Male Circumcision, Peer Effects, and Risk Compensation[*]

Hyuncheol Bryant Kim[†], Booyuel Kim[‡], Cristian Pop-Eleches[§], Jaehyun Jung[¶]

August 2021

Abstract

We study the impact of male circumcision on risky sexual behaviors and the role that peers play in the decision and consequences of being circumcised. It is based on a two-stage experiment that randomly offered free circumcision with transport support within 33 secondary schools in Malawi. We find an increase in the demand for circumcision for those assigned to treatment and evidence of peer effects in the decision to get circumcised among untreated students. A four-year follow-up using biomarkers of sexually transmitted infections shows evidence of risk compensation for those who circumcised due to the intervention, but not for those induced by peer effects.

(JEL: C93, O33, I12)

[*] We wish to thank Douglas Almond, John Cawley, Esther Duflo, Josh Graff Zivin, James Habyarimana, Ofer Malamud, Doug Miller, Roger Moon, Matthew Neidell, Harsha Thirumurthy, Miguel Urquiola, as well as seminar participants at Columbia University, Cornell University, Georgetown University, Hitotsubashi University, Hong Kong University of Science and Technology, Seoul National University, University College London, University of Seoul, University of Southern California, Yonsei University, NBER summer institute, and APPAM and AEA annual meetings. We also thank all members of 'Project Malawi' of Africa Future Foundation (AFF) for their excellent fieldwork. This research was supported by the Korean International Cooperation Agency (KOICA), the research grant of KDI School of Public Policy and Management, and the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018S1A3A2075117). All views expressed are ours, and all errors are our own. The study was approved by ethical review committees at the National Health Sciences Research Council (Malawi NHSRC Protocol #902), Columbia University (USA, IRB-AAAL8400), and Cornell University (USA, 1310004153).

[†] hbkim@ust.hk, Department of Economics, Hong Kong University of Science and Technology, Hong Kong and Department of Policy Analysis and Management, Cornell University, Ithaca, NY 14835.

[‡] booyuel@snu.ac.kr, Department of Environmental Planning, Graduate School of Environmental Studies / Environmental Planning Institute (EPI), Seoul National University, Seoul, Republic of Korea 08826.

[§] cp2124@columbia.edu, School of International and Public Affairs, Columbia University, New York, NY 10027.

[¶] jjung6@kipf.re.kr, Korea Institute of Public Finance, Sejong, Republic of Korea 30147.

1. Introduction

HIV/AIDS is one of the world's most serious health challenges and huge efforts have been put into practice worldwide to address this problem. Although 25.4 million out of 38 million people living with HIV received antiretroviral therapy by the end of 2019, HIV/AIDS prevention remains an important challenge since the number of new infections in 2019 was 1.7 million, only a 23% decline since 2010 (UNAIDS (2020)).

Male circumcision has become an integral strategy for HIV prevention after three efficacy trials showed that male circumcision can reduce HIV transmission risk by about 50 percent after the procedure (Auvert et al. (2005), Bailey et al. (2007), Gray et al. (2007)). In addition, male circumcision also reduces herpes simplex virus type 2 (HSV-2) and human papillomavirus (HPV) infection (Sobngwi-Tambekou et al. (2009), Tobian et al. (2009)). To promote the demand for male circumcision, the World Health Organization (WHO) strongly recommends medical male circumcision as a key intervention for reducing female to male transmission of HIV (WHO/UNAIDS (2007)). There has been a global mobilization for scaling up male circumcision especially in countries with high HIV prevalence and low prevalence of male circumcision (WHO (2019)).

However, there are two major concerns related to scaling up male circumcision: weak demand for the procedure and potential risk compensation. First, the demand for male circumcision is still very low even with a heavily subsidized price and proper information (Chinkhumba et al. (2014)). Among the major barriers discussed in the literature are financial constraints, lack of information, awareness, and accessibility, fear of pain, as well as religious and cultural norms (Evens et al. (2014)). Second, even if a scale-up project is successful in increasing the take-up of male circumcision, such programs might have limited impacts if circumcised men are more likely to engage in risky sexual behaviors (Bingenheimer and Geronimus (2009), Brooks et al. (2010), Cassell et al. (2006), Eaton and Kalichman (2009), Kalichman et al. (2007), Kremer (1996), Mattson et al. (2008), Padian et al. (2008), Sawires et al. (2007), Weiss et al. (2010), WHO/UNAIDS (2007)).

This paper attempts to identify the magnitude of peer effects in the decision to get circumcised as well as to evaluate the long-term impact of male circumcision on risky sexual behaviors. The motivation for this paper is twofold. The first is to understand peer effects, which have been shown to affect behaviors in a wide range of areas including education (Foster (2006), Kremer and Levy (2008), Lyle

(2007), Sacerdote (2001), Zimmerman (2003)), health (Chong et al. (2020), Godlonton and Thornton (2012), Miguel and Kremer (2004), Oster and Thornton (2012)) and labor (Bandiera et al. (2010), Mas and Moretti (2009)). More specifically, there is an interest among researchers and policy makers in the role that peer effects may play in the take-up of certain interventions or the adoption of certain technologies. This is because take-up is often considered suboptimal and potential spillovers to peers might justify subsidizing an intervention (Miguel and Kremer (2004)). The role of peer effects in the demand for male circumcision is theoretically ambiguous. It may be positive if friends provide emotional support that reduces psychological costs or if they share private information about the benefits of the circumcision procedure. On the other hand, a negative experience following male circumcision (i.e. complication or pain) might decrease a friend's demand for male circumcision.

The second motivation is to understand whether in addition to the direct biological effects of male circumcision on the probability of transmission of HIV, there are also behavioral responses that could reduce or amplify the impact of getting circumcised. The evidence on risk compensation from the projects that implemented the early efficacy trials in Kenya, South Africa and Uganda is based on short run follow-up studies and does not show a strong relationship between male circumcision and changes in risky sexual behaviors (Auvert et al. (2005), Bailey et al. (2007), Gray et al. (2007), Mattson et al. (2008), Wilson et al. (2014)). In addition, long-term experimental studies of risk compensation in these settings were not possible because circumcision was provided to those in the control group at the end of the efficacy trials.[1]

There are several reasons why the behavioral responses in the scale-up project that we study might be different from those in the original efficacy trials. First, the efficacy trials reviewed above, and as is more generally the norm with these types of trials, were combined with intensive health education and individual HIV counseling interventions throughout all the follow-up visits up to 24 months. This suggests that risk compensation in the original efficacy trials may be more limited compared to the current male circumcision scale-up projects where knowledge of the benefits of male circumcision is widely accepted and follow-up HIV counseling is often more limited. Second, current beliefs and

[1] For example, Gray et al. (2012) and Kong et al. (2012) try to explore risk compensation in the long-term by comparing risky sexual behaviors between circumcised and uncircumcised men in the post-trial follow-up surveys in Uganda. These studies do not find compelling evidence of increased risky sexual behaviors among circumcised men. However, these results may be driven by confounders that may affect both circumcision take-up and risky sexual behaviors, since they are not based on the experimental variation of the trials. Gao et al. (2021) perform a meta-analysis on the association between male circumcision and HIV risk compensation and do not find evidence of risk compensation. However, we note that none of the reviewed articles used bio-makers for measuring risk compensation behaviors.

attitudes to circumcision at the time of our study could be substantially different compared to the time of the efficacy trials when the biological impacts of male circumcision on HIV and HSV-2 infections were neither established nor well understood. Third, the selection of subjects into clinical trials could be different than in a scale up project that targets a broader population. Therefore, if a scale-up project encourages specific groups of people to take up circumcision, the effects of the circumcision on risky sexual behaviors could differ from those in a clinical trial. Finally, risky sexual behaviors could be different in the short and longer term following circumcision, especially for an adolescent population.

