general information about the risk of HIV but did not inform teenagers of the risk distribution in the population, had no impact on the incidence of unprotected sex, as measured by pregnancy rates. Self-reported sexual behavior data collected among a subset of the sample is consistent with the pregnancy results, and suggest that teenage girls who received information on the relative riskiness of older partners substituted away from older partners and towards condom-protected sex with same-age partners.

The findings of this paper suggest that teenagers are responsive to risk information, but their sexual behavior is more elastic on the intensive than on the extensive margin. These findings have important policy implications. Public health interventions often focus their efforts on the extensive margin of a risky behavior: they aim at the complete elimination of the behavior and urge complete abstinence from the activity. Accordingly, they rarely provide information on the relative riskiness of different varieties of a risky activity – information that would enable people to reduce the intensity of their exposure to risk while remaining active. For example, despite the fact that three randomized medical trials in sub-Saharan Africa have recently demonstrated that circumcision considerably reduces the risk of HIV transmission for men, some African governments are wary of disseminating such information and of promoting circumcision, because they are afraid it could induce circumcised men to increase their demand for unprotected sex. However, the amount of information that a prevention campaign should provide in order to maximize its health impact depends on the relative size of two elasticities: the elasticity between high- and low-risk varieties of an activity and the elasticity between activity and no activity at all. The empirical evidence presented in this paper suggests that, in the case of sexual behavior, the former appears the

26

larger of the two. This result suggests that HIV education campaigns may achieve a wider health impact if they include both risk reduction and risk avoidance information.

References

[1] Bundy, Donald (2002). *Education and HIV/AIDS: A Window of Hope.* Washington DC, The World Bank.

[2] Central Bureau of Statistics (CBS) [Kenya], Ministry of Health (MOH) [Kenya], and ORC Macro. (2004). *Kenya Demographic and Health Survey 2003.* Caverton, Maryland: CBS, MOH, and ORC Macro.

[3] Clark, Shelley (2004) "Early Marriage and HIV risks in Sub-Saharan Africa". *Studies in Family Planning* 35[3]: 149–160.

[4] Dee, Thomas and William Evans (2001). "Teens and Traffic Safety", in J. Gruber, editor, *Risky Behavior Among Youths: An Economic Analysis*, University of Chicago Press, pp. 121-165.

[5] Duflo, Esther, Pascaline Dupas, Michael Kremer and Samuel Sinei (2006) "Education and HIV prevention: Evidence from Western Kenya", Policy Research Working Paper 4024, World Bank.

[6] Gallant, Melanie and Eleanor Maticka-Tyndale (2004), "School-based HIV prevention programmes for African youth", *Social Science and Medicine*, 2004, 1337-1351

[7] Gersovitz, Mark, Hanan G. Jacoby, F. Seri Dedy and A. Goze Tape, (1998) "The Balance of Self-Reported Heterosexual Activity in KAP Surveys and the AIDS Epidemic in Africa," *Journal of the American Statistical Association*, vol. 93, pp. 875-883.

[8] Glynn JR, Caraël M, Auvert B, Kahindo M, Chege J, Musonda R, Kaona F, Buvé A. (2001) "Why do young women have a much higher prevalence of HIV than young men? A study in Kisumu, Kenya, and Ndola, Zambia". *AIDS*, 15 (suppl 4): S51-S60.

[9] Gray RH, Li X, Kigozi G, Serwadda D, Nalugoda F, Watya S, Reynolds SJ, Wawer M.(2007). "The Impact of Male Circumcision on HIV Incidence and Cost per Infection Prevented: A Stochastic Simulation Model from Rakai, Uganda". *AIDS*. Apr 23;21(7):845-850.

[10] Gregson S, Nyamukapa CA, Garnett GP, Mason PR, Zhuwau T, Caraël M, Chandiwana SK, Anderson RM. (2002) "Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe". *The Lancet*; Vol 359.

[11] Grimes, David A. (2003). "Unsafe abortion: the silent scourge". *British Medical Bulletin* 67:99-113

[12] Gruber, Jonathan and Jonathan Zinman (2001). "Youth smoking in the United States: Evidence and Implications" in J. Gruber, editor, *Risky Behavior Among Youths: An Economic Analysis*, University of Chicago Press, pp. 69-120.

