Relative Risks and the Market for Sex: Teenagers, Sugar Daddies and HIV in Kenya

Dupas, Pascaline

October 2005

Relative Risks and the Market for Sex: Teenagers, Sugar Daddies and HIV in Kenya*

Pascaline Dupas†
Dartmouth College

First Draft: October 2005
This Draft: August 2006

Abstract

An information campaign that provided Kenyan teenagers in randomly selected schools with the information that HIV prevalence was much higher among adult men and their partners than among teenage boys led to a 65% decrease in the incidence of pregnancies by adult partners among teenage girls in the treatment group relative to the comparison. This suggests a large reduction in the incidence of unprotected cross-generational sex. The information campaign did not increase pregnancies among teenage couples. These results suggest that the behavioral choices of teenagers are responsive to information on the relative risks of different varieties of a risky activity. Policies that focus only on the elimination of a risky activity and do not address risk reduction strategies may be ignoring a margin on which they can have substantial impact.

*I am extremely grateful to Esther Duflo, Michael Kremer and Rohini Pande for their generous advice and support. I also wish to thank Thomas Chaney, David Evans, James Habyarimana, Seema Jayachandran, Debraj Ray, Jonathan Robinson, Kudzai Takavarasha, Chris Udry, as well as seminar participants in various seminars, for very helpful comments and discussions. Special thanks to ICS Africa and the Ministry of Education for their collaboration. This project would have been impossible without the dedication of Grace Makana, Susan Nduku, and Carolyne Nekesa, as well as the entire HIV field and data entry teams in Busia Kenya. This project was supported by a grant from the World Bank. All views expressed are mine, and do not necessarily reflect the opinions of the World Bank. All errors are my own.

†Economics Department, Dartmouth College, 318 Rockefeller Hall, Hanover, NH 03755. E-mail: pascaline.dupas@dartmouth.edu.

1 Introduction

Public health interventions often aim at the complete elimination of a risky behavior. Accordingly, they often do not include information that would help people reduce their exposure to risk *within* the practice of a risky activity. This may be due to the fear that clients would opt for the low-risk practice of the activity rather than for complete abstinence. If, however, risk behaviors are more elastic with respect to which of the different practices of a risky activity to engage in, rather than with respect to whether to engage in the activity at all, then policies that focus only on the elimination of a risky activity may be ignoring an important margin along which people could reduce their exposure to risk. There has, so far, been little evidence on the relative size of these two elasticities. This paper presents empirical evidence in one setting: HIV prevention among teenagers in a high-prevalence environment.

In sub-Saharan Africa the prevalence of HIV is considerably higher among young women than among young men. In Kenya and Zambia, for example, prevalence in the 15-19 age group has been found to be at least five times higher among the girls (Glynn et al. 2001).[1] Recent studies suggest that this discrepancy is due, in part, to the high incidence of consensual, unsafe cross-generational sexual relationships — that is, unprotected sex between teenage girls and adult men five or more years their senior.[2] 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 a relationship leads to pregnancy (Bergstrom and Bagnoli 1993). 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 provide information only on the average 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 adult men are more likely to have HIV than younger men.

In this paper, I argue that, when HIV prevalence among men is correlated with age, and as a consequence with income, providing information on the average HIV prevalence in the population without disaggregating the prevalence by age group may not reduce the risk for young women. Providing information on the average HIV risk increases the implicit price of

[1] This in part because the risk of male-to-female HIV transmission is greater than the risk of female-to-male transmission (Peterman et al. 1988). This biological factor, however, accounts for only part of the gap observed (Gregson et al. 2002).

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

unprotected sex, causing men who are poor but low-risk (teenage boys) to be driven out of the market, and thus increasing the proportion of richer but riskier men (adult men) in the pool of potential partners. Consequently, fewer partnerships may be formed, but each of them is, on average, riskier than those formed in the absence of information on HIV.

Providing information on HIV prevalence disaggregated by gender and age group, on the other hand, could help young women reduce their risk. Information on relative risks among potential partners would enable young women to ensure that they receive full compensation for the actual risk they assume by engaging in unprotected sex with adult men, rather than just for the average risk, which is lower. Those adult men who are unable to compensate young women for the full risk would be driven out of the market, reducing the incidence of unprotected cross-generational partnerships, and along with it new HIV infections among young women. The total amount of sexual activity might increase, however, if teenage girls who learn that sex with teenage boys is relatively safer decide to engage in unprotected sex with them, instead of remaining abstinent. This might have negative public health consequences, both in terms of teen pregnancies and in terms of lifetime HIV risk and its epidemiological implications. This may be why most HIV prevention campaigns aimed at youths choose not to provide information on relative risks.

I test the effects of information on relative risks using data from a randomized experiment conducted in 328 primary schools in Kenya. An information campaign conducted by an NGO in 71 randomly selected schools provided teenagers with information on the prevalence of HIV disaggregated by age and gender group. Teenagers in the remaining 257 schools did not receive this information and serve as the comparison group. The randomized design ensured that there would be no systematic difference in the prior information held by the students across the treatment and comparison groups before the start of the information campaign. This allows me to identify the impact of the additional information by comparing behaviors and outcomes across groups 6 to 12 months after the intervention.

I find that the information campaign led to a 65% decrease in the incidence of pregnancies by adult partners among teenage girls in the treatment group relative to the comparison, suggesting a large reduction in the incidence of unsafe, cross-generational sex. The information campaign also led to an increase in self-reported sexual activity among teenagers with partners from within their cohort, suggesting a substitution effect. But there was no increase in pregnancies among teenage couples, since condom use also increased. Data to measure the exact impact on the incidence of HIV infections is not currently available, but a long-term follow-up using biomarkers of sexually transmitted infections could be conducted in the future.

These results show that the behavioral choices of teenagers are responsive to information on the relative riskiness of potential partners. Overall, the information campaign led to an increase in reported sexual activity, but to a decrease in unsafe sex, suggesting 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. Thus, discouraging teenagers from engaging in high-risk activities might be more effective than only asking them to abstain. These results suggest that policies that focus only on eliminating a risky activity without discussing risk reduction strategies may be ignoring a margin on which they could have substantial impact.

Furthermore, I estimate the role of social interactions by exploiting the random variation across secondary schools in the proportion of students coming from a treatment primary school. This design allows me to overcome the endogeneity problem usually inherent to identification of peer effects (Manski, 1993 and 1995; Evans et al., 1992). I find evidence that the sexual behavior of teenage girls is affected by the behavior of others around them, suggesting the presence of peer effects or social norms leading to conformity in behavior, as in Hofferth (1987).

The remainder of the paper is organized as follows. Section II describes the Kenyan context and the information campaign. Section III derives the predictions of a simple model of sexual partnership formation in the presence of risk heterogeneity and incomplete information. Section IV describes the data and the estimation strategy. Section V presents the results and the cost-effectiveness of the program. Section VI discusses spillover and peer effects, and Section VII concludes.

2 Background and Program Description

2.1 HIV prevalence in Kenya

The principal mode of transmission of HIV in Kenya is heterosexual contact (Baltazar et al. 2001). The 2003 Kenya Demographic and Health Survey (KDHS), which included HIV testing of survey respondents, 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.

2.2 HIV-related knowledge among teenagers

Knowledge of the existence of HIV and its modes of transmission is widespread. Ninety percent (90%) of Grade 8 pupils mention HIV/AIDS when prompted for sexually transmitted diseases and 80% declare having a relative who has HIV or AIDS or has died from AIDS (Duflo, Dupas, Kremer and Sinei 2005). Knowledge of risk reduction strategies, however, is more limited. Only 45% of girls and 66% of boys know that condoms can protect from HIV infection, and only 29% of girls know that older men are more likely than young men to infect them with HIV (Table 1, Panel A).[3]

2.3 Sexual partnerships

At the baseline of this study in 2004, 21% of girls (aged 15.1 on average) and 48% of boys (aged 15.5 on average) enrolled in Grade 8, the final year of primary school, reported having had sex (Table 1, Panel A). Around 6.4% of teenage girls enrolled in Grade 8 in a given year will have begun childbearing by the end of the following year (Panel B). Though the great majority (88%) of teen pregnancies are unplanned, the male partner will usually marry the girl, or at least support her financially, in the event of pregnancy (Panel C). The older the male partner, the more likely pregnancy results in a marriage. While the rate of marriage is 42% if the teenage girl gets pregnant by a man 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 (Panel C). In three quarters of teen pregnancies, the partnership involves regular cash transfers from the male to the female partner even before the pregnancy (Panel C).

Overall, 44% of teenage pregnancies are among girls whose partner is more than 5 years older, and 10% among girls whose partner is more than 10 years older (Panel C). However, 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, which suggests that, compared to teenage boys, adult men are more willing to pay for unprotected sex with teenage girls. Thus, even though 44% of observed pregnancies involve adult male partners, it is likely that less than 44% of sexually active teenage girls actually have adult partners. When asked for the age of their partners, only 2% of Grade 8 girls report having a partner older than 20, while 21% of girls report being sexually active (Panel A). This suggests that about 10% of sexually active Grade 8 girls are involved in a sexual relationship with an adult partner. It is possible, however, that teenage girls underreport their involvement with older men.

[3] This last observation is consistent with a study conducted by Longfield et al.(2002), which found that young Kenyan women involved with older men believe that "*older men are low-risk partners because they are less likely to be promiscuous and more likely to remain faithful to younger partners and wives*".

2.4 School-based HIV education

The Kenya Ministry of Education, Science and Technology (MoEST) has integrated HIV/AIDS education into the primary school curriculum. It includes information on the biology of HIV/AIDS, its transmission channels, the consequences of the epidemic for families, schools, and the nation at large, and how to care for people with AIDS. In addition, it includes a prevention section that emphasizes abstinence until marriage, moral values, and refusal skills. The curriculum provides only limited scope for discussing contraception or safer sex in response to student questions. The curriculum does not include discussion of partner selection, and, while love relationships between same-age boys and girls are discussed, cross-generational relationships (and their associated risk) are not mentioned in the textbooks. The proposed strategies to avoid infection are to "Avoid Sex" and to "Say NO to sex before marriage".[4] All sexual activity outside of marriage, irrespective of the age of the partner, is thus considered equally risky.

2.5 The Relative Risks Information Campaign

To assess the elasticity of sexual behavior to information on the distribution of HIV by group, I evaluate the impact of a program run by the NGO International Child Support Africa (ICS) in Kenya. The program involved around 13,000 pupils initially enrolled in Grade 8 in 2004. Out of a pool of 328 primary schools, 71 primary schools were chosen to participate in a Relative Risks Information Campaign (also called the "Sugar Daddy Awareness Campaign"). Between June and November 2004, a trained ICS officer visited each of these schools and, with the authorization of the teachers, spoke to Grade 8 pupils 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.[5] After the survey, students were shown a 10-minute educational video on the risk of pregnancy associated with partnerships with adult men.[6] The video screening was followed by an open discussion about cross-generational sex. During the discussion, the ICS 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.[7] In particular, the ICS 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, as published by the WHO

[4] 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).

[5] The results of this survey were discussed in Section 2.3 and are presented in Table 1, Panel A.

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

[7] In accordance with the Kenyan government policy, ICS officers did not volunteer information on condoms nor demonstrate how to use condoms.

in 1997 and reported in the Kenyan Government's brochure *AIDS in Kenya* (Baltazar et al, 2001).[8]

The 328 schools involved in this study are also involved in a separate randomized evaluation designed to test the effectiveness of Kenya's national HIV prevention curriculum for primary schools.[9] In that study, the sample of schools was stratified by location, test scores, and gender ratio, and half of the schools were randomly chosen to receive teacher training on the HIV curriculum in 2003. The training was implemented jointly by ICS, the Kenya Institute of Education and the MoEST.[10] In 2004, the 71 schools chosen for the Relative Risks Information Campaign were selected randomly after stratifying by participation in the HIV teacher training program, as well as by location, test scores and gender ratio. Thus, the Relative Risks Information Campaign was implemented both in schools where teachers had been trained and in schools where teachers had not been trained, which allows me to compare the impact of the Relative Risks Information Campaign with the impact of the teacher training reinforcement on the national HIV prevention curriculum. Information on the distribution of HIV infections by age and gender is typically not given to adolescents by their teachers because it is not covered by the school curriculum and was not included in the teacher training. Moreover, male teachers are sometimes involved in cross-generational relationships themselves, and may not be in a position to discuss the issue freely with their students.