We take advantage of a large-scale HIV prevention program implemented in a sample of 3,970 9th to 11th graders attending 124 classrooms in 33 public secondary schools in Malawi by the Africa Future Foundation (AFF), an international non-governmental organization (NGO). Given the capacity constraints of the partner hospital, AFF used a phase-in design to allow the entire study population to get free male circumcision with transport support at the assigned hospital. Specifically, during Round 1 of the study (December 2011 to May 2013), classrooms were randomly assigned into three groups: *100% Treatment*, *50% Treatment*, or *No Treatment* classrooms. All male students in *100% Treatment* classrooms received a free circumcision offer with transport support in Round 1, while no students in *No Treatment* classroom received the intervention in Round 1. In *50% Treatment* classrooms, half of the students were randomly selected to receive the intervention in Round 1. In Round 1 of the intervention, the free circumcision offer with transport support was given to a total of 1,972 male students.

In Round 2 (June 2013 to December 2013), which was implemented after the short-term follow-up survey was completed, the remaining 1,998 male students who were temporarily not assigned to treatment in Round 1 also received the free circumcision offer with transport support intervention. However due to funding constraints of the collaborating NGO that were not expected at the beginning of the study, the frequency and duration of the intervention were significantly reduced in Round 2 in comparison to Round 1. In order to study long-term outcomes for the different treatment groups, the two youngest cohorts (9^{th} and 10^{th} graders at baseline) were selected for the long-term follow-up survey that was implemented after four years from the baseline survey. Two main strengths of the long-term follow-up survey are a re-interview rate of 86.8% and the collection of biomarkers for sexually transmitted infections (STIs) including HIV and HSV-2.

We have four main sets of results. We first find that at the end of Round 1, male students who received the free circumcision offer with transport support are approximately three times more likely to take-up male circumcision. These results, which are based on administrative data from hospital records, are also confirmed from self-reported data on circumcision in the short-term follow-up survey. Moreover, as a result of the shortened intervention in Round 2, take-up rates of circumcision for those offered treatment in Round 2 were lower after the completion of Round 2 and these take-up differences between Round 1 and Round 2 persist even four years after the start of the study.

The second finding is that peers play an important role in the decision to get circumcised. Untreated students in *50% Treatment* classrooms were 79% more likely to get circumcised than students in *No treatment* classrooms, suggesting a positive externality in the demand for male circumcision. This type of peer effect, which refers to the effect on untargeted individuals who are exposed to treated individuals, can be interpreted as a *diffusion* effect (Fafchamps, Vaz, and Vicente (2020) and Fafchamps and Vicente (2013)). Since we collected a roster of three best friends within the classroom at baseline, we also use an alternative empirical strategy based only on students in the 50% Treatment classrooms that uses the experimental variation in the fraction of friends who were offered treatments to further understand peer effects. The results using this method to capture peer effects are also positive but less precise. However, when we use the same variation in the fraction of close friends who received the intervention in 50% treatment class, we find evidence that an individual who is assigned to the intervention is more likely to get circumcised, when a larger share of his close friends also received the intervention. This type of peer effect has been defined in the literature as a *reinforcement* effect (Fafchamps, Vaz, and Vicente (2020) and Fafchamps and Vicente (2013)).

Our third set of findings use the data from the four-year follow-up to estimate credible long-term causal effects of male circumcision on bio-markers as well as risky sexual behaviors. In a nutshell, we find evidence that is consistent with risk compensation among those who received the free circumcision offer with transport support intervention in Round 1. Specifically, we find higher rates of HSV-2 infection measured by IgG (reflecting lifetime infection) and IgM (reflecting recent infections or reinfections) antibodies for students who were offered the more intense intervention in Round 1 (with a higher take up rate of circumcision) compared to those students who received the less intense intervention in Round 2 (with a lower take up rate of circumcision) in the long-term follow-up survey. These results are corroborated to some extent by other risky sexual behaviors (such as using condoms for casual sex), measured by the item count technique (as in Coffman et al. (2017)). In our analysis we also provide results from a simulation exercise showing that relatively large treatment effect of the intervention on HSV-2 infections can be explained by a combination of a selection effect into circumcision as well as risk compensation behavior after circumcision.

The fourth main finding is that risky sexual behaviors among those induced to get circumcision because of the direct intervention (those treated in both 100% *Treatment* and 50% *Treatment* classrooms) are different from those who took up circumcision due to the indirect peer effect (those untreated in *50% Treatment* classrooms). In our setting, when a student got circumcised as a result of peer effects, we do not observe any evidence of risk compensating behaviors.[2] We further explore this striking difference by analyzing the heterogeneity of male circumcision take-up using two approaches. First, we try to understand if there are any observable characteristics of compliers between the two groups that get circumcised by the direct intervention and by peer effect channels. Secondly, we apply machine-learning techniques proposed by Chernozhukov et al. (2018) to explore treatment take-up heterogeneity and relevant characteristics driving this heterogeneous take-up in the two groups. Our results using both approaches support the view that take-up of male circumcision is heterogeneous and driven by observable characteristics such as risk preferences and personality traits.

These findings contribute to two main strands of the literature. First, to the best of our knowledge, it is the first to explore long-term experimental effects of male circumcision on risk compensation. We are also the first to study these questions in a scale-up project setting rather than in the context of a clinical trial. The existing literature on this subject is based on either data from projects that implemented the short-term efficacy trials (Auvert et al. (2005), Bailey et al. (2007), Gray et al. (2007), Mattson et al. (2008), Wilson et al. (2014)) or longer-term observational studies that compare those who get circumcision and those who do not (Gray et al. (2012), Kong et al. (2012)). This study is more broadly related to the risk compensation literature also known in the economics literature as the "Peltzman effect" (see Peltzman (1975) and Evans and Graham (1991) for applications to seat belts, and Thompson et al. (1999) for an application to bicycle helmets).[3]

Second, this paper contributes to the literature on peer effects. Arguably most important are the

[2] This result is similar to Duflo and Saez (2003) in the case of retirement plan decisions. Their findings suggest that peer-effect compliers may have different treatment effects than incentive compliers.

[3] Our study also contributes to the literature on how to improve male circumcision take-up. For example, previous trials show that a financial compensation of $8.75 increases take-up four times in Kenya (Thirumurthy et al. (2014)). Similar magnitudes have been found in another study from an $11 compensation in Kenya (Evens et al. (2016) as well as from a $10 conditional compensation in South Africa (Wilson et al (2016). Chinkhumba et al. (2014) using a randomized price and information interventions in Malawi find increases in circumcision rates of 3%.

implications of our study of the different long-term behaviors between male students who took up circumcision directly through the program and those who got circumcised through peer effects. Regarding the role of peer effects in the take-up of technologies or interventions, many studies aim to understand short-term take-up decisions but they fail to validate their short-term outcomes in the longer-term period. We show that take-up of male circumcision through two different mechanisms (direct inducement of transport intervention and indirect inducement through peer effect) can lead to very different behavioral responses in the long-term.

The paper is organized as follows. Section 2 provides background information on the Malawian context and the male circumcision intervention. In Section 3 we describe the data followed in section 4 by the empirical strategies. Section 5 presents the main results. The final section presents conclusions.

2. Background and Experimental Design

2.1. HIV and Male Circumcision in Malawi

Malawi has been heavily affected by the HIV pandemic with an HIV prevalence rate of 8.8% for people aged 15 to 49 (NSO (2017)) and a life expectancy at birth of 61 years (WHO (2018)). The prevalence of male circumcision in Malawi is relatively low with 28% of men being circumcised (NSO (2017)).[4] Male circumcision is practiced mainly for religious and cultural reasons in Malawi. For example, 98% of Muslims are circumcised while only 19.8% of Christians practice circumcision (NSO (2017)). Culturally, 90.3% of those belonging to the Yao ethnic group practice male circumcision while other ethnic groups have much lower levels of male circumcision (NSO (2017)). The baseline circumcision rate in the sample is 10.5%, which reflects the fact that majority of residents in the study area are from the Christian Chewa ethnic group.[5]

Medical male circumcision for HIV prevention has become an important component of Malawi's national HIV prevention program since 2011 (NAC (2012)). Nevertheless, despite recent

[4] The true prevalence of complete male circumcision could be lower because many of those reporting being circumcised practice incomplete circumcision, which only removes part of the foreskin. Incomplete circumcision may not have the full protective benefits of male circumcision (Bengo et al., 2010).