[13] Jensen, Robert (2007): "The Perceived Returns to Education and the Demand for Schooling". Mimeo, Harvard University.

[14] Kelly RJ, Gray RH, Sewankambo NK, Serwadda D, Wabwire-Mangen F, Lutalo T, Wawer MJ. (2003). "Age differences in sexual partners and risk of HIV-1 infection in rural Uganda". *Journal of Acquired Immune Deficiency Syndrome*, 32:446-451.

[15] Kirby, Douglas (2008) "The Impact of Abstinence and Comprehensive Sex and STD/HIV Education Programs on Adolescent Sexual Behavior" *Sexuality Research and Social Policy* (2008;5(3):6-17)

[16] Laga M, Schwärtlander B, Pisani E, Sow PS, Caraël M. (2001) "To stem HIV in Africa, prevent transmission to young women." *AIDS* 15: 9310934.

[17] Levitt, Steve and Lance Lochner (2001). "The determinants of Juvenile Crime", in Jonathan Gruber, editor, *Risky Behavior Among Youths: An Economic Analysis*, University of Chicago Press, pp. 327-374.

[18] Luke, Nancy (2003) "Age and economic asymmetries in the sexual relationships of adolescent girls in sub-Saharan Africa." *Studies in Family Planning*, Jun;34(2):67-86.

[19] Luke, Nancy and Kathleen M. Kurz (2002) "Cross-generational and Transactional Sexual Relations in Sub-Saharan Africa: Prevalence of Behavior and Implications for Negotiating Safer Sexual Practices." Washington. DC: International Center for Research on Women.

[20] Magruder, Jeremy (2007). "Marital Shopping and Epidemic AIDS." Mimeo, Yale University.

[21] Ministry of Health, Kenya (2001) *AIDS in Kenya: Background, Projections, Impact, Interventions and Policy*, eds. 6th Ed, Nairobi, AIDS Control Unit of Ministry Of Health and National AIDS Control Council.

[22] Peterman, T, Stoneburner R, Allen J, Jaffe H, and Curran J. (1988) "Risk of HIV transmission from heterosexual adults with transmission associated infections". *JAMA*; 259: 55–58.

[23] Population Services International (2005). "The Sugar Daddy Syndrome: African Campaigns Battle Ingrained Phenomenon". Available online at: http://www.psi.org/resources/pubs/cross-gen.pdf (accessed November 22, 2008)

[24] Rogo, Kama O. (1993). "Induced abortion in sub-Saharan Africa" *East African Medical Journal.*70(6):386-95.

[25] Thornton, Rebecca L. (2008). "The Demand for and Impact of Learning HIV Status: Evidence from a Field Experiment". *American Economic Review* 98 (5): 1829-63.

[26] Trenholm, Christopher, Barbara Devaney, Ken Fortson, Lisa Quay, Justin Wheeler, and Melissa Clark. "Impacts of Four Title V, Section 510 Abstinence Education Programs". Final report, Mathematica Policy Research Inc.. Availble online at: http://aspe.hhs.gov/hsp/abstinence07/index.htm (accessed November 28, 2008).