3 Partner's choice and risk information: a simple model

This section uses a simple matching model with competition between men to predict and compare the impact of providing two types of information: 1) the average prevalence of HIV in the population, and 2) HIV prevalence disaggregated by gender and age group.

I consider a market for sex with three groups: teenage girls, teenage boys and adult men. There are two risks associated with sex: the risk of HIV infection for those who are not infected, and the risk of pregnancy for girls. Both risks can be reduced by using condoms.

[8] 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 city of Kisumu is the capital of Nyanza province, which is predominantly Luo, but the NGO program 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.

[9] This other study is conducted by Esther Duflo, Michael Kremer, Samuel Sinei and myself. See Duflo et al. (2006).

[10] The KIE and the MoEST have trained a number of trainers to provide in-service courses for teachers on HIV/AIDS education methodology. The training is phased-in over a large period of time, which allowed randomization at a large scale.

In the model, a teenage girl randomly meets both a teenage boy and an adult man, but she can start a sexual partnership with only one of them. The two males compete via prices; that is, via the transfer to the girl in compensation for sex. Adult men differ from teenage boys in three ways. First, their group may have a different HIV prevalence rate. Second, they have greater financial resources. Third, they derive relatively less utility from condom-protected sex, an assumption I motivate below.

I solve for the types of partnerships formed, for the use of condoms, and for minimum transfers to girls, under various sets of information. I show that information that increases the perceived average HIV prevalence might have a perverse effect on the riskiness of partnerships formed if it is not accompanied by disaggregated information on the prevalence in each group. However, for certain distributions of the parameters of the model, disaggregated information might have a perverse effect on the amount of within-cohort sexual activity. An empirical test is thus needed to resolve the type of information to provide in order to maximize public health benefits.

3.1 Base case: Partnerships chosen by virgins

I consider the decision made by teenage girl G who meets teenage boy B and adult man A.

Assumptions

Everyone holds the same beliefs about each person's preferences and HIV status, and about the effectiveness of condoms and the risk of pregnancy.

A person who gets infected with HIV incurs a cost H, but a person who infects his or her partner incurs no cost. I ignore the harm caused by reinfection - that is, I assume that having sex with an infected person is harmful only to the uninfected people.

Girl G is a virgin and everyone knows that her HIV risk is zero: $\tilde{p}_G = 0$. No one knows the HIV statuses of boy B and adult man A, but everyone holds common beliefs about them, namely, the prevalence of HIV is $\tilde{p}_B$ among teenage boys and $\tilde{p}_A$ among adult men. For simplicity, I assume that transmission rates from male to female or from female to male are equal. I call the transmission rate t. For person i who has unprotected sex with person j, the expected health cost is $(1 - \tilde{p}_i)\tilde{p}_j tH$.

While the HIV risk is symmetric across gender, the cost of pregnancy is borne entirely by women, in the absence of transfers by male partners). If G has unprotected sex, the expected pregnancy cost is C (equal to the probability of pregnancy times the cost of rearing a child).

I assume that adult A is richer than both boy B and girl G, and that there is diminishing marginal utility of money. As such, even if all three individuals derive the same pleasure from unprotected sex, A's willingness to pay for an unprotected sexual partnership is higher than B's and G's. I call W_A the willingness to pay of A for unprotected sex, and W_Y the willingness of

G and B to pay for unprotected sex ($W_Y < W_A$). Note that everyone's reservation utility is 0.

Condoms exist and reduce both the risk of HIV transmission and the risk of pregnancy by a factor e ($0 < e < 1$).[11] For A and B (but not G), condom use reduces the utility from sex. I make the simplifying assumption that adult men who select themselves into the "teenage" market have an outside option – for example, protected sex with professional sex workers or unprotected sex with older women, – that brings them a higher utility than protected sex with teenage girls.[12] Thus A derives zero utility from protected sex with G. On the other hand, B derives positive utility from protected sex: condom use reduces B's utility from sex by $D < W_Y$. Furthermore, I assume that $W_Y < (1-e)C$, or that G is unwilling to transfer money to either A or B for protected sex.

Partnerships

G's payoff from an unprotected partnership with A is $W_Y - \tilde{p}_A tH - C + P^A$, where P^A is the transfer that A makes to G. The transfer that makes G indifferent between unprotected sex with A and remaining a virgin is $\tilde{p}_A tH + C - W_Y$. Thus A and G may have an unprotected partnership as long as $W_Y + W_A \geq C + \tilde{p}_A tH$. Similarly, G and B may have an unprotected partnership as long as $2W_Y \geq C + \tilde{p}_B tH$. Last, G and B may have a condom-protected partnership as long as $2W_Y - D \geq (1-e)(C + \tilde{p}_B tH)$.

Comparative Statics

I solve for the partnerships formed under three different sets of information:[13] "No HIV," where people believe that there is no HIV; "Average HIV," where people have information on the mean prevalence in the population; and "HIV by Group," where people have information on the prevalence in each group, namely, $\tilde{p}_A > p > \tilde{p}_B$.

Set 1: No HIV: People believe that the prevalence of HIV in the population is 0. Then:

(i) G engages in an unprotected partnership with A as long as his willingness to pay is higher than the sum of B's willingness to pay for condom-protected sex and the extra pregnancy risk: $W_A \geq W_Y - D + eC$.

(ii) If A's willingness to pay is below the threshold given in (i), B's willingness to pay for unprotected sex is also below the threshold given in (i), and G has protected sex with B if $2W_Y \geq (1-e)C + D$. Otherwise she will choose to remain a virgin.

Set 2: Average HIV: People believe that the average HIV prevalence in the population is p,

[11] While I model e as the actual rate of protection, it is only the perceived rate of protection that will affect behavior. The perceived e may be lower than the actual e.

[12] Imagine that there are two types of adult men, those who like condoms and those who don't like condoms. Those who don't like condoms are more likely to look for partners whose risk of HIV infection is low (ideally, virgins), and thus to be on the "teenage" sex market.

[13] Because it is not important for my purpose, I do not model the bargaining game between two partners in an incentive-compatible partnership, and so do not solve for the division of the surplus between them (i.e., whether the adult man makes a transfer equal to the girl's reservation price or larger).

and assign this average to both adult men and teenage boys; that is, $\tilde{p}_A = \tilde{p}_B = p$. *Then*:

(i) G engages in an unprotected partnership with A if A's willingness to pay is large enough for an unprotected partnership to be welfare improving ($W_A + W_Y \geq C + ptH$), and large enough to outweigh the payoff that G could get from protected sex with B ($W_A > W_Y - D + eC + eptH$). This means that G engages in an unprotected partnership with A as long as the perceived prevalence p is below the threshold $\bar{p}_1 = \min\{\frac{W_Y + W_A - C}{tH}, \frac{W_A - W_Y + D - eC}{etH}\}$.

(ii) If the prevalence p is higher than the threshold $\bar{p}_1$, G engages in a protected partnership with B as long as the perceived average prevalence is lower than the threshold $\bar{p}_2 = \frac{2W_Y - D - (1-e)C}{(1-e)tH}$.

Set 3: HIV by group: *People believe that the prevalence among adult men is* $\tilde{p}_A > p$ *and the prevalence among teenage boys is* $\tilde{p}_B < p < \tilde{p}_A$. *Then:*

G discriminates in price (she requests a higher minimum compensation from A than from B) and:

(i) G engages in an unprotected partnership with A as long as the perceived prevalence among adult men $\tilde{p}_A$ is below the threshold $\bar{p}_3 = \min\{\frac{W_Y + W_A - C}{tH}, \frac{W_A - W_Y}{tH} + \tilde{p}_B, \frac{W_A - W_Y + D - eC}{tH} + (1 - e)\tilde{p}_B\}$.

(ii) If the perceived prevalence among adult men $\tilde{p}_A$ is higher than the threshold $\bar{p}_3$, G engages in an unprotected partnership with B as long as $\tilde{p}_B$ is lower than the threshold $\bar{p}_4 = \min\{\frac{2W_Y - C}{tH}, \frac{D - eC}{etH}\}$

(iii) If the perceived prevalence among adult men $\tilde{p}_A$ is higher than the threshold $\bar{p}_3$, and if the perceived prevalence among teenage boys $\tilde{p}_B$ is higher than the threshold $\bar{p}_4$, G engages in a protected partnership with B as long as $\tilde{p}_B$ is lower than the threshold $\bar{p}_2$ (as defined above).

Considering that $\bar{p}_1$ and $\bar{p}_3$ are increasing in W_A, we can infer the following proposition:

Proposition 1 *For any set of true prevalence rates* p_A *among adult men,* p_B *among teenage boys, and* p *on average such that* $p_B < p < p_A$*, we can solve for thresholds* $\bar{W}_L(p)$ *and* $\bar{W}_H(p_A, p_B)$ *in A's willingness to pay such that* $\bar{p}_1(\bar{W}_H) = p_A$ *and* $\bar{p}_3(\bar{W}_L) = p$*, and predict that:*

(i) If adult man A's willingness to pay is sufficiently large $(W_A \geq \bar{W}_H)$*, neither information on the average HIV prevalence nor information on HIV prevalence by group can deter girl G from engaging in unprotected sex with A.*

(ii) If A's willingness to pay is sufficient to compensate for the average risk but not sufficient to compensate for the true risk associated with cross-generational partnerships $(\bar{W}_L < W_A < \bar{W}_H)$*, information on the average HIV prevalence does not deter G from engaging in unprotected sex with A, but information on HIV prevalence by group does.*

(iii) If A's willingness to pay is sufficiently low $(W_A < \bar{W}_L)$*, both information on the average HIV prevalence and information on HIV prevalence by group deter girl G from engaging in unprotected sex with A.*

The proof is presented in the appendix. Proposition 1 implies that, as the perceived average HIV prevalence increases, only relatively richer men are able to compensate for the increased risk, which decreases the total number of unprotected sexual partnerships (see Figure 2). However, if richer men are also more likely to have HIV, each partnership formed is, on average, riskier than before. (Pool B on Fig. 2 is smaller than Pool A, but has a higher average HIV prevalence). In this context, informing teenage girls that the real risk associated with those men is higher than the average risk improves their welfare: girls who have information on HIV rates by group will not enter an unprotected relationship with an adult man unless they are fully compensated for the associated HIV risk. If adult men cannot offer adequate compensation for the risk (which is possible since the risk is high), girls decide not to enter the market for unprotected sex with adult men. However, providing information on HIV rates by group might not always be optimal from the point of view of a public health-maximizing policy marker, as suggested by the following Proposition:

Proposition 2 *Providing information on the average HIV prevalence never triggers teenage girl G to engage in unprotected sex with teenage boy B. However, providing information on HIV prevalence by group causes G to engage in unprotected sex with B if the following three conditions hold:*
(i) adult man A's willingness to pay is not sufficient to compensate for the full HIV risk $\left(W_A < \bar{W}_H\right)$
(ii) G's and B's willingness to pay for unprotected sex is sufficiently large $\left(W_Y > \frac{\tilde{p}_B tH + C}{2}\right)$
(iii) B's disutility from condom use is sufficiently large compared to the effectiveness of condoms $\left(\frac{D}{e} > \tilde{p}_B tH + C\right)$

The proof is straightforward and derives directly from Proposition 1 and the analysis of behaviors under Set 3 presented above. Proposition 2 implies that, if people know the average prevalence of HIV in the population, providing additional information on HIV rates by group might increase the amount of sexual activity between teenage girls and teenage boys, since it leads girls to revise downwards their beliefs about the risk associated with teenage boys.