[5] Most people belonging to the Yao ethnic group live in the southern part of Malawi, while the Chewa are the main ethnic group in the central region where our project is located.

growth, the number of medical male circumcisions was still small at the time when this study was implemented in 2012.[6] The estimated number of male circumcisions performed between 2008 and 2011 was only 0.7% of the target number (2.1 million) needed to achieve a male circumcision prevalence of 80% in Malawi.

2.2. Study Setting and Experimental Design

This study was conducted in four catchment areas of Lilongwe district by the partner NGO (AFF), as part of a larger development program that also included an HIV/AIDS education component and a cash transfer program for female secondary students.[7, 8] The target population of the study was male secondary school students identified through a school-based baseline survey. Our sample includes 3,970 male students enrolled in grades 9 to 11 in 124 classrooms of 33 secondary schools within the catchment area.[9] Ethnically, the five largest ethnic groups in the sample were Chewa (52.7%), Ngoni (17.8%), Lomwe (9.3%), Tumbuka (9.0%), and Yao (7.4%).[10]

The medical male circumcision clinic within Daeyang Luke Hospital located in the catchment area of the study was established by AFF. The clinic was equipped with a self-contained surgical unit and surgical beds to perform modern and safe medical circumcisions by appropriately trained medical personnel. Medical personnel using the standard forceps-guided method performed the medical male circumcision surgeries under local anesthesia in the assigned clinic.

Table 1 summarizes the two-stage randomization design of the study. 124 classrooms within the 33 schools were stratified by grade and randomly assigned into three groups: *100% Treatment*, *50% Treatment*, and *No Treatment* classrooms. All male students in the *100% Treatment* classrooms received the free male circumcision offer with transport support during Round 1 (Group 1). No students in the *No*

---

[6] Starting from very small numbers (589 in 2008, 1,234 in 2009, and 1,296 in 2010), there have been sizeable increases of medical male circumcision in Malawi since 2011 (11,881 in 2011, 21,250 in 2012, 40,835 in 2013, 80,419 in 2014, 108,672 in 2015, 129,975 in 2016, and 166,350 in 2017) (WHO (2018)).

[7] The partner NGO's catchment areas include TA Chimutu, TA Chitukula, TA Tsbango, and TA Kalumba.

[8] The impact of the cash transfer program on girls was analyzed in Kim et al. (2018). Both the cash transfer for girls and the HIV/AIDS education program were cross-randomized across classrooms. We discuss their possible impacts on male circumcision behavior in the analysis.

[9] Most schools (28 out of 33 schools) have one class per grade and there are limited cross-grade school activities. While our main analysis uses all 33 schools, our main long term results discussed in Section 5.3 are robust when we are restricting the analysis to the 28 schools in our sample that do not have multiple classrooms per grade.

[10] According to the 2018 Malawi Census, the five largest ethnic groups were Chewa (34.4%), Lomwe (18.9%), Yao (13.3%), Ngoni (10.4%), and Tumbuka (9.2%). Although our sample over-represents Chewa people and under-represents Lomwe and Yao people, the ethnic diversity within the sample is close to that of the population in Malawi.

Treatment classrooms received the intervention during Round 1 (Group 4). In 50% Treatment classroom, we randomly selected half of students for the treatment (Group 2), and the remaining students were not treated during Round 1 (Group 3). Assignment to the treatment in *50% Treatment* classrooms was done at the individual level.[11] This two stage experimental design that randomizes treatment both across and within classes allows us to measure not only the direct effect but also peer effect of the male circumcision offer.[12] As summarized in Table 1, 41 classrooms (across 24 schools) were assigned to *100% Treatment*, 41 classrooms (across 25 schools) were assigned to *50% Treatment*, and 42 classrooms (across 28 schools) were assigned to the *No Treatment* group in Round 1. Two additional features of the program that are worth noting are that prior to receiving their circumcision offers, students in treated classrooms in both Round 1 and Round 2 received a short information session about male circumcision and while not made explicit by AFF, the free circumcision and complication check-up services were available to students from all groups throughout the study period who visited the clinic.

During Round 1, AFF provided a male circumcision offer that consisted of free access to the surgery, complication check-ups, and transportation support to the clinic to the selected students (Group 1 and Group 2). Students receiving the transportation offer could choose either a direct pick-up service provided by AFF from the school to the clinic or a transportation voucher that was reimbursed after the circumcision surgery at the hospital was performed. The amount of the voucher varied according to the distance between the clinic and a student's school. These vouchers were student specific and student participants were not allowed to trade these vouchers among themselves. In Round 2, due to budgetary constraints of the implementing NGO, the remaining temporarily untreated male students in Round 1 received a less intense intervention. The main difference between Rounds 1 and 2 is the duration and frequency of the support program.

The transport vouchers for independent travel to the clinic were given out twice during Round 1. The first voucher did not specify an expiration date but specified six possible dates available to use for the pick-up service to the clinic (Panel A of Figure A1). The second voucher specified an expiration date (the end date of Round 1) and one possible pick-up date (Panel B of Figure A1). During Round 1, the NGO offered 138 surgery dates over a period of 18 months and 78 pickup dates by bus. In the

---

[11] The partner NGO wanted to keep the demand constant due to the capacity constraints. During the study period, the maximum number of surgeries per day was 19.

[12] This standard two-step randomized design to estimate peer effects was first proposed in Duflo and Saez (2003).

beginning of Round 2, Groups 3 and 4 were also given the intervention. The intervention lasted only for 6 months, the travel vouchers were offered only once per student (Panel C of Figure A1) and the NGO offered only 44 surgery dates and 38 pickup dates by bus.

After the medical male circumcision surgeries, post-circumcision care was provided after three days and seven days at the student's school to check and disinfect the wound, and record complications. The program provided intensive pre- and post-surgical counseling sessions based on the guideline of WHO/UNAIDS (2008) on male circumcision under local anesthesia for adolescent boys and men (Figure A2). It lasted 30 to 45 min, and was provided in small groups that averaged around 3 students. It is worth noting that circumcision participants were advised that male circumcision provides only partial protection against HIV infection and thus practicing safe sex after the surgery was encouraged and recommended.

The timeline of the project was as follows. After the baseline survey that was collected between October 2011 and May 2012, male students assigned to treatment in Groups 1 and 2 received the vouchers during December 2011 to May 2013 (Round 1). The short-term follow-up survey was conducted between January and June of 2013. Since our research design was based on a phase-in design, at the time of the short-term follow-up survey, the untreated students during Round 1 (Groups 3 and 4) had not received the intervention yet. Groups 3 and 4 received the less intensive intervention (than Groups 1 and 2) from June to December 2013 (Round 2).[13] A long-term follow-up survey was implemented approximately 4 years after the baseline survey between October 2015 and August 2016. Due to the budget constraints of the partner NGO, we focus only on the 2,663 students who were 9th and 10th graders at baseline for the long-term follow-up survey.

3. Data

We use four main sources of data for this research: baseline, short-term and long-term follow-up survey data as well as administrative data from the male circumcision clinic. At baseline, we started with a list of 10,715 enrolled students at the 33 participating schools. We managed to successfully get consent and complete the baseline survey for 74.4% of the students in the school roll-call lists.[14] Of the 7,967

---

[13] One may worry that the timing of the circumcision offers are also different between the treatment and control group, and it could be problematic if risk compensation pattern is non-linear over time. In Section 5.3.1, we address this concern using bio-markers that capture lifetime and recent infections separately.