Figure 1
HIV Prevalence by Gender and Age Groups, Kenya

% HIV Positive
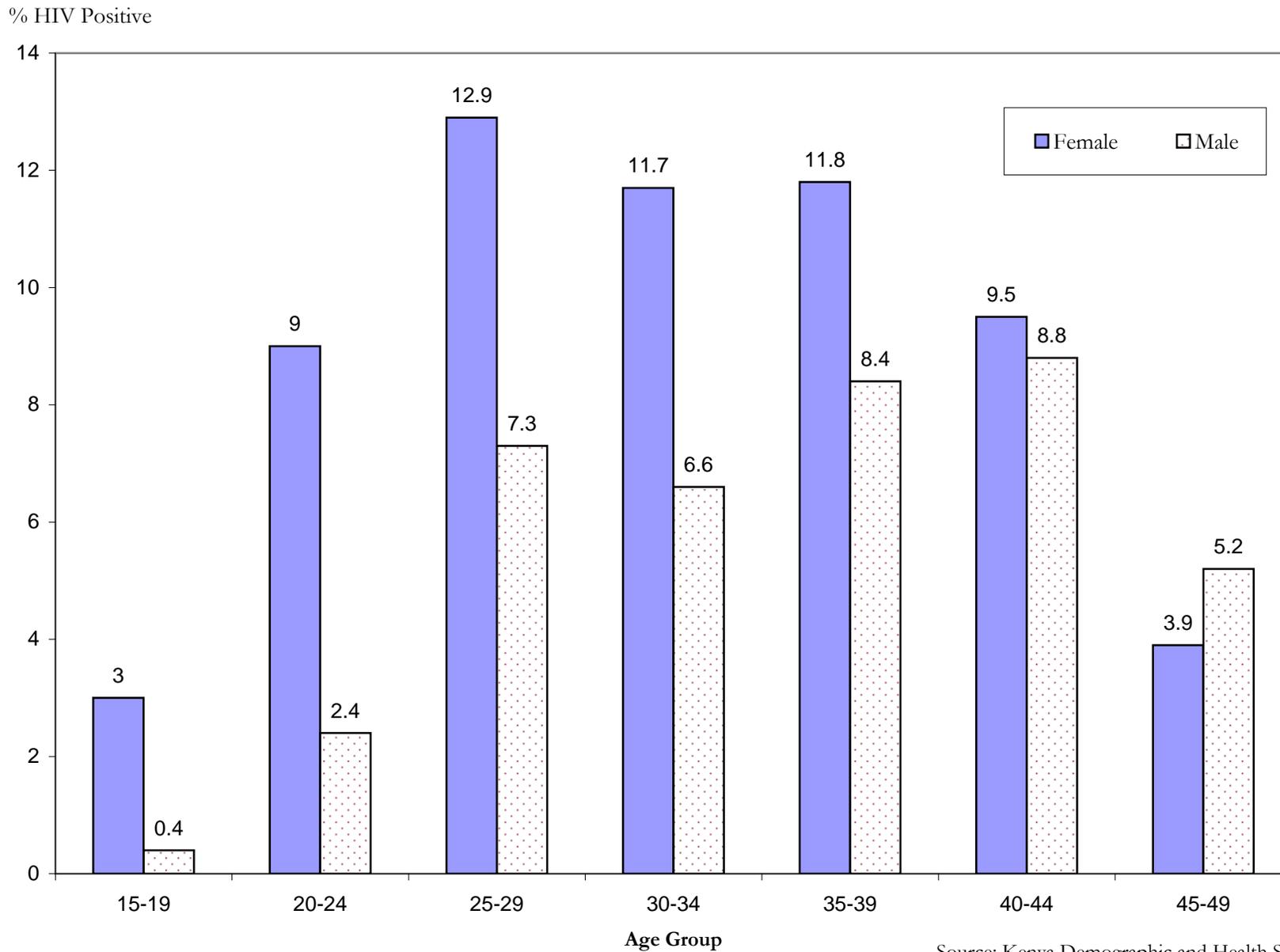
Source: Kenya Demographic and Health Survey, 2003

Table 1 *Summary Statistics on Knowledge and Behavior among Adolescents in Study Area*

Panel A: Baseline Knowledge, Attitudes and Practice (Self-reported) among 8th graders

	(1) Girls	(2) Boys
Average Age	15.10	15.52
	(1.2)	(1.5)
Share thinking condoms can prevent pregnancy	0.46	0.71
Share thinking condoms can prevent HIV infection	0.45	0.66
Share thinking men above 25 have a higher HIV infection rate than teenage boys	0.29	0.25
Share reporting having had sex	0.21	0.48
Share reporting that some girls in the class have a partner who is not a student	0.61	0.57
Number of observations	1176	1246

Panel B: Partnership Survey (Girls who started childbearing within a year of starting 8th grade)

Share reporting that the pregnancy was wanted	0.13
Share reporting age difference with male partner > 5 years	0.43
Share reporting age difference with male partner > 10 years	0.11
Share reporting that the partnership was consensual	0.99
Share reporting that the male partner made regular cash payments to the teenage girl prior to the pregnancy	0.73
Share reporting that the male partner is currently providing financial support to the teenage girl	0.75
Share married if age difference < 5 years	0.42
Share married if 5 years <age difference <10 years	0.63
Share married if age difference > 10 years	0.79
Number of Observations	362

Notes: Standard deviations are presented in parentheses. Panel A: Self-reported data collected among teenagers enrolled in Grade 8 in 2004, prior to RR information campaign. The survey was self-administered. Panel B: Childbearing data collected in August 2005 for teenage girls enrolled in a RR comparison school at baseline (2004), and who had begun childbearing by July 2005. In 55% of cases, the teenage girl was interviewed herself; in the rest of cases, she was not at home on the day of the enumerator's visit and a relative (typically her mother) answered questions on her behalf.