This effect on within-cohort sexual activity might have negative health consequences. While individuals fully internalize the cost of getting infected with HIV themselves, they do not 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.

Proposition 3 *A public health maximizing policy maker may want to provide information on the average HIV prevalence but withhold information on HIV rates by group if A's willingness to pay is insufficient to compensate for the average risk* $\left(W_A < \bar{W}_L\right)$*.*

Proof. If $W_A \geq \bar{W}_H$, G engages in unprotected sex with A under both sets of information. However, providing information on HIV prevalence by group enables G to discriminate in price and thus to extract full compensation for the risk she is taking, while providing information on the average risk enables her to extract compensation only for the average risk, which is lower. Thus providing HIV prevalence by group improves G's economic welfare without increasing her risk. If $\bar{W}_L < W_A < \bar{W}_H$, providing information on HIV prevalence by group deters G from engaging in unprotected sex with A. Consequently, G might switch to unprotected sex with B, but her risk decreases compared to when she only receives information on the average HIV prevalence, which does not deter her from engaging in unprotected sex with A. Thus providing HIV prevalence by group decreases G's risk. If $W_A < \bar{W}_L$, G refuses to engage in unprotected sex with A under both sets of information. However, providing information on HIV prevalence by group increases the likelihood that G engages in unprotected sex with B, thus increasing her health risk compared to providing information only on the average HIV prevalence. ■

Thus, for a given set of prevalence rates (p, p_A, p_B), the overall impact of providing information on HIV prevalence by group compared to providing only information on the average HIV prevalence will depend on the distributions of W_A, W_Y, D and C, that is, on the proportions of matches in each case highlighted above. In Section 5, I examine these comparative statics empirically in Kenya and measure the effects of HIV-rate information on the incidence of cross-generational sex, condom use, and transfers to girls.[14]

3.2 Extensions

3.2.1 Predictions for non-virgins

Non-virgin girls who receive information on prevalence rates will update their beliefs about their own probability of infection to $p_G = 1 - (1 - t\tilde{p}_A)^n - (1 - t\tilde{p}_B)^m$, where n is the number of unprotected intercourse with past adult partners and m the number of unprotected intercourse with past teenage partners. This means that an increase in the perceived prevalence among potential partners will be mitigated by a decrease in non-virgin girls' cost of HIV infection (since they are more likely to be already infected). In particular, for teenage girls who had adult partners in the past (henceforth Type 1), learning that the true prevalence among adult men ($\tilde{p}_A$) is higher than previously thought might not necessarily increase the marginal cost of engaging in unprotected sex with an adult man. That is, if they believe that their probability of infection is very high (because they have had many adult partners in the past), they might (rationally) become fatalistic and increase their activity on the market for unsafe sex. For those who have had fewer adult partners in the past and do not become fatalistic, the additional information will increase the cost of unprotected sex with adult men but to a lesser extent than

[14] While I do not have data on cash transfers, marriage likely represents a transfer from the man to the woman, so I examine marriage rates. Polygamy is legal in Kenya, and a bride-price is typically paid by the groom to the family of the bride.

for virgins or for girls who only had a teenage partner in the past (Type 0). Thus, providing disaggregated information on HIV rates will have a larger impact on Type 0 than on Type 1 girls, and the pool of teenage girls who are willing to engage in unprotected sex with adult men will become smaller but riskier on average. On the other hand, adult men and teenage boys will also update their beliefs about the riskiness of Type 1 girls, and so decrease their willingness to pay for unprotected sex with them. Overall, the provision of disaggregated information on prevalence rates in the presence of complete information on the sexual histories of girls may have a negative impact on Type 1 girls insofar as they will only be able to match with very risky partners.[15]

3.2.2 Incomplete information on the sexual history of teenage girls

I have so far assumed that everyone knew the sexual histories of teenage girls; that is, whether a particular girl is a virgin (Type 0) or has had adult partners (Type 1). If, instead, I assume that people cannot infer information on the sexual history of a girl from the price at which she is willing to engage in unprotected sex (for instance, because girls are heterogeneous in their disutility for pregnancy), then providing information on the average prevalence will reduce the willingness of teenage boys and adult men to pay for unprotected sex with teenage girls, which is likely to reduce the overall amount of unprotected sexual activity in the population.

In this context, providing disaggregated information could create a market for lemons (Akerlof, 1970). Men who learn that some teenage girls are Type 1 but do not know their identity are less willing to pay for unprotected sex. Type 0 girls will withdraw from the market for unprotected sex if men's willingness to pay falls too low. This will increase the average infection rate in the pool of teenage girls willing to engage in unprotected sex, and thus further reduce men's willingness to pay. Eventually the market for unprotected sex with teenage girls may collapse.

3.2.3 General equilibrium effects

The model looks only at the market for sex with teenage girls. To fully assess the epidemiological consequences of providing information on HIV risk by group, it would be necessary to look at general equilibrium effects. In particular, one should address the question of what becomes of the adult men who are turned down by informed teenage girls. On one hand, they might reduce their sexual activity. On the other hand, they might become fatalistic (Kremer, 1996) and engage in unprotected sex with commercial sex workers, which could have negative epidemiological

[15] If HIV testing is available, high-risk types who turn out negative will be able to match with low-risk partners.

consequences. I do not study this issue in this paper, as I do not have data on adult men's behavior.

4 Evaluation Strategy and Data

4.1 Evaluation Strategy

The randomized design of the information campaign provides a straightforward source of identification. 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 pre-treatment school averages for a series of school and pupils outcomes. Except for class size, which is lower on average in treatment schools, all other differences in pre-treatment averages are small and insignificant. Furthermore, since the information campaign was unannounced and conducted in the middle of the school year, students could not self-select into or out of the treatment group. Thus, the treatment status of a student is not correlated with her observed or unobserved characteristics. Given this, the effect of the information campaign can be measured by comparing outcomes of students in treatment schools to those of students in comparison schools.

4.2 The Sample

The Relative Risks Information Campaign was phased into the 71 treatment schools over 4 months, from July to October 2004.[16] 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 began secondary school in February 2005.

Hereafter I call all students who were enrolled in Grade 8 at the time of the Relative Risks Information Campaign (that is, in 2004) the "treatment cohort", and all students who were either a year ahead or a year below are considered as part of the "control cohorts". I call "treated students" those in the treatment cohort who come from a treatment school. There are about 13,000 teenagers in the treatment cohort, and 2,500 of them were treated. All students in the treatment 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. Table 3 shows information on their status in July 2005, averaged by school and broken down by gender and treatment status. At that time, the schooling status of teenagers in the treatment group was not significantly different to that of teenagers in the

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

comparison group (Column 3). Almost half of the sampled teenagers were enrolled in secondary school, 26% to 29% were out of school, and the others were repeating Grade 8 (Columns 1 and 2).

4.3 The Data

Three types of data were collected: (1) self-reported sexual behavior, (2) childbearing and marital status, and (3) characteristics of sexual partners of the girls who had started childbearing. Data on the HIV status of teenagers in the sample is presently unavailable.

4.3.1 Self-reported sexual behavior

To measure the impact of the information campaign on the behavior of teenagers after they left primary school, a post-survey was administered to all students enrolled in a secondary school in the study area between May and July 2005, about 7 to 9 months after the intervention. 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. Summary statistics on the characteristics of the survey respondents are presented in Table 4. Overall, 55% of the secondary school students who completed the post-survey came from a primary school that participated in the study (10% from a treatment primary school, and 45% from a comparison primary school). Enumerators also collected data on observable school characteristics, such as total enrollment, average performance of students on the national exam, tuition, gender ratio, and location.

Teenagers who joined secondary school may not be 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. For this reason, out-of-school teenagers were asked to come to their former primary school for a half-day in March 2005 to complete a post-survey similar to the post-survey administered in secondary schools. Summary statistics presented in Table 4 show that, as expected, out-of-school respondents were older, came from poorer families and were more likely to be orphans than secondary school respondents. Only 50.2% of the out-of-school teenagers sampled for the survey participated in the survey.[17] Because those who complied are likely to differ from those who did not, the out-of-school survey data is likely to suffer from large selection biases and cannot be use for the evaluation.[18]

[17] Compliance was higher among boys than among girls (56.4% versus 43.3%).

[18] There are two potential selection biases. First, teenagers who left their home village to go and work in town are probably underrepresented, though this selection bias should be similar in the treatment and comparison groups. Second, girls who are pregnant or got married are also likely to be underrepresented. If the treatment affected the incidence of childbearing, this would bias the estimate of the treatment effect on the adoption of safer behaviors upwards (since out-of-school girls who did not get pregnant are more likely to be abstaining or using

In addition to attrition biases, self-reported data on sexual behavior may suffer from reporting biases. Indeed, self-reported sexual behavior 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). Relying solely on self-reported behavior to estimate the impact of the information campaign could thus be misleading, particularly if the campaign changed the patterns of the reporting bias in the treatment relative to the comparison group. For this reason, I look at a biological outcome - childbearing.

4.3.2 Childbearing and marriage status

In addition to being interesting in itself, childbearing is a good proxy for risky sexual activity among primary school girls, as some measures that prevent HIV transmission, such as abstinence and condom use, also prevent pregnancy. Information on childbearing and marital status was obtained for all girls in the sample during three visits to each treatment and comparison school in the 12-month period following the intervention. 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 form, the following questions were asked: Is *Gladys* 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? [19]

Childbearing is a good but imperfect proxy for risky sexual behavior. Certain practices, such as abortion and anal sex, reduce the correlation between the incidence of childbearing and the risk of HIV infection. Since it can be used as a substitute for contraception, abortion is of particular concern for the evaluation of the Relative Risks Information Campaign. Pregnancy by a teenage boy is less likely to result in marriage or child support, as seen in Section 2.3. As such, teenage girls who get pregnant might be more likely to abort if the father of the child is another teenager. If the information campaign increased the incidence of unprotected sex with teenage partners, it might also have increased the incidence of abortion in the treatment group. A comparison of the incidence of childbearing across groups would, therefore, overestimate the program impact on the incidence of unprotected sex. Data on abortion rates is unavailable, so I compare mortality rates among girls across groups. Since abortion is illegal in Kenya, abortions

condoms).

[19] 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 in the treatment and comparison groups.

that do happen tend to be so unsafe that they often result in maternal death.[20] I find that the mortality rate among girls between July 2004 and July 2005 was similar across groups (0.08%). This suggests that the incidence of abortion is not greater in the treatment group than in the comparison group, and thus, that it is reasonable to use childbearing data to estimate the effect of the information campaign on risky sexual behavior.

4.3.3 Age of partners of girls who have begun childbearing: Home Follow-up Visit

Another reason why the incidence of childbearing is not a perfect proxy for the risk of HIV transmission is that an adolescent girl who has a long-term relationship with one partner is less likely to get HIV but more likely to get pregnant compared to an adolescent girl who has several short-term relationships. To better understand the circumstances surrounding teen pregnancy, enumerators conducted a home follow-up visit with female students who had been reported to have started childbearing by July 2005. This home 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. Even though it is self-reported, this data is much less likely to suffer from reporting biases than the behavior survey administered in school for two reasons. First, once a teenage girl is pregnant or has given birth (and so is already out of school), she is much less likely to hide her sexual activity or to hide information about her partner. Second, the interview was conducted by a female enumerator in the confidentiality of the respondent's house.

4.4 Reduced form regressions

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 school s in cohort c. $TreatSchool_s$ is the treatment status of school s and $TreatCohort_c$ the dummy for being in the treatment cohort. $HIVcurriculum_s$ 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 :

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

where I_i is a vector of controls for individual characteristics. The average effect of coming from a treatment school ($TreatSchool_s = 1$) versus a comparison school ($TreatSchool_s = 0$) is captured

[20] 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).

by β_1. Since $TreatSchool_s = 1$ was randomly assigned, we should expect $E(\varepsilon_{is}|TreatSchool_s) = 0$ 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 $HIVcurriculum_s = 1$ was randomly assigned, we should expect $E(\varepsilon_{is}|HIVcurriculum_s) = 0$ so that the ordinary least squares estimator of χ_1 is also unbiased. By comparing β_1 and χ_1, I can thus compare the impact of the two information sets.[21] 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).