[14] For the 15% of the students in our sample under the age of 15 at the time of the baseline survey, we received parental

secondary students who completed the baseline, 3,997 were girls and 3,970 were male students.

The baseline survey was designed to measure detailed background characteristics including student information about HIV knowledge, sexual behaviors, and their friendship network. At the end of the survey we gave students 10 kwacha (6 cents), and offered them a chance to buy condoms at the subsidized price of 5 kwacha to measure the demand for safe sex as in Thornton (2008).[15]

In Table 2 we present summary statistics and balance tests of baseline observable characteristics. Columns (1) to (4) and (5) to (8) of Table 2 refer to the full sample (9th to 11th graders) and the long-term follow-up sample (9th and 10th graders), respectively. In Columns (1) and (5), we show the mean characteristics for those in the control group. As shown in Column (1), the average age of study participants in the control group is 16.8 years and 16% belong to ethnic groups that practice circumcision and 10% are already circumcised. In general, study participants showed a high level of HIV/AIDS knowledge: the average number of correct answers to the HIV/AIDS knowledge questionnaire was 17.4 out of 20 questions. They also have a relatively high knowledge of the medical benefit of male circumcision (67.1%). In addition, 36% of the control group students believe that male circumcision is painful and 15% believe that male circumcision is only appropriate for Muslims. 31% ever had sex and 9% are currently engaged in a sexual relationship. Columns (2) to (4) and (6) to (8) present mean differences between the treatment groups (Groups 1 to 3) and the control group (Group 4). The results confirm that classroom randomization was well implemented: the proportion of statistically significant mean difference at the 10% significance level is 9 out of 60 (15%) in the full sample (Columns (2) to (4)) and 9 out of 60 (15%) in long-term follow-up sample (Columns (6) to (8)).

We also asked students to list their three best friends within the classroom regardless of sex. In the analysis, we reconstructed friendship data by reordering best male friends after excluding friends without baseline survey information and female friends. Table A1 shows summary statistics for the friendship networks. In Panel A, Column (1) includes the original friendship network data, and Column (3) presents reranked male friends after excluding female friends and those who did not participate in the baseline survey. Almost all (3,832 out of 3,844) have at least one male friend, and around 80% (3,135 out of 3,844) have two male friends among their three best friends. Panel B presents the

---

consent.

[15] At the time when the study was implemented, condoms were generally available for free at health centers. However, long travel times to health centers were a major barrier to widespread use.

treatment distribution among eligible male friends. It shows substantial variation in the fraction of best male friends who got treated. In addition, Table A2 confirms that the fraction of treated friends is well balanced across the baseline characteristics.

The main outcome variables are an indicator for male circumcision take-up, self-reported sexual behaviors, and bio-marker test results (HIV and HSV-2), which were specified, in the pre-analysis plan.[16] One advantage of our setting is that for the measurement of circumcision take-up, we collected a hospital administrative dataset containing the complete records of all the medical circumcisions performed at the assigned clinic. As a result, we do not need to rely on self-reported data, which could easily be affected by social desirability bias. The administrative data records the name and age of the patient as well as the medical information related to the procedure, which includes the date of the surgery, the dates of any follow-up visits as well as information on complications related to the procedure.[17] In addition the administrative record also included information on whether circumcision participants used the direct pick-up service or redeemed the transport voucher. During the study period, 502 boys were circumcised in Round 1 and 124 in Round 2, which likely reflects the lower intensity of treatment in Round 2.[18]

We also use the short- and long-term follow-up surveys implemented one and four years after the baseline survey to measure sexual behaviors, and STIs (HIV and HSV-2). The short-term follow-up survey was first conducted at school. 68.0% of the baseline sample students (2,698 students) completed the survey at school. For the remaining 32.0% who were lost because of absence, transfer, or dropout, we randomly selected 15%, and For more details on the estimation and selection of the baseline controls see the pre-analysis plan for the long-term follow-up at AEA RCT Registry AEARCTR – No.0001335 (Kim et al., 2016) implemented an intensive community survey.[19] The long-term follow-up survey was a

[16] Biological data were not collected at baseline because the NGO was concerned that students who tested HIV positive at school may not be able to continue their studies. Instead, HIV testing was done when students came to the clinic for the circumcision procedure.

[17] It is unlikely for students to get circumcised in other medical facilities at the time of our study because there were few facilities nearby that provided male circumcision on a regular basis. One exception is the Banja La Mtsogolo (BLM) clinic located in the TA Chitukula area, which is one of our four catchment districts. However, this clinic charged around USD 10 for the surgery and complication check-ups, which limited the demand for the circumcision procedure, especially for the population of secondary school students.

[18] In Round 2, 1,998 male students from Group 3 and Group 4 were offered the free circumcision and transportation treatment. The number of students in Group 1 and 2 is roughly similar to the 1,972 students in Group 1 and 2. Among the students in the long-term follow-up sample (grade 9 and 10 at baseline) 303 were circumcised in Round 1 and 91 in Round 2.

[19] While the long-term follow-up survey was implemented for the entire baseline sample of 9th and 10th graders, the data collection for the short-run follow-up was conducted in two stages. In the first stage, we revisited the schools in our sample

community survey. It included several novel measures of sexual behaviors and STIs, which are the main outcomes variables used to analyze the impacts of the circumcision intervention. In terms of bio-markers, HSV-2 and HIV infections were evaluated using rapid test kits.[20]

The main outcome for measuring risky behavior is the HSV-2 test for Sexually Transmitted Diseases (STDs), a test that has been successfully used in a number of related studies (Baird et al. (2012), Duflo et al. (2015)). Since HSV-2 is almost exclusively sexually transmitted, our rapid test results provide a good measure of risky sexual behaviors. For HSV-2, we use two measures of infection: the IgG is a permanent marker of HSV-2 infection that has occurred at any point during a person's lifetime (Obasi et al. (1999)), while IgM shows recent infections or reinfections (LeGoff, Péré, and Bélec (2014), Workowski and Bolan (2015)). Therefore, the IgG marker is our primary outcome because it captures any HSV-2 infection up until the follow-up survey. The secondary outcome of interest is the HIV infection rate, and the HSV2 infection measured by IgM is the tertiary outcome, which were predetermined in the pre-analysis plan.

One potential complication with using the results from HSV-2 and HIV tests to capture changes in risky sexual behavior after circumcision is that male circumcision prevents HSV-2 and HIV infections by about 50 percent (Auvert et al. (2005), Bailey et al. (2007), Gray et al. (2007)). For example, the discussion in the medical literature (Tobian et al. (2009)) suggests that both anatomical and cellular factors could explain biological mechanisms that would reduce the HSV-2 infections for circumcised individuals who engage in unprotected sexual activities. Therefore, as a result of this biological mechanism, any effect of the circumcision intervention on HSV-2 prevalence is likely a lower bound of the true effects induced by risk compensation.