Table 2 *Verifying Randomization:*
School Characteristics and Students Outcomes at Baseline, by Treatment Group

	(1)	(2)	(3)	(4)	(5)	(6)
	RR Information			TT on HIV/AIDS Curriculum		
	Comparison Group (C)	Treatment Group (T)	Difference T-C	Comparison Group (C)	Treatment Group (T)	Difference T-C
Panel A. School Characteristics at Baseline						
Class Size	38.2	34.4	-3.8	37.4	37.3	-0.06
	[15.9]	[17.4]	(1.540)**	[16.9]	[15.7]	(1.281)
Pupils sex ratio (Girls/Boys)	1.07	1.12	0.049	1.06	1.10	0.040
	[0.489]	[0.668]	(0.072)	[0.476]	[0.586]	(0.059)
Teacher-pupil ratio	0.026	0.026	0.000	0.025	0.027	0.003
	[0.026]	[0.022]	(0.003)	[0.021]	[0.028]	(0.003)
Teachers sex ratio (Females/Males)	1.033	0.921	-0.112	1.003	1.014	0.011
	[0.914]	[0.777]	(0.119)	[0.92]	[0.852]	(0.099)
KCPE results (2003)	251.0	249.4	-1.6	252.2	249.0	-3.2
	[29.0]	[27.4]	(3.9)	[28.6]	[28.5]	(3.2)
Sampled for TT on HIV/AIDS Curriculum	0.50	0.49	-0.003	0.00	1.00	
			(0.067)			
Sampled for RR information	0.00	1.00		0.22	0.22	-0.002
						(0.046)
Panel B. Girls enrolled in 8th grade in 2003 (Control Cohort for RR): Outcomes at the end of 2004						
Percent repeating class 8	0.245	0.208	-0.04	0.236	0.238	0.001
			(0.021)*			(0.018)
Percent in Secondary School	0.448	0.455	0.007	0.471	0.428	-0.044
			(0.026)			(0.021)**
Percent in Professional Training	0.037	0.036	0.000	0.037	0.036	0.000
			(0.008)			(0.007)
Percent out of School	0.259	0.288	0.029	0.246	0.286	0.040
			(0.022)			(0.018)**
Percent married	0.112	0.121	0.009	0.106	0.121	0.015
			(0.014)			(0.011)
Percent who had begun childbearing	0.122	0.115	-0.008	0.110	0.132	0.022
			(0.018)			(0.015)
Observations	4783	1212	5995	3016	2979	5995
Panel C. Boys enrolled in 8th grade in 2003 (Control Cohort for RR): Outcomes at the end of 2004						
Percent repeating class 8	0.23	0.22	-0.006	0.23	0.22	-0.004
			(0.022)			(0.018)
Percent in Secondary School	0.52	0.51	-0.012	0.53	0.51	-0.021
			(0.027)			(0.023)
Percent in Professional Training	0.01	0.01	-0.006	0.02	0.01	-0.007
			(0.004)			(0.004)*
Percent out of School	0.23	0.25	0.025	0.22	0.25	0.036
			(0.023)			(0.019)*
Percent married	0.023	0.014	-0.009	0.022	0.020	-0.002
			(0.008)			(0.006)
Observations	4845	1229	6074	3079	2995	6074
Number of Schools	252	71	323	163	160	323

Notes: These are school averages. School characteristics collected in 2004. Students outcomes collected in 2004 for the control cohort (Grade 8 of 2003). Standard deviations in brackets. Columns 3 and 6: Standard errors in parenthesis. Significantly different than zero at 1 (***), 5 (**), and 10 (*) percent. Five schools did not have an 8th grade class in 2003 and therefore are excluded from the table.