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

$$\begin{aligned} Y_{isc} &= \alpha_2 + \beta_2 \times TreatSchool_s \times TreatCohort_c \\ &\quad + \delta \times TreatCohort_c + \theta \times TreatSchool_s \\ &\quad + \chi_2 \times HIVcurriculum_s + I_i'\gamma_2 + \omega_{isc} \end{aligned}$$

Comparing the single-difference to the difference-in-differences estimates is useful for two reasons. First, if the randomization was not perfect, the difference-in-differences will adjust for potential pre-existing random differences between treatment and comparison schools. Second, the difference-in-differences allows the inclusion of school fixed effects (either primary school fixed effects or secondary school fixed effects, depending on the data set), which allows to control for unobservable school characteristics. However, the double-difference estimates could be biased in the presence of treatment spillover across cohorts. This issue should be kept in mind while analyzing the results.

When the outcome is binary, I also estimate the effect of the program on the probability that the outcome occurs using a Logit model:

$$\Pr(Y_{is1} = 1) = [1 + e^{-(\alpha_3 + \beta_3 \times TreatSchool_s + \chi_3 \times HIVcurriculum_s + I_i'\gamma_3 + \varkappa_{is})}]^{-1}$$

5 Results

5.1 Impact on Incidence of Teen Childbearing

Table 5, Columns 1 to 4 show the estimates of the treatment effect on the incidence of childbearing with four different regression specifications: the simple difference with a linear probability

[21] 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.

model (OLS); the simple difference with a Logit model; the OLS estimate of the difference-in-differences; and the OLS estimate of the difference-in-differences with school fixed effects. The information campaign reduced the incidence of childbearing by 1.7 percentage points among treated girls relative to girls in the comparison group (Table 5, Column 1). The childbearing rate in the comparison group is 5.4%, and thus the treatment effect corresponds to a 31% decrease in the incidence of childbearing. The sign and magnitude of the effect are robust to all specifications. In the difference-in-differences without school fixed effects, the estimate of the coefficient for "Treatment School" is close to zero, confirming the absence of ex-ante difference between treatment and comparison schools (Column 3, Row 1).

Table 5, Columns 5 to 12, show estimates of the treatment effect on childbearing broken down by marital status. The bulk of the decrease in the incidence of childbearing in the treatment group corresponds to a decrease in childbearing outside of marriage (unmarried childbearing decreased by 60%), while the incidence of childbearing within marriage decreased only slightly and not significantly. This means that the proportion of girls who are married among girls who started childbearing 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 are consistent with the model presented in Section 3, and suggest that, relative to girls in the comparison groups, treated girls 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.

The teacher-training on the HIV curriculum had no impact on the incidence of childbearing (Table 5, 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. 2005). 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 marrying and having children at a young age.

5.2 Types of Pregnancies Averted

To determine the extent to which the observed decrease in the incidence of childbearing in the treatment group corresponds to a decrease in the incidence of unprotected sex with adult men, I look at the age differentials between girls who have started childbearing and their partners. The data is available for two cohorts: the treatment cohort (Grade 8 of 2004) and one control cohort (Grade 7 of 2004). Since the information campaign reduced the incidence of childbearing in the treatment group, the data is available for differentially selected subsamples of each group. Nevertheless, the ratio of cross-generational to intra-generational pregnancies should be the same across subsamples, unless the treatment had an effect on age differentials between partners.

The dependent variable in Table 6, Columns 1 and 2, is the age difference between the respondent and her baby's father. The treatment effect is negative and significant: among girls who had begun childbearing, the average age gap with the baby's father is 1.75 years smaller for treated girls compared to girls who did not receive the treatment (Column 1). The coefficient in the difference-in-differences specification is even larger (Column 2).[22] 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 (23 percentage points), and significant at the 95% confidence level (Column 3). These results contrast with the insignificant impact of the teacher training on the national HIV curriculum (Row 4).

Table 7 shows the share of cross-generational pregnancies among all averted pregnancies. I consider a normalized case in which 100 pregnancies occur in the comparison group. Of these, 47.6 are by adult men. In the treatment group, we observe 68.6 pregnancies, 16.7 of which by adult men. Thus, the Relative Risks Information Campaign averted 30.9 pregnancies by adult partners in the treatment group. This means that the incidence of cross-generational pregnancies declined by 64.8% in the treatment group relative to the comparison group, while intra-generational pregnancies remained stable. Overall, 98% of averted pregnancies would have been by adult partners.

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 adult men, but did not lead to an increase in the incidence of unprotected sex between teenage girls and teenage boys. This strongly suggests that the information campaign has reduced the teenagers' exposure to the risk of HIV infection. In contrast, the HIV curriculum training had no impact on the incidence of childbearing by adult partners, that is, did not reduce the incidence of unprotected sex with adult men. The differential impact of the two information sets is consistent with the model presented in Section 3, and suggests that adult men have enough resources to compensate teenage girls for the average risk of HIV infection, but not for the actual risk that adult men pose.

5.3 Cost-Effectiveness of the Relative Risks Information Campaign

The information campaign reached about $1,300$ girls and $1,400$ boys in 71 schools and cost less than US$$2,000$ (Table 8, Panel A). The campaign reduced the incidence of childbearing by 1.7 percentage points in the treatment group, which means that a total of 22 $(1,300 \times 1.7)$

[22] Using outcomes for the post-treatment cohort as a baseline group, however, could lead to an overestimation of the effect of the program if the intervention led to a crowding-in of adult men with younger cohorts, as suggested by the coefficients of the "Treatment School" dummy in the difference-in-differences specificiations (Row 1, Columns 2, 5 and 8).

pregnancies were averted thanks to the program. All of these (98%, see Section 5.2) would have resulted from a cross-generational partnership. Thus, the overall cost per cross-generational pregnancy averted is US $86.

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. Instead, I compute cost-effectiveness estimates using three hypothetical ratios: 5/100, 15/100, and 25/100.[23] Panel B in Table 8 shows the cost per HIV infection using these ratios. For a ratio of 15/100, US $86 per cross-generational pregnancy averted corresponds to a cost of US $576 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).

The next step in the evaluation of the epidemiological impact of the information campaign will be to conduct a follow-up study using biomarkers (such as HSV2 and HIV tests) in 2007/2008. By that time, the incidence of sexually transmitted infections in the sample might be high enough to detect a difference between the treatment and the comparison with the current sample size.

5.4 Impact on Self-reported Sexual Behavior

Did the decrease in the incidence of childbearing by adult partners 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 adult men (high-risk option), did they substitute towards teenage boys (low-risk option) or towards abstinence (zero-risk option)? Data to answer these questions is unavailable. As I mentioned in Section 4.1.1, self-reported data on the sexual behavior of out-of-school teenagers could not be collected due to large compliance biases. However, self-reported data are available for teenagers who joined a secondary school in the study area. This subgroup is not representative of all teenagers in the sample.[24] Nevertheless, studying the impact of the Relative Risks Information Campaign on the self-reported sexual behavior of secondary school students is interesting in itself and can provide partial answers to these questions.

The first three columns of Table 9 presents the estimates of the treatment effect on self-

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

[24] 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 make transfers to girls in return for sex.

reported sexual activity. The results of three specifications are shown: the OLS estimate of the simple difference (Column 1); the Logit estimate of the simple difference (Column 2); and the OLS estimate of the difference-in-differences with school fixed effects (Column 3). Panel A shows the coefficient estimates for girls and Panel B shows the estimates for boys.

The Relative Risks Information Campaign led to an increase in the proportion of teenagers who report being sexually active. While 16% of girls coming from comparison schools report ever having had sex, the treatment increased this proportion by 8.6 percentage points, an increase of more than 50% over the comparison group (Table 9, Panel A, Column 1). This result is robust to the logit and to the difference-in-differences specification. In the terms of the model developed in Section 3, this result suggests that the information campaign decreased the perceived riskiness of teenage boys, and therefore some girls who were abstinent before the campaign decided to enter the market for sex with teenage boys.

While 51% of boys coming from comparison schools report ever having had sex, 63% $(0.51 + 0.12)$ of treated boys report ever having had sex (Table 9, Panel B, Column 1). This corresponds to a 24% increase, also robust to all specifications. Assuming that the information campaign led treated boys to update their beliefs on the prevalence of HIV among teenage girls upwards, the increase in the sexual activity of treated boys suggests that boys discriminate between Type 0 and Type 1 girls (as defined in Section 3.2.1), otherwise they would have reduced their willingness to pay for all girls and their overall sexual activity would have decreased.

In contrast, training teachers on the national HIV curriculum decreased the likelihood that teenagers report being sexually active. Girls who were enrolled in a primary school that participated in the training program are 3.8 percentage points less likely to report they ever had sex (Table 8, Panel A, Row 3, Column 3). This difference is significant at the 90% confidence level and corresponds to a 25% decline. Among boys, the effect is also negative, although smaller and not significant. These self-reported differences in behavior, however, are not consistent with the lack of impact on childbearing rates.

Table 9, Columns 4 through 9 show estimates of the treatment effect on sexual activity broken down by ever use of condoms. The increase in sexual activity for both boys and girls does not 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 4 to 6). Rather, the bulk of the increase in sexual activity was matched by an increase in condom use (Columns 7 to 9). Since the Relative Risks Information Campaign neither emphasized nor demonstrated condom use, this result suggests that teenagers know how to use and where to find condoms.

The observed increase in self-reported sexual activity by both treated girls and treated boys could, in part, be due to reporting biases. First, teenagers who openly discussed the issue of sex with an officer from the implementing NGO in 2004 may be more likely to open up when the

NGO officer returns with a survey 6 months later. Second, treated girls may be less likely than other girls to hide that they are sexually active if being involved with a partner from within one's cohort is more socially acceptable than being involved in a cross-generational relationship and if the treatment induced girls to switch from adult partners to teenage partners. Since the treatment did not increase childbearing by teenage couples, the hypothesis of a reporting bias is very plausible. More generally, self-reported data are always susceptible to endogenous reporting biases. This underscores the importance of collecting other data that cannot be manipulated by individuals, such as childbearing rates or biological markers of sexually transmitted infections.

6 Spillover Effects

6.1 Measuring spillover effects

Section 5.4 showed that the information campaign increased self-reported sexual activity among secondary school students who previously attended a treatment school. This change in behavior among the treated might have had an impact on the behavior of their untreated classmates (Hofferth 1987, Sacerdote 2001). To measure the importance of such spillover effects, I use the variation in the density of treated students across secondary schools.

In May 2005 (six months after the information campaign), almost half (46%) of the teenagers in the sample were enrolled in Grade 9, the first year of secondary school. Once in secondary school, they were mixed with students from various primary schools.[25] In each secondary school, I distinguish three types of Grade 9 students: (1) students who previously attended a treatment primary school (T); (2) students who previously attended a comparison primary school (C); and (3) students who attended an unsampled primary school (U) (i.e., a primary school outside the pool of 328 schools included in the study). T and C students are part of the study sample, or baseline, while U students are not in the study sample. For students who did not receive the information on relative risks (C and U students), the intensity of the interaction with students who received the information (T students) can be measured by the proportion $\frac{T}{T+C+U}$. This proportion, however, is likely to be correlated with school characteristics. For example, better performing secondary schools are likely to attract students from far away, and thus may have a higher ratio of non-baseline students (U) over baseline students ($T+C$).