In addition, we implemented an alternative measure of risky sexual behavior using the item

---

and attempted to re-interview all the students in our initial sample. We successfully interviewed 68.0% of the baseline sample students (or 2,698 students) using interviews conducted at school. During the second stage, we implemented a more extensive tracking exercise as in Thomas et al. (2001) and Thomas et al. (2012). First, we selected a subsample of 15% of the students (191 students) among those who did not participate in the school follow-up survey and attempted to visit them at home to complete the survey. The home survey follow-up rate was 74.9%. Combining the school and home visits results in an effective survey follow-up rate of 91.9%. The effective survey rate (ESR) is a function of the regular school follow-up rate (RFR) and intensive home-visit follow-up rate (HFR) as follows: ESR = RFR + (1 - RFR) × HFR. Overall, ESR is 91.9% (68.0% + 32.0% ×74.9%). We run a weighted regression with a weight 6.67 for home-visit survey since we randomly selected 15% students from the attrition sample (Baird et al., 2012)

[20] Although we have implemented HIV testing during the long-term follow-up survey, given the low incidence of HIV, our power calculations indicated that our study was not powered to detect differences in HIV rates between the treatment and the control groups.

count technique (ICT) (Coffman et al. (2017), Miller (1984)). This is to further account for potential measurement error in self-reported responses, since numerous studies point out that self-reported answers are poor proxies of true attitudes about sensitive topics such as sexual behaviors (Minnis et al. (2009), Palen et al. (2008)). The ICT methodology asks respondents to report the total number of true statements in a set of questions that may include a sensitive item, instead of directly endorsing it. Participants were randomly given one of two sets of questions. One set included only four non-sensitive items (*short-form*) and the other set included four non-sensitive items and an extra sensitive statement of interest (*long-form*). We used two sensitive questions of interest: "I think I have to use a condom in case of sex with somebody that I do not know well" (Using condoms for casual sex) and "I had sex with more than two people in last 12 months" (Multiple sexual partners). We can measure the impact of assignment to treatment on the sensitive statement by estimating the differential proportion of respondents who agree with the sensitive statement between the treatment and control group. Details of the ICT questions are described in Figure A3. Lastly, at the end of the follow-up surveys we again gave students 10 kwacha (6 cents), and offered to sell them condoms at a subsidized price of 5 kwacha in order to measure the demand for condoms.

In the baseline and two follow-up surveys we also included questions that are typically asked in surveys that are based on self-reported sexual behaviors. These include questions about attitudes toward condom use, inconsistent use of condoms, unprotected sex with a recent partner and sex with multiple sexual partners. Using these secondary outcomes, we created an 8-item risk score to summarize self-reported sexual behaviors into a single index.

We conclude this section with an analysis of attrition behavior during the follow-up surveys as well as take-up of the biomarker tests. We were able to track about 91.9% and 86.8% of study subjects during the short- and long-term follow-up surveys, respectively. A smaller percentage (78.2%) participated in the bio-marker tests. Table A3 suggests that attrition from the sample is not correlated with treatment assignment and there are no statistically significant differences between the treatment groups (F-tests results) in the long-term follow-up survey and the take-up of the biomarker tests.[21][22]

[21] In Columns (1)-(2) of Table A3, there are statistically different attrition rates in the short-term follow-up survey, but we only use the sexual behaviors that were self-reported in the short-term follow-up survey in Panel A of Table 8.
[22] We have also explored and found no evidence of differential attrition by baseline characteristics across the different intervention arms.

4. Estimation Strategy

We employ a number of empirical strategies to capture the direct effect of being assigned to the more intense male circumcision intervention as well as possible peer effects in this setting. The first empirical strategy estimates the following model:

(1) $Y_{ijgs} = \beta_0 + \beta_1 \cdot G1_{jgs} + \beta_2 \cdot G2_{jgs} + \beta_3 \cdot G3_{jgs} + \gamma' X_{ijgs} + \delta_g + \varepsilon_{ijgs}$

where Y_{ijgs} denotes an outcome of interest such as male circumcision take-up, sexual behaviors, and bio-marker test results for individual student *i* attending classroom *j* in grade *g* at school *s*. G1, G2, and G3 refer Group 1, Group 2, and Group 3, respectively. Group 4 is the reference group. The control vector, *X*, includes age, circumcising ethnicity, circumcising religion (Muslim), orphan status, parent's education, parent's job, household assets and school type, and the assignments to HIV/AIDS education and girl's cash transfer program. δ captures grade fixed effects, and ε is a random error. In some specifications we also include school fixed effects. Errors are clustered at the classroom level, which is the randomization unit. We also present heterogeneous treatment effects by three different priors such as knowledge on the medical benefit of male circumcision, fear of pain, and religious norms.

In these specifications, both β_1 and β_2 capture the direct effects of being assigned in Group 1 and Group 2. The difference between β_1 and β_2 also captures possible peer (diffusion) effects, given that Group 2 has peer who were not assigned to treatment in the classroom (Group 3) while everybody in Group 1 was assigned to the treatment. In addition, β_3 might capture possible peer (diffusion) effects, given that a key difference between Group 3 and Group 4 is existence of peer(s) who received the free male circumcision offer and transport support (Group 2) within the classroom. It is worth noting that although both β_1 - β_2 and β_3 capture potential diffusion effects, they are conceptually not the same since β_1 - β_2 reflects the peer effects when comparing classrooms where all and 50% of peers have received the treatment while β_3 reflects the peer effects comparing classrooms where 50% and 0% of peers have received the treatment.

In some of our specifications, we also restrict ourselves only to the 50% Treatment classrooms. An alternative way to estimate the main direct effect of the intervention is by comparing the difference in outcomes between students in Group 2 who received the intervention and students in Group 3 who did not receive the intervention during Round 1. Both of these groups (Group 2 and Group 3) are within the same classrooms and contain students who are exposed to the same peer effects since 50% of their peers are treated.

Next, we extend the analysis of peer effects by using our experimental variation in treatment intensity with the data from the friendship rosters. We try to measure these effects in the restricted sample of *50% Treatment* classrooms because there is no within class variation in *100% Treatment* and *No Treatment* classrooms. The following linear regressions are estimated:

(2) $Y_{ijgs} = \alpha_0 + \alpha_1 \cdot G2_{jgs} + \alpha_2 \cdot Peer_{ijgs} + \gamma' X_{ijgs} + \psi_g + \Phi_s + \varepsilon_{ijgs}$,

(3) $Y_{ijgs} = \kappa_0 + \kappa_1 \cdot G2_{jgs} + \kappa_2 \cdot Peer_{ijgs} + \kappa_3 \cdot (G2_{jgs} \cdot Peer_{ijgs}) + \gamma' X_{ijgs} + \psi_g + \Phi_s + \varepsilon_{ijgs}$,

where $Peer_{ijgs}$ is the fraction of male friends on the friendship roster that was offered treatment in Round 1. ψ and Φ is a grade and school fixed effect, respectively, and errors are clustered at the classroom level. Since receiving the offer within our 50% Treatment classrooms is randomly assigned, the fraction of friends that received the treatment in the classroom is also random. It is worth noting that the peer effect that α_2 in Equation (2) captures is different from the peer effect β_3 in Equation (1) for several reasons. First, α_2 captures possible peer effects coming from the smaller group of best friends while β_3 reflects peer effects from the larger group of classroom students. Second, even though we did not explicitly provide information on free male circumcision opportunity to both Group 3 and Group 4 in Round 1, Group 3 can find this opportunity more easily than Group 4 due to their classroom peers from Group 2.

In Equation (3), we estimate another type of peer effect, resulting from potential complementarities between the subject's offer and his friend's offers. This type of peer effect could also be defined as a reinforcement effect (Fafchamps, Vaz and Vicente (2020) and Fafchamps and Vicente (2013)), and is captured by the coefficient of the interaction term κ_3. In our setting, it is possible for such reinforcement effects to be present when peers make a decision together to get circumcised.

Since we are considering many outcomes for risky sexual behaviors simultaneously, there may be a probability of observing a few large estimates due to sampling variability, even if there was no true effect. To account for these multiple comparisons, we follow Finkelstein et al. (2012)'s approach. The first approach is to group outcome measures into a domain and to take an average standardized treatment in each domain, as suggested in the pre-analysis plan.[23] We group three bio-makers (HSV-2 IgG, HSV-2 IgM, and HIV) into the STD infection domain. Then, we stack the data for individual outcomes within each domain and estimate a single regression equation while clustering standard errors both at the classroom and individual level in order to compute the average standardized treatment effect. Next, we use the free step-down resampling method for multiple hypotheses testing to adjust the family-wise error rate (Westfall and Young, 1993)[24]. The three hypotheses (HSV-2 IgG, HSV-2 IgM, and HIV) arise from examining multiple outcomes in the STD infection domain.[25]

5. Results

5.1 Impacts of intervention on the take-up of male circumcision

Figure 1 shows the cumulative prevalence of male circumcision over time based on hospital records data (Panel A) and self-reported data in the long-term follow-up survey (Panel B) for baseline 9 and 10th graders.[26] We rely more on hospital records data since self-reported data may suffer from recall bias.[27] Hospital administration data is not available after December 2013 because the clinic was closed after the program was ended. At the beginning of the project, male circumcision prevalence among our study participants was 10.5%.[28] Due to the intervention, circumcision take-up among treated students in *100% Treatment* (Group 1) and *50% Treatment* classrooms (Group 2) significantly increased during Round 1. Untreated students in *50% Treatment* classrooms (Group 3) were also more likely to take-up male circumcision during Round 1 than the control group (Group 4).