Table 3 *Probability that Girls have started Childbearing*

SPECIFICATION MODEL	Has started childbearing				Has started childbearing, Unmarried				Has started childbearing, Married			
	SD OLS (1)	SD PROBIT (ME) (2)	DD OLS (3)	DD-FE OLS (4)	SD OLS (5)	SD PROBIT (ME) (6)	DD OLS (7)	DD-FE OLS (8)	SD OLS (9)	SD PROBIT (ME) (10)	DD OLS (11)	DD-FE OLS (12)
RR Information	-0.015 (0.008)**	-0.014 (0.007)**	0.004 (0.011)		-0.011 (0.005)**	-0.009 (0.004)**	0.006 (0.008)		-0.004 (0.006)	-0.002 (0.005)	-0.002 (0.009)	
RR Information x 2004 Cohort			-0.024 (0.013)*	-0.024 (0.013)*			-0.017 (0.009)*	-0.019 (0.008)**			-0.006 (0.010)	-0.004 (0.010)
TT on HIV/AIDS Curriculum	0.006 (0.006)	0.008 (0.006)	0.000 (0.005)		0.006 (0.004)	0.006 (0.003)*	0.002 (0.003)		0.000 (0.005)	0.002 (0.004)	-0.002 (0.004)	
Sample												
Control Cohort Included (2003 cohort)			Yes	Yes			Yes	Yes			Yes	Yes
Controls												
Individual Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Primary School Characteristics	Yes	Yes	Yes		Yes	Yes	Yes		Yes	Yes	Yes	
Primary School Fixed Effects				Yes				Yes				Yes
Observations	5989	5989	10970	10970	5989	5989	10970	10970	5989	5989	10970	10970
Mean of Dep Var (RR=0)	0.054	0.054	0.054	0.054	0.021	0.021	0.021	0.021	0.033	0.033	0.033	0.033

Notes: Specifications: SD = simple difference; DD = difference-in-difference; FE= school fixed effects. Only the 2004 cohort was affected by the RR Information program. The dependent variables are at individual-level dummies. Robust standard errors in parentheses, clustered at the school level. Significantly different than zero at 1(***), 5(**) and 10(*) percent. Columns 2, 6 and 10 report mean marginal effects. Individual controls include: schooling status, age, and cohort when applicable. School controls include: gender ratio among pupils, average school performance on the national KCPE exam, location, timing of follow-up visit. Follow-up visits were conducted after 15 to 21 months for the 2003 Cohort, and after 12 to 17 months for the 2004 Cohort. The timing of visits was balanced across treatment and control schools, with an average gap of 16 months between baseline and follow-up in both groups (in both groups: 17.6 months for the 2003 Cohort and 14.5 months for the 2004 Cohort).

Table 4 *Age Gap between Girls who have started Childbearing and their Partner*

SPECIFICATION	Age Difference between Teenage Girl and her partner		Age gap >5 years			Age gap >10 years		
MODEL	SD OLS	DD OLS	SD OLS	SD PROBIT (ME)	DD OLS	SD OLS	SD PROBIT (ME)	DD OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
RR Information	-1.734	1.035	-0.221	-0.229	0.161	-0.066	-0.078	0.171
	(0.571)***	(0.858)	(0.109)**	(0.090)**	(0.121)	(0.061)	(0.052)	(0.084)**
RR Information x 2004 Cohort		-2.548			-0.350			-0.236
		(1.061)**			(0.187)*			(0.109)**
TT on HIV/AIDS curriculum	-0.811	-0.409	0.083	0.103	0.027	-0.078	-0.066	-0.031
	(0.711)	(0.472)	(0.081)	(0.076)	(0.062)	(0.059)	(0.053)	(0.038)
<u>Sample</u>								
Control Cohort Included (2005 cohort)		Yes			Yes			Yes
<u>Controls</u>								
Individual characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Primary School Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	120	250	134	134	278	134	134	278
Mean of Dep Var (RR=0)	5.84	5.84	0.48	0.48	0.48	0.16	0.16	0.16
Std. Dev.	(4.21)	(4.21)						

Notes: Specifications: SD = simple difference; DD = difference-in-difference. Only the 2004 cohort was affected by the RR Information program. The dependent variables are at the individual level. Robust standard errors in parentheses, clustered at the primary school level. Significantly different than zero at 1(***), 5(**) and 10(*) percent. Columns 4 and 7 report mean marginal effects. Individual controls include: schooling status, age, and cohort when applicable. School controls include: gender ratio among pupils, average school performance at national KCPE exam, location.