To overcome this problem, I use the proportion $\frac{T}{T+C}$ as an instrument for $\frac{T}{T+C+U}$. Because of the random assignment of primary schools to the treatment and the comparison groups, the variation in the proportion $\frac{T}{T+C}$ can be considered exogenous, assuming that (1) the intervention did not affect the decision to attend secondary school, and (2) the intervention did not affect

[25] The ratio of primary schools to secondary schools in western Kenya is about 4 to 1.

the choice of a secondary school. Information on the current schooling status of teenagers in the sample suggests that the first assumption above is true: at the time of the secondary school survey, 44% of both treated and comparison girls were enrolled in secondary schools (see Table 3). The second assumption is also reasonable: the intervention took place shortly before the KCPE exam (the primary school exit exam), and admission into a specific secondary school depends primarily on the student's score at the KCPE, as well as the financial resources of the parents.

Because the HIV Curriculum Training Program had a small negative impact on sexual activity, I also include the proportions of girls and boys coming from schools that were sampled for the training on the right-hand side (i.e. $\frac{HIV_g}{HIV_g+nonHIV_g+U_g}$ and $\frac{HIV_b}{HIV_b+nonHIV_b+U_b}$). Since the variation in these proportions is not exogenous either, I instrument them with $\frac{HIV_g}{HIVg+nonHIV_g}$ and $\frac{HIV_b}{HIV_b+nonHIV_b}$.

6.2 Spillover Results

The change in behavior among the treated may have impacted the behavior of their peers in two main ways. First, through peer effects: if there is social learning or peer pressure, the behavior of treated students is likely to affect the behaviors of their schoolmates of the same sex. To capture the role of such social interactions, I use the predicted proportion of treated among sampled classmates of the same sex as an independent variable (for girls: $\widehat{\frac{T_g}{T_g+C_g+U_g}}$; for boys: $\widehat{\frac{T_b}{T_b+C_b+U_b}}$).

Second, price (or market) effects: if there is a change in the transfer that treated girls (treated boys) are requesting (resp., willing to pay) for unprotected sex, this will change the probability that boys (resp. girls) who meet treated teenagers engage in sex. To capture these effects, I use the predicted proportion of treated students among sampled classmates of the opposite sex as an independent variable. If, however, there are cross-gender peer effects, the data available do not allow to disentangle these effects from price effects.

The results are presented in Table 10. I report the results of both the OLS and two-stage least squares (IV-2SLS) regressions.

6.2.1 Peer Effects

I find a clear contrast between genders with regards to peer influence. While girls' behavior seems substantially informed by the behavior of other girls, boys do not seem to be affected by the behavior of other boys around them. The results in Column 2 can be interpreted as follows: a teenage girl who is in a school where the share of treated girls is equal to the average (0.14) is 10 percentage points more likely to report having ever had sex than a girl who is in a school with no treated girls (Panel A, Row 1, Column 2: $0.14 \times 0.739 = 0.10$). This difference is significant at the 99 percent confidence level. This effect suggests a high level of conformity in behavior among girls, which can be understood either through a social learning effect (non-treated girls learned that teenage boys are low-risk by observing that treated girls had teenage boyfriends) or as a norm effect (non-treated girls are more likely to have or to report having a teenage boyfriend since it became more common).

In contrast, a teenage boy who is in a school where the share of treated boys is equal to the average (0.12) is 4 percentage points less likely to report having ever had sex, but this difference is not significant (Panel B, Row 2, Column 2: $0.12 \times (-0.329) = -0.04$). The OLS results are indistinguishable from the 2SLS results.

The large conformity in behavior observed among girls applies not only to sexual activity, but also to condom use. While the probability that a girl reports ever having had sex but never having used a condom does not increase significantly with the share of treated girls in her school (Columns 3 and 4), the share of girls who report ever having protected sex is 7 percentage points higher (0.14×0.476) in schools with the average share of treated girls compared to schools with no treated girls (Column 6).

6.2.2 Price Effects

In Table 10, the coefficients of the price effects can be found in Row 2 in Panel A and Row 1 in Panel B.

First we see that treated boys are avoiding sex with untreated girls. Indeed, in schools with a higher proportion of treated boys, girls are less likely to be sexually active (Panel A, Row 2, Column 2). In terms of the model presented in Section 3, this finding suggests that treated boys consider that secondary school girls pose a high risk (that they are Type 1).

Second, in schools with a high share of treated girls, boys are more likely to be sexually active than boys in schools with a low share of treated girls (Panel B, Row 1, Column 2). In a school with the average share of treated girls (0.14), teenage boys are 11 percentage points more likely to have had unprotected sex than teenage boys in a school with no treated girls (Panel B, Row 1, Column 4). This difference is significant at the 90 percent confidence level. This result is also consistent with the model: girls who have received information on relative risks are requesting lower transfers from teenage boys (for unprotected sex) than girls who did not receive the information.

6.3 Spillover Discussion

Given that the information campaign had an indirect impact on the untreated through both peer and price effects within secondary schools, 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 in Section 5.1 is an underestimation of the overall effect of the information campaign. On the other hand, if adult men responded to the change in price charged by treated teenage girls by moving away from treatment schools and towards the surroundings of comparison schools when looking for sex 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. Considering the fact that 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 in Section 5.1.

7 Conclusion

I used a randomized information campaign to study the change in the sexual behavior of Kenyan teenagers in response to information on the relative risk of HIV infection by type of partner. Providing this information led to a 65% decrease in the incidence of pregnancies by the older (riskier) partners among teenage girls in the treatment group relative to the comparison group and triggered an increase in self-reported condom-protected sexual activity among teenagers with partners their own age, suggesting a substitution effect. These results suggest that teenagers are responsive to information on relative risks and that, overall, the information campaign helped teenagers reduce their risk of HIV infection. In contrast, a program that provided general information about the risk of HIV but did not inform teenagers of the risk distribution in the population had no impact on sexual behavior, as measured by pregnancy rates.

The findings of this paper have important 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 abstention 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. But 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 the low-risk variety and no activity at all. The empirical evidence presented in this paper suggests that, in the case of sexual behavior, the former is larger than the latter. This result suggests that HIV education campaigns may achieve a wider health impact if they include both risk reduction and risk avoidance information.

Further work to assess the overall welfare implications of providing information on relative risks will be important in order to allow governments, as well as health and education specialists, to design the most effective information campaigns.

References

[1] Akerlof, George A.(1970). "The Market for 'Lemons:' Quality Uncertainty and the Market Mechanism", *Quarterly Journal of Economics*, August, 84, 488-500.

[2] Baltazar, G. M et al. (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.

[3] Becker, G. and K. Murphy (1988) "A theory of rational addiction". *Journal of Political Economy* 96:675-700.

[4] Bergstrom, T.C. and Bagnoli M. (1993). "Courtship as a Waiting Game", *The Journal of Political Economy*, Vol. 101, No. 1. (Feb), pp. 185-202.

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

[6] 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.

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

[8] Duflo, E., et al. (2006) "Education and HIV prevention: Evidence from Western Kenya", Mimeo, MIT Poverty Action Lab.

[9] Evans, W., Oates W., and Schwab, R. (1992) "Measuring Peer Group Effects: A Model of Teenage Behavior," *Journal of Political Economy*, 966-991.

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

[11] Gersovitz, M. et al, (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.

[12] Gertler, P., Shah, M., and Bertozzi, S., (2005). "Risky Business: The Market for Unprotected Commercial Sex", *Journal of Political Economy* 113: 518-551

[13] Glynn, J.R. et al. (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.

[14] Gregson, S. et al. (2002) "Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe". *The Lancet*; Vol 359.

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

[16] Hofferth, S. (1987) "Factors Affecting Initiation of Sexual Intercourse". In *Risking the Future: Adolescent Sexuality, Pregnancy and Childbearing*, ed. by S. Hofferth and C. Hayes. Washington: Nat. Acad. Press.

[17] Kelly, R. J, et al. (2003). "Age differences in sexual partners and risk of HIV-1 infection in rural Uganda." *Journal of Acquired Immune Deficiency Syndrome*, 32:446-451.

[18] Kremer, M. (1996) "Integrating Behavioral Choice into Epidemiological Models of AIDS." *The Quarterly Journal of Economics*, Vol. 111, No 2: 549-573.

[19] Laga, M., et al. (2001) "To stem HIV in Africa, prevent transmission to young women." *AIDS* 15: 9310934.

[20] Longfield, K., et al. (2002). "Cross-generational relationships in Kenya: Couples' motivation, Risk Perception for STIs/HIV and Condom Use". PSI Research Division, Working Paper No. 52.

[21] 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.

[22] Luke, Nancy. (2005) "Exchange and Condom Use in Informal Sexual Relationships in Urban Kenya," forthcoming, *Economic Development and Cultural Change*.

[23] Luke, N. and K. 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.

[24] Manski, Charles, "Identification of Exogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, LX (1993), 531-542.

[25] Manski, Charles, *Identification Problems in the Social Sciences* (Cambridge: Harvard University Press, 1995).

[26] Miguel, E. and M. Kremer (2004)."Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities " *Econometrica*, 72 (1), 159-217

[27] 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.

[28] Philipson, T. and R. Posner (1993). *Private Choices and Public Health: the AIDS epidemic in an Economic Perspective*, Cambridge: Harvard University Press.

[29] Rao, V., et al. (2003). "Sex Workers and the Cost of Safe Sex: The Compensating Differential for Condom Use Among Calcutta Prostitutes." *Journal of Development Economic* 71(2): 585-604.

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

[31] Sacerdote, Bruce (2001). "Peer Effects With Random Assignment: Results for Dartmouth Roommates." *Quarterly Journal of Economics*, Vol. 116.

8 Appendix: Proof of Proposition 1

The true prevalence rates are: p on average, p_B among teenage boys and p_A among adult men, ranked as follows: $p_B < p < p_A$. The beliefs are denoted: $\tilde{p}$, $\tilde{p}_B$ and $\tilde{p}_A$. I compare partnership formation under three sets of information:

1. In Set 1, the beliefs are $\tilde{p}_A = \tilde{p}_B = \tilde{p}$ and $\tilde{p} < p$, defined such that unprotected sex between G and A dominates protected sex between G and B: $W_A \geq \tilde{p}tH + C - W_Y$ and $W_A \geq e\tilde{p}tH + eC + W_Y - D$. Thus under Information Set 1 G always engages in unprotected sex with A.

2. In Set 2, the beliefs are $\tilde{p}_A = \tilde{p}_B = \tilde{p} = p$.

3. In Set 3, the beliefs are $\tilde{p}_A = p_A$, $\tilde{p}_B = p_B$, and $\tilde{p} = p$.

Proof.

I pose $W_H = \left\{ \begin{array}{l} C + p_AtH - W_Y \quad \text{if } \{2W_Y - C - \tilde{p}_BtH < 0 \text{ and } 2W_Y - (1-e)[C + \tilde{p}_BtH] - D < 0\} \\ (p_A - p_B)tH + W_Y \quad \text{if } 2W_Y - C - \tilde{p}_BtH \geq 0 \text{ and } \tilde{p}_B \leq \frac{D-eC}{etH} \\ (p_A - (1-e)p_B)tH + eC + W_Y - D \text{ if } 2W_Y - (1-e)[C + \tilde{p}_BtH] - D \geq 0 \text{ and } \tilde{p}_B > \frac{D-eC}{etH} \end{array} \right\}$

and I pose $W_L = \left\{ \begin{array}{l} C + ptH - W_Y \quad \text{if } \{2W_Y - C - \tilde{p}_BtH < 0 \text{ and } 2W_Y - (1-e)[C + \tilde{p}_BtH] - D < 0\} \\ W_Y \quad \text{if } 2W_Y - C - \tilde{p}_BtH \geq 0 \text{ and } \tilde{p}_B \leq \frac{D-eC}{etH} \\ eptH + eC + W_Y - D \quad \text{if } 2W_Y - (1-e)[C + \tilde{p}tH] - D \geq 0 \text{ and } \tilde{p}_B > \frac{D-eC}{etH} \end{array} \right\}$

- If $2W_Y - C - \tilde{p}_BtH < 0$ and $2W_Y - (1-e)[C + \tilde{p}_BtH] - D < 0$, G engages in unprotected sex with A if $W_A \geq C + \tilde{p}_AtH - W_Y$ (Inequality 1).
  - Under Information Set 2, Inequality 1 is true only if $W_A \geq W_L$ since $\tilde{p}_A < p_A$.
  - Under Information Set 3, Inequality 1 is true only if $W_A \geq W_H$ since $\tilde{p}_A = p_A$.
  - The minimum compensation that G can extract under set 2 is $P_2 = C + ptH - W_Y$, which is lower that the minimum compensation that she can extract under set 3 ($P_3 = C + p_AtH - W_Y$) since $p < p_A$.