In Panel A of Table 3 we present estimates of the impact of the intervention on the demand for male circumcision that are based on Equation (1), while in Panel B we combine the two treatment arms

[23] For more details on the estimation and selection of the baseline controls see the pre-analysis plan for the long-term follow-up at AEA RCT Registry AEARCTR – No.0001335. The pre-analysis plan (Jung et al. 2016) can be found at the following link: https://doi.org/10.1257/rct.1335-2.0
[24] Specifically, we used the Stata command 'wyoung' developed for Jones et al. (2019).
[25] We conducted 10,000 simulations for each multiple hypothesis test with and without baseline covariates, respectively. In order to make our multiple hypotheses testing results replicable, we use the seed number of 20.
[26] Results for the whole sample show a similar pattern (Figure A4).
[27] For example, only 17% and 43 % of those who circumcised at the study hospital recall the month-year and year of the circumcision correctly at the time of the follow-up survey, respectively.
[28] We have estimated the impacts on HSV2 and HIV infections after removing those students who were already circumcised at baseline in Table A9 and found that our results do not change.

(Group 1 and Group 2) in order to increase statistical power. In Columns (1)-(6) we present results from the full sample (9th-11th graders) while in Columns (7)-(12) those from the 9th and 10th graders included in the long-term follow-up survey. The dependent variables are male circumcision take-up from hospital administration data at the end of Round 1 (Columns (1)-(3) and (7)-(9)), and Round 2 (Columns (4)-(6) and (10)-(12)). For most specifications in the result section, we present regression results using three different sets of controls: (1) no controls, (2) grade fixed effects, socio-demographic controls and controls for AFF's HIV/AIDS education and girls' CCT programs[29], and (3) same controls as in (2) but also school fixed effects. Our preferred specifications use the controls in group (2) as they follow the pre-analysis plan.

The results confirm that the free circumcision and transport intervention for male circumcision significantly increased the demand for male circumcision during Round 1. At the end of Round 1, Group 1 and Group 2 were 13.8 and 17.0 percentage points (288% and 354%) more likely to get circumcised than those in control classrooms (Group 4), respectively (Column (2)). These effects are sustained until the end of Round 2 (Columns (4)-(6)), although they become smaller, which is not surprising given the less intense intervention during Round 2. We find similar pattern for 9th and 10th graders (Columns (7)-(12)). In Panel B of Table 3 we show that for our preferred specification when we use the pooled treatment groups (Group 1 and Group 2), the demand for circumcision for being assigned to treatment in Round 1 increases by 14.9 percentage points at the end of Round 1 and by 10.8 percentage points at the end of Round 2. In Table A4, we also present similar regressions to the models presented in Table 3 but instead in Columns (4)-(6) we use the self-reported circumcision status measured at the time of the 2nd follow-up survey as our outcome variable. The results confirm that those exposed to the free-circumcision and transport intervention in Round 1 had also higher circumcision rates in the long term. The point estimates are larger than those using the administrative data at the end of Round 2, although they are less precisely estimated.[30] [31]

We further test whether the take-up of male circumcision is heterogeneous by prior beliefs such as knowledge about the benefits of male circumcision, beliefs about how painful male circumcision is,

---

[29] We have also implemented specifications where these additional two interventions were also interacted with the circumcision treatments and did not find that they make a difference to our results. The results from these interactions with the HIV/AIDS education program are reported in Appendix Table A6.
[30] The correlation between hospital reports and self-reports of circumcision status is about .7 and does not vary by treatment arms.
[31] Hospital administrative data is not available during the time when the long-term follow-up survey was implemented.

and beliefs that male circumcision is appropriate only for Muslims (Table A5). In general, those who think that male circumcision is painful or that it is only for appropriate for Muslims are less likely to receive male circumcision in general. The interactions of these variables and the male circumcision variable are also negative, suggesting that individuals with these prior beliefs are less responsive to the circumcision intervention, although we note that many of the coefficients are imprecisely estimated.

The estimated impacts of our intervention on male circumcision take-up are generally larger than the previous studies on this topic reviewed earlier that are based on financial incentives (Thirumurthy et al. (2014), Evens et al. (2016), Wilson et al (2016), Chinkhumba et al. (2014)). The size of our impacts is closer in magnitude to the soccer based incentives intervention of Kaufman et al (2016), a study that just like ours is based on an adolescent sample. However, while the age of the sample as well as other factors could potentially explain the size of our effects, it is worth emphasizing that our implementing NGO gave our student population multiple options over a period of time to reach the circumcision clinic.[32]

5.2. Peer effects on male circumcision take-up

As discussed previously, we estimate different types of peer effects based on Equations (1), (2), and (3). We start with results presented in Panel A of Table 3 that use Equation (1). Those in Group 3, containing untreated students in *50% Treatment* classrooms, were about 4 percentage points more likely to receive male circumcision than students in Group 4 (*No treatment*) classroom (Columns (1)-(3)). We interpret this result as the spillover effect within the classroom from the other half of classmates who received the offer of male circumcision in Round 1. Interestingly, the difference between Group 3 and Group 4 increases to about 6-7 percentage points at the end of Round 2 (Columns (4)-(6)). It implies that those in Group 3, who had peers who experienced male circumcision within their classroom (Group 2), were more responsive to the offer than the control group (Group 4). These differences also persist after 4 years when we use self-reported circumcision status as the outcome in Table A4.

A second way to look at the existence of spillovers is to test whether take-up rates of Group 1 and Group 2 are different given that everybody in Group 1 classroom are treated in Round 1 while Group 2 had peer who did not received an offer initially. The peer effect here is different from the one measured by Group 3 and Group 4. For Group 3, half of the peers received the treatment, while no one

[32] See also the summary by Sgaier et al. (2016) of seven studies in a special issue of JAIDS (2016) on this topic.

in Group 4 received the treatment. Somewhat surprisingly, we do not find evidence of spillovers using these two groups, suggesting nonlinear peer effects related to the take-up decision. The difference in take up is not statistically significant and if anything, the take-up rate is larger in Group 2 than Group 1.[33]

While our interpretation of the results for Group 3 in Table 3 is that they capture peer effects in the decision to get circumcised, an alternative interpretation is that students in Group 3 benefited from the information session about male circumcision during Round 1 that was provided to 50% treatment classrooms (Group 2 and 3), and this might have induced them to get circumcised. We provide evidence against this alternative explanation by using the fact that AFF randomly provided HIV/AIDS information and education as part of their HIV/AIDS prevention program. This HIV/AIDS education intervention was more comprehensive and intensive than the information session on male circumcision given to the treatment classrooms, but the component about circumcision in the broad education intervention was identical to the circumcision information session (the one provided to Group 1, Group 2, and Group 3, and not to Group 4). Since the HIV/AIDS education intervention was cross-randomized, we can look at the interaction of being assigned to the Group 3 and being assigned to the education intervention. These results are presented in Table A6, and they show that both the HIV/AIDS education coefficient and the interaction between HIV/AIDS education and being assigned to Group 3 are small or statistically insignificant. These results suggest that information provided in school by the NGO is unlikely to account for the take-up of male circumcision for the students assigned to Group 3.