Table 5 *Overall Treatment Effects on Incidence of Childbearing by Male Partners More than 5 Years Older*

Panel A. Effect of Relative Risks Information

| | Sampled for RR Information? | | | |
	NO	YES	Treatment Effect in %	Change
	Base=100			
# Teen Pregnancies	100.0	72.3	-27.7%	-27.7
Share of Pregnancies by men >5 years older	48%	26%	-22.1%	
# Pregnancies by men >5 years older	47.6	18.4	-61.3%	-29.2
# Pregnancies by men ≤5 years older	52.4	53.8	2.8%	1.4

Panel B. Effect of Teacher Training on HIV/AIDS curriculum

| | Sampled for TT on national curriculum? | | Treatment | |
	NO	YES	Effect	Change
	Base=100			
# Teen Pregnancies	100.0	111.1	11.1%	11.1
Share of Pregnancies by men >5 years older	48%	56%	8.3%	
# Pregnancies by men >5 years older	47.6	62.1	30.5%	14.5
# Pregnancies by men ≤5 years older	52.4	49.0	-6.5%	-3.4

Notes: In each panel: First row: effect on number of teen pregnancies reported from Table 3, column 1. Second row: effect on share of pregnancies by older men reported from Table 4, column 3.

Table 6 *Self-Reported Sexual Behavior for Students who joined Secondary School*

Panel A: Girls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Has had sex with multiple partners		Has a regular partner		Partner >5 years older		Ever had sex		Ever had sex but never used a condom		Used a condom at last sexual intercourse	
MODEL	OLS	PROBIT (ME)	OLS	PROBIT (ME)	OLS	PROBIT (ME)	OLS	PROBIT (ME)	OLS	PROBIT (ME)	OLS	PROBIT (ME)
RR Information	0.005	0.005	0.130	0.118	-0.070	-0.079	0.085	0.075	0.020	0.019	0.121	0.111
	(0.015)	(0.012)	(0.066)*	(0.055)**	(0.041)*	(0.044)*	(0.029)***	(0.024)***	(0.026)	(0.023)	(0.072)*	(0.063)*
TT on HIV/AIDS curriculum	-0.010	-0.009	0.122	0.123	0.003	-0.004	-0.013	-0.013	-0.016	-0.017	-0.006	-0.002
	(0.012)	(0.010)	(0.046)**	(0.043)***	(0.035)	(0.033)	(0.022)	(0.021)	(0.017)	(0.017)	(0.057)	(0.055)
Observations	2209	2209	368	368	258	258	2209	2209	2209	2209	373	373
Mean of Dep Var (RR=0)	0.04	0.04	0.21	0.21	0.08	0.08	0.16	0.16	0.11	0.11	0.24	0.24
Controls												
Individual Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Secondary School Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: Boys

	Has had sex with multiple partners		Has a regular partner		Age of partner	Ever had sex		Ever had sex but never used a condom		Used a condom at last sexual intercourse	
MODEL	OLS	PROBIT (ME)	OLS	PROBIT (ME)	OLS	OLS	PROBIT (ME)	OLS	PROBIT (ME)	OLS	PROBIT (ME)
RR Information	0.104	0.097	-0.01	-0.008	-0.373	0.125	0.126	0.062	0.061	0.014	0.014
	(0.031)***	(0.027)***	(0.044)	(0.044)	(0.338)	(0.035)***	(0.036)***	(0.039)	(0.038)	(0.032)	(0.031)
TT on HIV/AIDS curriculum	-0.004	-0.004	0.013	0.013	0.471	-0.023	-0.023	0.003	0.003	-0.013	-0.013
	(0.023)	(0.022)	(0.035)	(0.035)	(0.416)	(0.025)	(0.025)	(0.023)	(0.023)	(0.027)	(0.026)
Observations	2730	2730	1423	1423	291	2730	2730	2730	2730	1438	1438
Mean of Dep Var (RR=0)	0.23	0.23	0.26	0.26	16.09	0.51	0.51	0.35	0.35	0.22	0.22
Controls											
Individual Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Secondary School Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variables are individual-level dummies. Robust standard errors in parentheses, clustered at the secondary school level. Significantly different than zero at 1(***), 5(**) and 10(*) percent. Columns 2, 4, 6, 8, 10 and 12 report mean marginal effects. Individual controls include: age and sampling status. Controls at the secondary school level include: location and gender ratio among pupils. Results are robust to the addition of all other available school controls (school size, average performance, school type (day or boarding), tuition costs).
Columns 3-4 and 11-12: sample restricted to those who declare ever having had sex. Columns 5-6: sample restricted to those who declare having a regular partner.