- If $2W_Y - (1-e)[C + p_BtH] - D \geq 0$ and $p_B > \frac{D-eC}{etH}$ and $2W_Y - (1-e)[C + ptH] - D < 0$:
  - Under Information Set 2, G engages in unprotected sex with A if $W_A \geq C + ptH - W_Y$, which is true only if $W_A \geq W_L$ since $p < p_A$.
  - Under Information Set 3, G engages in unprotected sex with A if $W_A - p_AtH - C \geq W_Y - D - (1-e)C - (1-e)p_BtH$, which is true only if $W_A \geq W_H$ since $W_H = (p_A - (1-e)p_B)tH + eC + W_Y - D$.
  - The minimum compensation that G can extract under Set 2 is $P_2 = C + ptH - W_Y$. The compensation that she can extract under Set 3 is $P_3 = (p_A - (1-e)p_B)tH + eC + W_Y - D$. We see that $P_3 > P_2 \iff [2W_Y - (1-e)[C + p_BtH] - D] + p_AtH > ptH - W_Y \Longrightarrow p_AtH > ptH$ since the first term on the left-hand side is positive and the second term on the right-hand side ($-W_Y$) is negative. Since $p_A > p$,we can infer that $P_3 > P_2$.

- If $2W_Y - (1-e)[C + ptH] - D \geq 0$ and $p_B > \frac{D-eC}{etH}$:
  - Under Information Set 2, G engages in unprotected sex with A if $W_A \geq eptH + eC + W_Y - D$ which is true only if $W_A \geq W_L$ since $p_A - p_B - e(p - p_B) \geq 0$.
  - Under Information Set 3, G engages in unprotected sex with A if $W_A \geq (p_A - (1 - e)p_B)tH + eC + W_Y - D$ which is true only if $W_A \geq W_H$.
  - The minimum compensation that G can extract under Set 2 is $P_2 = eptH + eC + W_Y - D$. The compensation that she can extract under Set 3 is $P_3 = (p_A - (1-e)p_B)tH + eC + W_Y - D$. We see that $P_3 > P_2 \iff p_A - (1-e)p_B > ep \iff p_A - p_B > e(p - p_B)$ which is always true since $e \leq 1$ and $p_A > p$.

- If $2W_Y - C - p_BtH \geq 0$ and $p_B \leq \frac{D-eC}{etH}$ and $2W_Y - C - ptH < 0$ and $2W_Y - (1-e)[C + ptH] - D \geq 0$:
  - Under Information Set 2, G engages in unprotected sex with A if $W_A \geq eptH + eC + W_Y - D$ which is true only if $W_A \geq W_L$ since $p_A - p_B - e(p - p_B) \geq 0$.
  - Under Information Set 3, G engages in unprotected sex with A if $W_A - p_AtH - C \geq W_Y - p_BtH - C$, which is true only if $W_A \geq W_H$ since $W_H = (p_A - p_B)tH + W_Y$.
  - The minimum compensation that G can extract under Set 2 is $P_2 = eptH + eC + W_Y - D$. The compensation that she can extract under Set 3 is $P_3 = (p_A - p_B)tH + W_Y$. We see that $P_3 > P_2 \iff (p_A - p_B)tH > eptH + eC - D$. Since $eC - D \leq ep_BtH$, $P_3 > P_2 \implies (p_A - p_B) > e(p - p_B)$ which is always true since $e \leq 1$ and $p_A > p$.

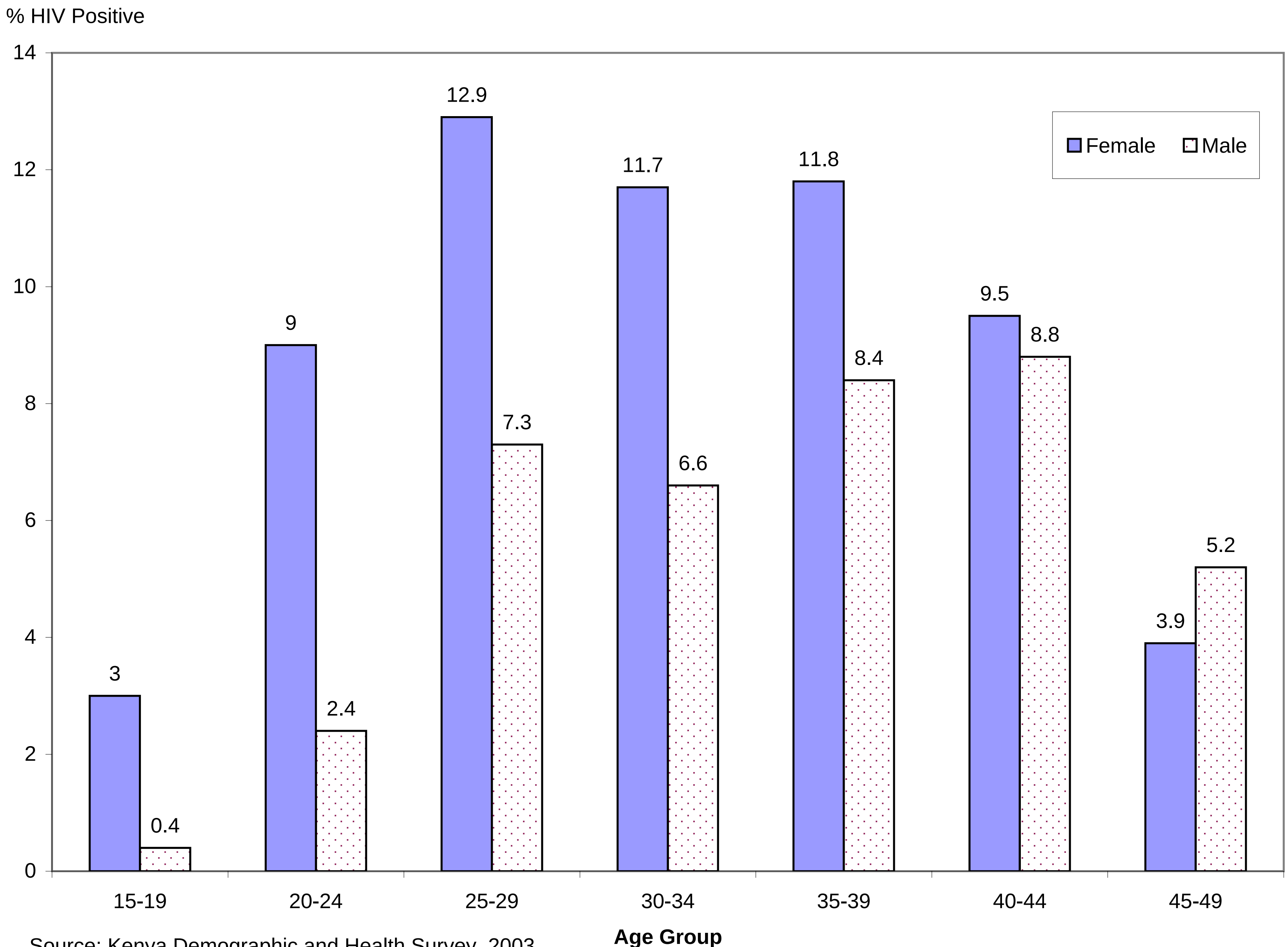
- If $2W_Y - C - p_BtH \geq 0$ and $p_B \leq \frac{D-eC}{etH}$ and $2W_Y - C - ptH < 0$ and $2W_Y - (1-e)[C + ptH] - D < 0$:
  - Under Information Set 2, G engages in unprotected sex with A if $W_A \geq C + ptH - W_Y$ which is true only if $W_A \geq W_L$.
  - Under Information Set 3, G engages in unprotected sex with A if $W_A - p_AtH - C \geq W_Y - p_BtH - C$, which is true only if $W_A \geq W_H$ since $W_H = (p_A - p_B)tH + W_Y$.
  - The minimum compensation that G can extract under Set 2 is $P_2 = C + ptH - W_Y$. The compensation that she can extract under Set 3 is $P_3 = (p_A - p_B)tH + W_Y$. We see that $P_3 > P_2 \iff 2W_Y - C - p_BtH > (p - p_A)tH$ which is always true since $2W_Y - C - \tilde{p}_BtH \geq 0$ and $p - p_A < 0$.

- If $2W_Y - C - ptH \geq 0$ and $p \leq \frac{D-eC}{etH}$:
  - Under Information Set 2, G engages in unprotected sex with A if $W_A \geq W_Y$ which is true only if $W_A \geq W_L$.
  - Under Information Set 3, G engages in unprotected sex with A if $W_A \geq (p_A - p_B)tH + W_Y$ which is true only if $W_A \geq W_H$.
  - The minimum compensation that G can extract under Set 2 is $P_2 = W_Y$. The compensation that she can extract under Set 3 is $P_3 = (p_A - p_B)tH + W_Y$. We see that $P_3 > P_2$ since $p_A > p_B$.

- Thus if $W_A \geq W_H$, G' behavior is identical under all information sets (she always engages in unprotected sex with A). However, since $P_3 > P_2$,the compensation that G can extract under Set 3 is higher than the compensation she can extract under Set 2 (assuming a similar bargaining structure under each set of information).

■

Figure 1
HIV Prevalence by Gender and Age Groups, Kenya



Source: Kenya Demographic and Health Survey, 2003

Figure 2

If low-prevalence men are poor and high-prevalence men are rich, an increase in the average perceived prevalence of HIV infection (*pp*) decreases the incidence of unprotected sex but increases the average riskiness of unprotected sexual partnerships that are formed: the pool of men who are competitive on the unprotected sex market under the belief $pp=p_{bar}$ (Pool B) is more infectious than the pool of men who are competitive under the belief $pp=0$ (Pool A). Under the belief $pp=p_A$, no man can provide sufficient compensation for the risk.