Another way to study peer effects is to estimate Equations (2) and (3) using the restricted sample of *50% Treatment* classrooms (Group 2 and Group 3). In this analysis we take advantage of the details of the self-reported friend networks. As Table 4 shows, we find that having a higher proportion of friends who are treated increases one's circumcision take up in general (Panel B) however the coefficients are not statistically significant. Panel C provides evidence of a complementarity between an offer to student and his friend that increase a student's take-up of circumcision. The impacts are large, statistically significant, and robust across the specifications, which suggests the existence of important reinforcement effects among peers in classrooms.

[33] One potential explanation of this (non-significant) difference might be the scarcity heuristic argument, which states that when a resource is less readily available people are more likely to perceive it as more valuable (Cialdini, (2009)).

In sum, the results in this section highlight that peer effects within classrooms played an important positive role in the decision to take up circumcision by the secondary school male students enrolled in the study and consistent with a reinforcement effect, they are even stronger when both friends are assigned to the male circumcision treatment.

5.3 Long-term impacts of male circumcision on STD infections and sexual behaviors

5.3.1 Impacts on STD infections

In this section, we study the long-term impacts of male circumcision on HSV-2 and HIV infections, sexual behaviors measured by Item Count Technique (ICT), and a range of self-reported sexual behaviors. As explained previously, one of the major concerns related to the scale up of male circumcision is possible risk compensation resulting from circumcised men engaging in risky sexual behaviors. In this analysis, we focus only on the baseline 9th and 10th graders who were included in the long-term follow-up survey.

The main results are presented in Table 5. The results from the preferred specification as specified in the pre-analysis plan are shown in Columns (2), (5), (8), and (11). Panel A of Table 5 shows that the cumulative probability of HSV-2 infection measured by IgG in Group 1 and Group 2 *increases* by about 1.9 and 4.2 percentage points after four years, respectively, and the corresponding increases for IgM are 1.3 and 2.0 percentage points. Interestingly, we do not find such evidence for Group 3 even though circumcision take up rate of Group 3 are comparable at the end of the experiment (Column (11) of Table 3) and the long-term follow-up survey (Panel B of Figure 1). Lastly, we do not find a long-term impact of the intervention on testing HIV positive, although we note that this study was not powered to detect impacts on HIV infection.

In Panel B, when we combine Group 1 and Group 2 as before, the results are similar and they show that male circumcision induced by the direct free circumcision and transport intervention leads to *increases* in STD infections. The results are statistically significant both when we use the standardized treatment effect in Column (11) of Panel B as well as individually for IgG (Column (2)) and IgM (Column (5)) where we use the free step-down resampling method for multiple hypotheses testing to adjust the familywise error.[34][35] Column (11) of Panel A and B also confirm that we can reject the hypothesis that the impact of the direct free circumcision offer with transport support on STD infections is the same as the indirect effect coming from peer effects (p-value for F-test is 0.004).

In Table A7 we further explore the impact of the intervention on HSV-2 infections. According to the medical literature, testing positive for IgG is a marker of having been infected at least once during your lifetime and it is also the case that this test takes several months to become positive after infection. In contrast, the IgM test can be detected quicker after infection and is a marker of a more recent infection. It also increases during HSV-2 recurrence (Morrow and Friedrich, 2006). As a result, a positive IgG and IgM test is capturing a new infection or recurrence.[36] Given the 4 year follow-up of our study, it would be unlikely for the intervention to have an impact on recent HSV-2 infections in the short period prior to the biomarker tests at the long-term follow-up (i.e. IgM positive without being IgG positive), but we should see IgM positive tests that occur with a positive IgG test. The regression results in Table A7 confirm this intuition. One can observe that the intervention had an impact on IgM only when it happened concurrently with a positive IgG result (Columns (7)-(9)) and not with a negative IgG result (Columns (4)-(6)).[37]

An additional way to explore these long-term impacts of male circumcision on STD infections is to present specifications that are similar to Table 4 using just the 50% classrooms and the friendship network information. In Panels A and B of Table 6 we confirm that STD infections for those in Group 1 and Group 2 in these classrooms receiving the direct free circumcision offer with transport support treatment significantly increases the standardized treatment effect in the STD domain (Column (11) of Panels A and B). However, when we turn in Panel C to specifications based on Equation (3), we observe that the positive reinforcement effects among peers in school on the decision to get circumcised shown

[34] We have also investigated how our estimates in Panel B of Table 5 are potentially impacted by attrition using Lee's bounds (Lee (2009)) in Table A10 and found no evidence of biased treatment effects due to nonrandom sample attrition.
[35] Another robustness check that we have implemented is to see if our main results in Table 5 hold when we are restricting the analysis to the 28 schools in our sample that do not have multiple classrooms per grade. The results are similar and if anything the point estimates are larger (results available on request). This alleviates concerns about potential spillovers between the three groups of classrooms.
[36] A reinfection of HSV-2 is possible but unlikely (Bernstein et al, 1986)
[37] We are not able to explore possible heterogeneous treatment effect by HSV-2 status at baseline, because data on HSV-2 status was only collected at the time of the follow-up. This exercise might be important especially if HSV-2 positive and negative individuals react differently to the intervention. Wilson et al. (2014) and Kerwin (2020) provide evidence that lower HIV transmission risks can cause safer sexual behavior. To the extent that such fatalism is important, one might argue that it is more important among those who are sexually active at the time of the baseline. We have implemented this heterogeneity analysis and have not found evidence for a differential response (results available upon request).

in Table 4 does not translate to noticeable changes in STD infections, although these coefficients are not precisely estimated. These results are consistent with the results in Table 5 in that they show that peer-driven increases in male circumcision do not translate into higher rates of STD infections.

5.3.2. Impacts on sexual behaviors

Next, we study impacts on sexual behaviors. Table 7 presents results using the Item Count Technique (ICT) to analyze responses to two sensitive risky sexual behavior: "I think I have to use a condom in case of sex with somebody that I do not know well" (Columns (1)-(3)) and "I had sex with more than two people in last 12 months" (Columns (4)-(6)). As a reminder, half of the survey respondents were randomly given a set of straightforward true/false statements, while the other half also received the additional sensitive statement about risky sexual behaviors.

Table 7 shows the regression results of the following estimation:

(4) $Y_{ijgs} = \beta_0 + \sum_{a=1}^{3} \beta_a \cdot G_{jgs} + \kappa \cdot long_{ijgs} + \sum_{a=1}^{3} \lambda_a \cdot G_{jgs} \cdot long_{ijgs} + \gamma' X_{ijgs} + \delta_g + \varepsilon_{ijgs}$,

where $long_{ijgs}$ is a binary indicator if individual i was assigned to a questionnaire with sensitive questions (*long form*). λ_a are the coefficients of interest that capture the difference in the fraction of respondents who respond positively to the sensitive question between the three treatment groups (Group 1, Group2, and Group3) and the control group (Group 4).

The ICT results suggest that the percent change in respondents who use condom in case of a casual sexual encounter decreases by 25.6% for Group 1 and 22.3% for Group 2 (Panel A). In Columns (4)-(6), we do not find statistically significant changes in reporting the multiple sex partner question. Table 7 also confirms that the changes in sexual behaviors for Group 3 that increased circumcision take-up through the peer channel are not significant. When we use both questions in the self-reported ICT sexual behavior domain, the standardized treatment effects reported in Columns (7)-(9) are not statistically significant.[38] The results in Panel B, which looks at the combined effect of Group 1 and Group 2, are similar to those reported in Panel A. Among those in the direct free circumcision and transport treatment group (Group 1 and Group 2), only 37.9% of students think that they have to use a condom in case of a casual sexual encounter while the corresponding number for the control group is

[38] We note that for the standardized treatment effects, the answers to "using condoms for casual sex" have been recoded so that a positive coefficient reflects an increase in reporting risky sexual behaviors.