Table 7 *Cost-Effectiveness of the RR Information Campaign*

	US$
Total Program Cost	$1,911
Total # of pregnancies averted	19.5
Cost per Pregnancy Averted	$98
# of cross-generational pregnancies averted	20
Cost per Cross-Generational Pregnancy Averted	$98
Scenario 1	
# of Primary HIV Infections Averted among Teenage Girls	4.88
Cost per Primary HIV Infection Averted among Teenage Girls	$392
Scenario 2	
# of Primary HIV Infections Averted among Teenage Girls	2.93
Cost per Primary HIV Infection Averted among Teenage Girls	$653
Scenario 3	
# of Primary HIV Infections Averted among Teenage Girls	0.98
Cost per Primary HIV Infection Averted among Teenage Girls	$1,960

Assumption in scenario 1: 25 cases of HIV infection per 100 cross-generational pregnancies

Assumption in scenario 2: 15 cases of HIV infection per 100 cross-generational pregnancies

Assumption in scenario 3: 5 cases of HIV infection per 100 cross-generational pregnancies

Appendix Table A *Status of Study Sample at Follow-up, by Treatment Group*

Panel A. Girls

	(1)	(2)	(3)	(4)	(5)	(6)
	RR Information			TT on HIV/AIDS Curriculum		
	Comparison Group (C)	Treatment Group (T)	Difference T-C	Comparison Group (C)	Treatment Group (T)	Difference T-C
Status unknow (attrited)	0.017	0.013	-0.004	0.018	0.014	-0.005
			(0.004)			(0.004)
Percent repeating class 8	0.235	0.212	-0.023	0.223	0.237	0.014
			(0.022)			(0.018)
Percent in Secondary School	0.461	0.450	-0.011	0.479	0.438	-0.041
			(0.026)			(0.021)*
Percent in Professional Training	0.044	0.052	0.007	0.046	0.045	-0.001
			(0.009)			(0.007)
Percent out of School	0.267	0.295	0.028	0.262	0.285	0.022
			(0.025)			(0.020)
Percent Dead	0.0014	0.0016	0.0002	0.0008	0.0021	0.0014
			(0.0016)			(0.0013)
Observations	5009	1189	6198	3230	2968	6198
Number of Schools	253	71	324	164	160	324

Panel B. Boys

	RR Information			TT on HIV/AIDS Curriculum		
	Comparison Group (C)	Treatment Group (T)	Difference T-C	Comparison Group (C)	Treatment Group (T)	Difference T-C
Status unknow (attrited)	0.015	0.013	0.021	0.014	0.015	-0.002
			(0.022)			(0.019)
Percent repeating class 8	0.219	0.189	0.000	0.211	0.214	0.001
			(0.002)			(0.001)
Percent in Secondary School	0.487	0.498	-0.030	0.491	0.487	0.004
			(0.019)			(0.016)
Percent in Professional Training	0.038	0.033	0.011	0.035	0.039	-0.003
			(0.025)			(0.021)
Percent out of School	0.263	0.283	-0.005	0.268	0.266	0.004
			(0.008)			(0.007)
Percent Dead	0.0006	0.0004	-0.0002	0.0011	0.0000	-0.0011
			(0.0007)			(0.0006)*
Observations	5490	1336	6826	3458	3368	6826

Notes: Status as of July 2005. Study sample composed of students enrolled in 8th grade in 2004. The table present school averages. Standard errors in parenthesis. Significantly different than zero at 1 (***), 5 (**), and 10 (*) percent. Four schools did not have an 8th grade class in 2004 and therefore are exluded from the table.