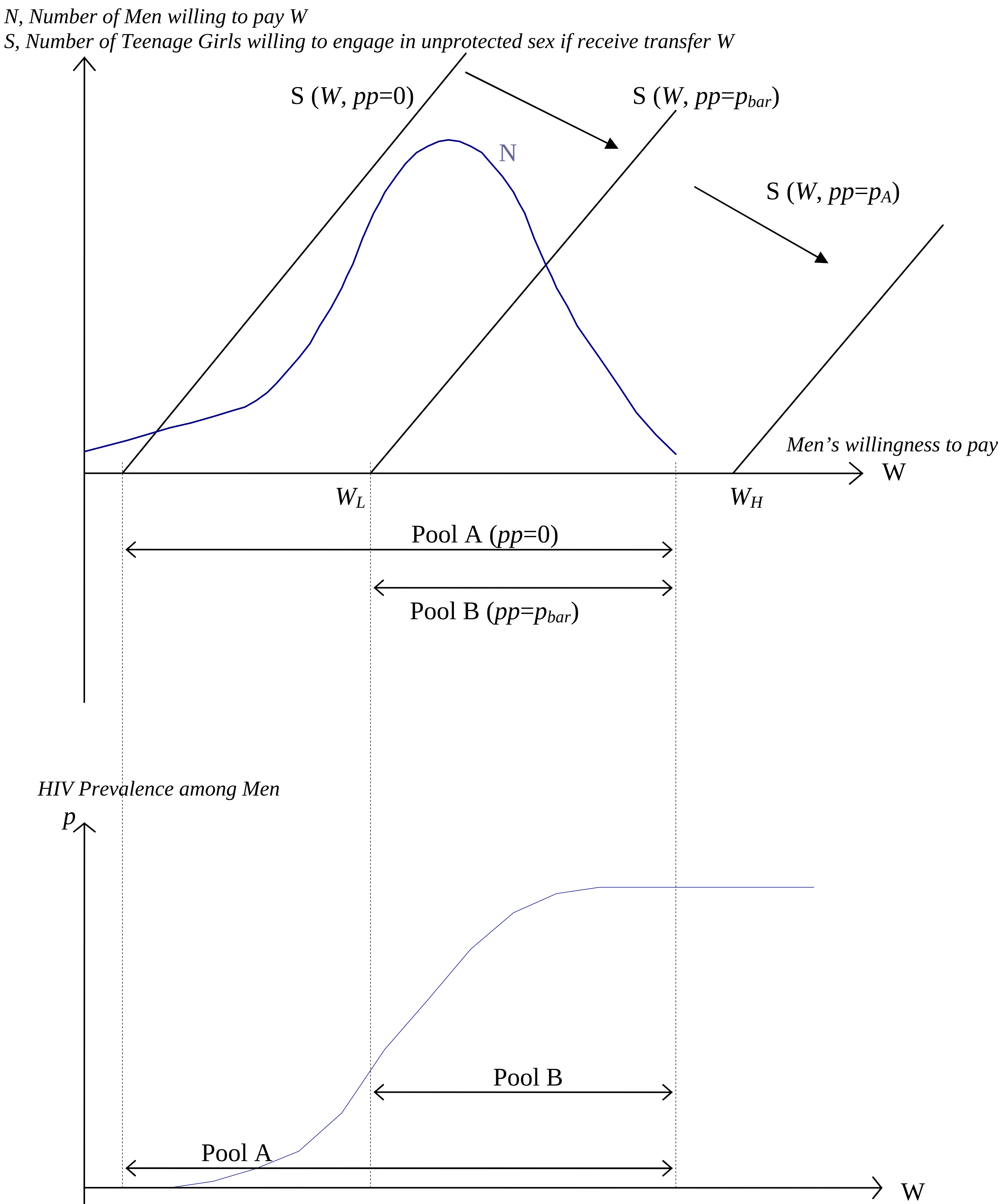


Table 1
Baseline Knowledge and Behavior

	Girls (1)	Boys (2)
Panel A: Knowledge, Attitudes and Practice (Self-reported)		
Average Age	15.1	15.5
	(1.2)	(1.5)
Percent thinking condoms can prevent pregnancy	0.46	0.71
Percent thinking condoms can prevent HIV infection	0.45	0.66
Percent thinking older people are more likely to infect them younger people	0.29	0.25
Percent reporting having had sex	0.21	0.48
Percent reporting that some girls in the class have a partner who is not a student	0.61	0.57
Percent reporting having a partner above 20 years of age	0.02	0.02
Number of observations	1176	1246
Number of schools	69	69
Panel B: Fertility Survey (Reported by former classmates)		
Percent married	0.059	0.007
Percent who has begun childbearing	0.064	N/A
Number of observations	4800	4853
Panel C: Partnership Survey (Interviews with girls who started childbearing)		
Age of teenage girl at time of follow-up (in years)	16.44	
	(1.21)	
Age of male partner at time of follow-up (in years)	22.00	
	(4.23)	
Age difference between partners > 5 years (% "yes")	0.44	
Age difference between partners > 10 years (% "yes")	0.10	
Percent reporting that the partnership was consensual	0.99	
Percent reporting that the pregnancy was wanted	0.13	
Percent reporting that the male partner made regular cash payments to the teenage g prior to the pregnancy	0.76	
If any cash payment was made, average monthly payment made to the teenage girl (in US$)	5.86	
	(7.08)	
Percent reporting that the male partner is currently providing financial support to th teenage girl	0.74	
Percent married if age difference < 5 years	0.42	
Percent married if 5 years <age difference <10 years	0.63	
Percent married if age difference > 10 years	0.79	
Number of Observations	286	

Notes: Standard deviations are presented in parentheses. Panel A: Self-reported data collected among teenagers enrolled in Grade 8 in 2004, prior to the program. The survey was self-administered. Panel B: outcomes for students in Pre-Treatment Cohort (enrolled in Grade 8 in 2003), collected through school visits in the fall 2004. Panel C: Childbearing data collected in August 2005 for teenage girls initially enrolled in a comparison school in 2004, and who had begun childbearing before 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 answered questions on her behalf.

Table 2
Average School Characteristics and Students Outcomes at Baseline, by Treatment Group

	(1) Treatment Group	(2) Comparison Group	(3) Difference T-C
<u>Panel a. School Characteristics</u>			
Class Size	35	38	-3.799**
			(1.55)
Pupils sex ratio (Girls/Boys)	1.11	1.07	0.041
			(0.07)
Teacher-pupil ratio	0.026	0.026	0.000
			(0.00)
Teachers sex ratio (Females/Males)	0.92	1.03	-0.112
			(0.12)
KCPE results (2003)	249	251	-2
			(4)
Received Teacher Training on HIV curriculum	0.493	0.498	
<u>Panel b. Outcomes at the end of 2004: Girls enrolled in 8th grade in 2003</u>			
Percent repeating class 8	0.22	0.25	-0.03
			(0.02)
Percent in Secondary School	0.47	0.46	0.01
			(0.03)
Percent in Professional Training	0.04	0.04	0.00
			(0.01)
Percent out of School	0.27	0.24	0.03
			(0.02)
Percent married	0.073	0.059	0.014
			(0.01)
Percent who has begun childbearing	0.060	0.064	-0.004
			(0.01)
<u>Panel c. Outcomes at the end of 2004: Boys enrolled in 8th grade in 2003</u>			
Percent repeating class 8	0.22	0.23	0.00
			(0.022)
Percent in Secondary School	0.51	0.52	-0.02
			(0.027)
Percent in Professional Training	0.01	0.02	-0.01
			(0.005)
Percent out of School	0.25	0.23	0.03
			(0.022)
Percent married	0.006	0.007	-0.001
			(0.004)
Number of Girls	**1215**	**4800**	**6015**
Number of Boys	**1235**	**4853**	**6088**
Number of Schools	**71**	**252**	**323**

These are school averages. School characteristics collected in 2004. Students outcomes collected in 2004 for the pre-treatment cohort (Grade 8 of 2003). Standard errors in parenthesis. Significantly different than zero at 99 (***), 95 (**), and 90 (*) percent confidence.

Table 3

Current Status of Teenagers in the Treatment Cohort, by Treatment Group

	(1) Treatment Group	(2) Comparison Group	(3) Difference T-C
Panel a. Situation in July 2005: Girls enrolled in 8th grade in 2004			
Percent repeating class 8	0.215	0.238	-0.024
			(0.023)
Percent in Secondary School	0.441	0.441	0.000
			(0.026)
Percent in Professional Training	0.048	0.043	0.005
			(0.009)
Percent out of School	0.291	0.263	0.028
			(0.02)
Panel b. Situation in July 2005: Boys enrolled in 8th grade in 2004			
Percent repeating class 8	0.194	0.222	-0.027
			(0.021)
Percent in Secondary School	0.481	0.463	0.018
			(0.026)
Percent in Professional Training	0.033	0.037	-0.003
			(0.008)
Percent out of School	0.282	0.264	0.018
			(0.023)
Number of Girls	**1279**	**5382**	**6661**
Number of Boys	**1378**	**5721**	**7099**
Number of Schools	**71**	**254**	**325**

These are school averages. Data is missing for 3 comparison schools.
Standard errors in parenthesis. Significantly different than zero at 99 (***), 95 (**), and 90 (*) percent confidence.

Table 4

Characteristics of Teenagers who participated in the Post-Survey on Sexual Behavior

	Secondary School Students				**Out-of-School Youths**	
	Boys		**Girls**		**Boys**	**Girls**
	Treatment Cohort (9th Grade)	Control Cohort (10th Grade)	Treatment Cohort (9th Grade)	Control Cohort (10th Grade)	Treatment Cohort Only	
	(1)	(2)	(3)	(4)	(5)	(6)
Average age (in years)	15.89	16.72	15.38	16.19	17.24	16.34
	(1.28)	(1.22)	(1.04)	(1.04)	(1.44)	(1.29)
Proportion coming from a Treatment primary school	0.10	0.10	0.11	0.11	0.33	0.37
Proportion coming from a Comparison primary school	0.44	0.43	0.47	0.44	0.67	0.63
Proportion coming from a primary school that received teacher training on the HIV curriculum	0.27	0.27	0.28	0.27	0.51	0.52
Proportion coming from a primary school outside the program area (unsampled)	0.46	0.47	0.43	0.45		
Proportion in a single-sex secondary school	0.38	0.36	0.31	0.31		
Proportion member of a health club in primary school	0.23	0.18	0.26	0.21		
Socio-Economic Status						
Proportion sleeping under a bednet at night	0.23	0.23	0.25	0.22	0.10	0.14
Proportion having a radio at home	0.91	0.91	0.94	0.94	0.81	0.89
Proportion having an iron roof at home	0.84	0.87	0.85	0.85	0.75	0.86
Proportion being maternal orphan	0.09	0.09	0.09	0.08	0.14	0.11
Proportion being paternal orphan	0.18	0.20	0.17	0.16	0.23	0.22
Number of Observations	3023	2333	2441	2152	499	330
Number of Secondary Schools	67	65	74	72		

Notes: Self-reported data collected through a self-administered survey, conducted in secondary schools (colums 1-4) and with out-of-school youths on the compound of their former primary schools (col. 5-6). Standard deviations are shown in parentheses.

Table 5
Probability that Teenage Girls have started Childbearing

	Dependent Variable											
	Has started childbearing				**Has started childbearing, Unmarried**				**Has started childbearing, Married**			
SPECIFICATION	SD	SD	DD	DD-FE	SD	SD	DD	DD-FE	SD	SD	DD	DD-FE
MODEL	OLS	LOGIT	OLS	OLS	OLS	LOGIT	OLS	OLS	OLS	LOGIT	OLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment School	**-0.017****	**-0.292**	0.005		**-0.013****	**-0.670****	0.002		**-0.005**	**-0.046**	-0.002	
	(0.009)	**(0.193)**	(0.012)		**(0.005)**	**(0.323)**	(0.009)		**(0.007)**	**(0.204)**	(0.009)	
Treatment School x Treatment Cohort			**-0.024***	**-0.023***			**-0.017**	**-0.018***			**-0.004**	**-0.002**
			(0.015)	**(0.015)**			**(0.015)**	**(0.015)**			**(0.015)**	**(0.015)**
HIV Curriculum	0.005	0.11	0.009		0.005	0.277	0.002		0.000	0.018	0.007	
	(0.006)	(0.134)	(0.006)		(0.004)	(0.195)	(0.004)		(0.005)	(0.154)	(0.005)	
Age	0.020***	0.385***	0.025***	0.025***	0.005***	0.202***	0.008***	0.008***	0.015***	0.467***	0.017***	0.018***
	(0.003)	(0.054)	(0.003)	(0.003)	(0.002)	(0.070)	(0.002)	(0.002)	(0.002)	(0.065)	(0.002)	(0.003)
<u>Sample</u>												
Control Cohort Included (8th Graders of 2003)			Y	Y			Y	Y			Y	Y
<u>Controls</u>												
Individual Characteristics	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Primary School Characteristics	Y	Y	Y		Y	Y	Y		Y	Y	Y	
Primary School Fixed Effects				Y				Y				Y
Observations	5924	5924	11423	11423	5924	5924	11423	11423	5924	3159	11423	11423
Mean of Dep Var in Comp. Group	0.054	0.054	0.054	0.054	0.022	0.022	0.022	0.022	0.032	0.032	0.032	0.032
Odds Ratio (Treatment)		0.747				0.512				0.955		
Odds Ratio (HIV curriculum)		1.116				1.319				1.019		

<u>Notes:</u> Treatment = Relative Risks Information Campaign. Specifications: SD = simple difference; DD = difference-in-difference; FE= school fixed effects. The dependent variables are at individual-level dummies. Robust standard errors in parentheses, clustered at the school level. Significantly different than zero at 99 (***), 95(**) and 90(*) percent confidence. Individual controls include: grade in 2003, current schooling status, 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 pre-Treatment Cohort, and after 12 to 17 months for the Treatment 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 pre-Treatment Cohort and 14.5 months for the

Table 6

Age Gap between Teenage Girls who have started Childbearing and their Partner

	Dependent Variable							
	Age Difference between Teenage Girl and her partner		**Age gap >5 years**			**Age gap >10 years**		
SPECIFICATION	SD	DD	SD	SD	DD	SD	SD	DD
MODEL	OLS	OLS	OLS	LOGIT	OLS	OLS	LOGIT	OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment School	**-1.753****	0.965	**-0.232****	**-1.539*****	0.13	**-0.058**	**-0.759**	0.165**
	(0.773)	(0.839)	**(0.107)**	**(0.585)**	(0.125)	**(0.057)**	**(0.771)**	(0.082)
Treatment School x Treatment Cohort		**-2.531****			**-0.323***			**-0.208****
		(1.126)			**(0.190)**			**(0.105)**
HIV Curriculum	-1.108	-0.476	0.091	0.741	0.033	-0.054	-0.743	-0.016
	(1.076)	(0.484)	(0.082)	(0.469)	(0.062)	(0.056)	(0.630)	(0.038)
Age of Teenage Girl	-0.992***	-0.743***	-0.104***	-0.641***	-0.077***	-0.054**	-0.791**	-0.027
	(0.358)	(0.215)	(0.039)	(0.233)	(0.025)	(0.025)	(0.322)	(0.018)
Sample								
Control Cohort Included (7th Graders of 2004)		Y			Y			Y
Controls								
Individual characteristics	Y	Y	Y	Y	Y	Y	Y	Y
Primary School Characteristics	Y	Y	Y	Y	Y	Y	Y	Y
Observations	120	247	133	121	274	132	103	273
Mean of Dep Var in Comparison Group	6.03	6.03	0.48	0.48	0.48	0.15	0.15	0.15
Std. Dev.	(5.21)	(5.21)						
Odds Ratio (Treatment)				0.21			0.47	
Odds Ratio (HIV Curriculum)				2.1			0.48	

Notes: Treatment = Relative Risks Information Campaign. Specifications: SD = simple difference; DD = difference-in-difference. The dependent variables are at the individual level. Robust standard errors in parentheses, clustered at the primary school level. Significantly different than zero at 99 (***), 95(**) and 90(*) percent confidence. Individual controls include: grade in 2003, cohort when applicable. School controls include: teacher-pupil ratio, gender ratio among teachers, gender ratio among pupils, average school performance at national KCPE exam, location.

Table 7
Overall Treatment Effect on Incidence of Childbearing by Adult Men

	Comparison Group Base=100	Treatment Group	# Averted	Treatment Effect
# Observed Teen Pregnancies	100	68.6	31.4	-31.4%
Share of Observed Pregnancies by Adult Men	48%	24%		-23.2%
# Observed Pregnancies by Adult Men	47.6	16.7	30.9	-64.8%
# Observed Pregnancies by Young Men	52.4	51.9	0.5	-1.0%
Share of Cross-Generational Pregnancies among Averted Pregnancies			98%	

Notes: Treatment = Relative Risks Information Campaign. First row: treatment effect on number of teen pregnancies reported from Table 5 (-0.17/0.53). Second row: treatment effect on share of pregnancies by adult men reported from Table 6, regression (3).

Table 8
Cost-Effectiveness of the Relative Risks Information Campaign

	Kenyan Shillings	**US$**
Panel A: Program Costs		
Program Officer Salary	35,000	467
Video and Power Equipment Rental	17,490	233
Transportation costs	84,000	1,120
Overhead (5%)	6,825	91
TOTAL COST	**143,315**	**$1,911**
Panel B: Cost-Effectiveness		
Total # of pregnancies averted		22
Cost per Pregnancy Averted		$86
# of cross-generational pregnancies averted		22
Cost per Cross-Generational Pregnancy Averted		$86
Scenario 1		
# of Primary HIV Infections Averted among Teenage Girls		5.53
Cost per Primary HIV Infection Averted among Teenage Girls		$346
Scenario 2		
# of Primary HIV Infections Averted among Teenage Girls		3.32
Cost per Primary HIV Infection Averted among Teenage Girls		$576
Scenario 3		
# of Primary HIV Infections Averted among Teenage Girls		1.11
Cost per Primary HIV Infection Averted among Teenage Girls		$1,729

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

Table 9

Sexual Activity: Self-Reported Behavior of Teenagers who joined Secondary School (9th Graders of 2005)

	Dependent Variable								
	Ever had sex			**Ever had sex but never used a condom**			**Ever had sex and ever used a condom**		
SPECIFICATION	SD	SD	DD-FE	SD	SD	DD-FE	SD	SD	DD-FE
MODEL	OLS	LOGIT	OLS	OLS	LOGIT	OLS	OLS	LOGIT	OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PANEL A: GIRLS									
Treatment School	**0.086*****	**0.547*****	-0.006	**0.019**	**0.179**	-0.003	**0.067*****	**0.880*****	-0.003
	(0.029)	**(0.172)**	(0.029)	**(0.025)**	**(0.242)**	(0.024)	**(0.020)**	**(0.216)**	(0.020)
Treatment School x Treatment Cohort			**0.079****			**0.010**			**0.069****
			(0.038)			**(0.031)**			**(0.027)**
HIV curriculum	-0.029	-0.204	-0.038**	-0.025	-0.256	-0.034**	-0.004	-0.05	-0.004
	(0.022)	(0.158)	(0.017)	(0.018)	(0.188)	(0.014)	(0.017)	(0.258)	(0.012)
Observations	2123	2123	3935	2123	2123	3935	2123	2123	3935
Mean of Dep Var (Comparison Group)	0.16	0.16	0.16	0.11	0.11	0.11	0.06	0.06	0.06
Odds Ratio (Treatment)		1.73			1.20			2.41	
Odds Ratio (HIV curriculum)		0.82			0.77			0.95	
PANEL B: BOYS									
Treatment School	**0.121*****	**0.314*****	0.014	**0.058**	**0.155**	0.042	**0.063****	**0.236*****	-0.028
	(0.034)	**(0.089)**	(0.037)	**(0.040)**	**(0.106)**	(0.036)	**(0.025)**	**(0.087)**	(0.030)
Treatment school x Treatment Cohort			**0.111****			**0.018**			**0.093****
			(0.047)			**(0.046)**			**(0.038)**
HIV curriculum	-0.027	-0.071	-0.017	0.01	0.026	0.001	-0.037**	-0.153**	-0.018
	(0.025)	(0.065)	(0.021)	(0.024)	(0.063)	(0.020)	(0.017)	(0.069)	(0.017)
Observations	2603	2603	4474	2603	2603	4474	2603	2603	4474
Mean of Dep Var (Comparison Group)	0.51	0.51	0.51	0.35	0.35	0.35	0.16	0.16	0.16
Odds Ratio (Treatment)		1.37			1.17			1.27	
Odds Ratio (HIV curriculum)		0.93			1.03			0.86	
Sample									
Control Cohort Included (10th Graders of 2005)			Y			Y			Y
Controls									
Individual Characteristics	Y	Y	Y	Y	Y	Y	Y	Y	Y
Secondary School Characteristics	Y	Y		Y	Y		Y	Y	
Secondary School Fixed Effects			Y			Y			Y

Notes: Treatment = Relative Risks Information Campaign. Specifications: SD = simple difference; DD = difference-in-difference; FE= school fixed effects. The dependent variables are individual-level dummies. Robust standard errors in parentheses, clustered at the secondary school level. Significantly different than zero at 99 (***), 95(**) and 90(*) percent confidence. "*HIV curriculum* " is a dummy indicating whether the respondent comes from a primary school that received teacher training on the HIV curriculum. Controls at the secondary school level include: location, average performance, school type (day or boarding), tuition costs, size, gender ratio among pupils. Individual controls include: cohort, socio-economic characteristics, orphan status, religion, participation in health club activities at primary school level, and whether the student comes from a primary school outside the program area.

Table 10
Spillover Effects: Peers of Treated Students who Joined 9th Grade

	Dependent Variable					
	Ever had sex		**Ever had sex, but never used a condom**		**Ever had sex and ever used a condom**	
MODEL	OLS	IV-2SLS	OLS	IV-2SLS	OLS	IV-2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A: GIRLS						
Share of Treated Girls [$T_g / (T_g+C_g+U_g)$]	0.554**	0.739***	0.139	0.263	0.415***	0.476**
	(0.221)	(0.243)	(0.180)	(0.175)	(0.144)	(0.186)
Share of Treated Boys [$T_b / (T_b+C_b+U_b)$]	-0.568**	-0.813***	-0.251	-0.425**	-0.317*	-0.388*
	(0.239)	(0.287)	(0.168)	(0.186)	(0.162)	(0.209)
Share of Girls from HIV curriculum school	0.042	0.058	0.101	0.05	-0.059	0.008
	(0.123)	(0.180)	(0.097)	(0.121)	(0.077)	(0.093)
Share of Boys from HIV curriculum school	0.003	-0.005	-0.034	-0.049	0.037	0.044
	(0.148)	(0.189)	(0.117)	(0.128)	(0.074)	(0.095)
Observations	1346	1346	1346	1346	1346	1346
Mean of Dep Var	0.17	0.17	0.11	0.11	0.06	0.06
PANEL B: BOYS						
Share of Treated Girls [$T_g / (T_g+C_g+U_g)$]	0.736*	0.888*	0.631*	0.779*	0.105	0.109
	(0.394)	(0.529)	(0.352)	(0.458)	(0.136)	(0.185)
Share of Treated Boys [$T_b / (T_b+C_b+U_b)$]	-0.234	-0.329	-0.206	-0.371	-0.028	0.042
	(0.397)	(0.569)	(0.357)	(0.487)	(0.166)	(0.219)
Share of Girls from HIV curriculum school	-0.047	0.280	0.022	0.224	-0.069	0.056
	(0.194)	(0.259)	(0.162)	(0.231)	(0.104)	(0.113)
Share of Boys from HIV curriculum school	0.09	-0.109	0.055	0.009	0.035	-0.118
	(0.153)	(0.187)	(0.141)	(0.155)	(0.072)	(0.098)
Observations	1695	1695	1695	1695	1695	1695
Mean of Dep Var	0.52	0.52	0.36	0.36	0.16	0.16
Controls						
Individual Characteristics	Y	Y	Y	Y	Y	Y
Secondary School Characteristics	Y	Y	Y	Y	Y	Y

Notes: Regressions include students coming from a primary school that was sampled neither for the Relative Risks Information Campaign, nor for the HIV curriculum Training. The dependent variables are individual-level dummies. Robust standard errors in parentheses, clustered at the secondary school level. Significantly different than zero at 99 (***), 95(**) and 90(*) percent confidence.

Controls at the secondary school level include: location, average performance, school type (day or boarding), tuition costs, size, gender ratio among pupils, average age of pupils in the calss. Individual controls include: cohort, socio-economic characteristics, orphan status, religion, participation in health club activities at primary school level, and whether the student comes from a primary school outside the program area.

The instrumental variables in regressions (2), (4) and (6) are the proportions of treated students among students coming from a sampled primary school by gender: Tg/(Tg+Cg) and Tb/(Tb+Cb), the proportions of students coming from schools that received HIV curriculum training by gender: HIVg/(HIVg+non-HIVg) and HIVb/(HIVb+nonHIVb), and the other explanatory variables.

In coed schools, the average Share of Treated Girls is 0.14 (std. dev: 0.12) and the average Share of Treated Boys is 0.12 (0.11). In single sex schools, the average Share of Treated Girls is 0.10 (0.09) and the average Share of Treated Boys is 0.08 (0.04).