62.2%.

We also measure the impact of the intervention on condom purchases and nine self-reported sexual behaviors, such as age at sexual debut or indicator variables for being sexually active or having multiple sexual partners (Table 8). The general pattern of results is that most coefficients on Group 1 and Group 2 are positive, suggesting that the direct intervention leads to riskier sexual behavior, but these are not statistically significant. When we combine all of these outcomes to estimate the standardized treatment effect from the domain of self-reported sexual behavior, the effects are also positive. But just as before, these impacts are not significant at conventional levels, except for the effect of Group 2 in the long-term follow-up (Column (11) of Panel B1) and for the joint effect of Group 1 and Group 2 (Column (11) of Panel B2)

In sum, although sexual behaviors are subject to social desirability bias, we find a general pattern of increases in risky sexual behaviors both when using the ICT as well as from self-reported sexual behaviors. In addition, it is worth noting that all our results that are statistically significant are related to condom use: The increase of reporting an inconsistent use of condom (Column (7)) and unprotected sex with last partner (Column (8)) are mirrored in the ICT results on the changes in attitude regarding condom use for casual sex (Column (1)-(3) of Table 7).

5.3.3. Simulation of STD infection rate with risk compensation

Our main findings so far show that those directly treated (Group 1 and Group 2) are more likely to be circumcised than the control group (Group 4) by 14.1% points (17.6% vs. 3.5%), however, those treated are more likely to be infected by HSV-2 than the control group by 2.7% points (11.1% vs. 8.4%). These results are larger than the existing literature reviewed previously, which generally finds either limited evidence of risk compensation (Auvert et al. (2005), Bailey et al. (2007), Gray et al. (2007), Mattson et al. (2008), Gray et al. (2012) and Kong et al. (2012)) or even the opposite of risk compensation (Wilson et al., 2014)[39]. In this section, we summarize the results from a simple simulation exercise that is explained in further detail in Appendix B, which tries to understand whether our large effects can be explained through reasonable levels of risk compensation taken from the available literature. In this model, the HSV Infection rate is defined as follows:

[39] Wilson et al. (2014) argue that the mechanism behind their finding is that circumcision could reduce fatalism about getting infected with HIV/AIDS.

(5) HSV-2 Infection rate = Frequency of sexual activity × Probability (risky sex) × Transmission rate

The HSV-2 infection rate is the product of the frequency of sexual activity during the study period, the probability of having sexual intercourse without protection, and the HSV-2 transmission rate. Male circumcision is known to decrease the HSV-2 transmission rate by half and thus lowers the overall HSV-2 infection level (Sobngwi-Tambekou et al., 2009). However, there are behavioral adjustments that might be associated with male circumcision. First, male circumcision could ex ante appeal to individuals with an inclination towards risky sex. This implies that there is a selection effect, namely, those who are more likely to engage in risky sex, have a higher tendency to decide to get the circumcision procedure. Second, male circumcision induces ex post risk compensating behaviors. With a lower probability of infection, the circumcised could be more likely to engage in risky sexual behaviors, they could increase the frequency of sex, or both.

Our simulation exercise examines the effect of male circumcision on the infection rate taking into account selection effects and risk compensation. Risk compensation is modeled as separate parameters, namely an increase in the probability of risky sex and a higher number of sex partners.

We make the following assumptions. First, we assume that both treatment and controls groups consist of risky and safe individuals that differ in the frequency of sexual activity and the probability of risky sex. It is a reasonable assumption under our RCT setting. Second, we also assume that the risky group is more likely to take-up male circumcision than the safe group. We present four scenarios, where the proportion of individuals who take-up circumcision from the risky and safe groups is (100%, 0%), (90%, 10%), (80%, 20%), and (70%, 30%), but (80%, 20%) will be our preferred scenario. Third, we also assume that the medical impact of male circumcision is that it reduces the transmission rate by half based on findings from the literature. However, the rate also differs by whether one uses a condom or not. We use the point estimates and confidence intervals from Magaret et al. (2016) about the HSV-2 transmission rate without and without a condom and parametrize the transmission rates, given that they are point estimates with large confidence interval.

We present results from two scenarios. In the first scenario, which is the most likely to find large treatment effects, we choose the largest possible transmission rate without protection (0.0044) and the smallest transmission rate with a condom (0.0002) shown in Magaret et al. (2016). In the second scenario, we choose the point estimates from Magaret et al (2016), which are 0.0017 with a condom and

0.0006 without a condom. Finally, we assume that the average frequency of sexual intercourse without circumcision is 60 during our study period, but note that our main conclusions are robust to any frequencies ranging between 40 and 300 (results are available by request).

The first step is to calculate the infection rate of the risky and safe groups by circumcision status. Then, the population infection rate is simply the weighted sum of the groups with weights determined by the take-up parameter. We run simulations with Python to find a set of parameters that generate results consistent with the findings from our study. The main results of this simulation exercise are best shown graphically. In Panels A and B of Figure 2, we show for our two scenarios how the treatment group infection rate varies with the percentage increase in the probability of having a risky sexual encounter and the increase in the percentage increase in the frequency of sexual encounters. In Panel A, the areas that are more yellow correspond to treatment group infection rates around 11.1%.[40] One can observe that such infection rates are achieved when there is a doubling in the frequency of sexual encounters and an increase in the probability of risky sexual behavior of about 35%. In Panel B, we show the results for the second scenario using the point estimates for infection using and not using condom protection. In this scenario, the maximum infection rate that we can reach with our simulation is 9.5%. While this result is smaller than the point estimate of our treatment effect (11.3%), we note that our treatment effect is estimated with noise, and therefore the confidence interval contains the treatment effects from our second scenario as well as other scenarios that require much smaller selection and risk compensation effects. In sum, this simulation exercise results suggest that the relatively large treatment effects that we estimate could be explained by changes driven by a combination of a selection effect as well as large but reasonable changes in risk compensation.

5.4 Heterogeneity of treatment effects

In the previous section, we show that the HSV-2 infection rate could significantly increase due to the risk compensation and selection. We try to explore potential reasons for these heterogeneous treatment responses in this section.

5.4.1 Understanding complier characteristics

In order to study why our biomarker outcomes from Group 1 and Group 2 (hereafter Group 1&2) and Group 3 are different, we first compare complier characteristics of Group 1&2 with Group 3 by

[40] In Figure A5, we also show the corresponding results for the (100%, 0%), (90%, 10%), and (70%, 30%).

restricting the sample to circumcision takers following Kim and Lee (2017) (Table 9). Since everyone has undergone circumcision in the restricted sample, any difference between Group 1&2 and Group 3 is due to the compositional change of circumcision takers. In this sample, circumcision takers of Group 1&2 are always takers and compliers driven by the free circumcision and transport support intervention while circumcision takers of Group 3 are always takers and compliers driven by peer effects. Thus, using the analysis with the restricted sample allows us to compare the characteristics of two different types of compliers. Our choice of characteristics for this analysis includes a SES index to capture broad differences in socio-economic status, as well as self-reported sexual behaviors measured in the baseline (HIV/AIDS knowledge, sexual behavior risk index). We also use a measure of risk preferences, based on the risk-domain lab-in-the-field experimental tools implemented by Choi et al. (2007), as well as measures of agreeableness and conscientiousness based on the ten-item Big Five personality inventory (Gosling et al. (2003)).[41]

The results presented in Table 9 provide suggestive evidence that is consistent with the view that the characteristics of compliers in Group 1&2 compared to Group 3 make the former more prone to engage in risky sexual behaviors, which helps explain the heterogeneous biomarker outcomes. More specifically, compliers of Group 1&2 are more likely to have risk neutral or risk seeking behaviors, they exhibit less agreeableness and less conscientiousness, but the difference in conscientiousness is not significant at the 5% level in our preferred specification. In fact, Schmitt (2004) documents associations between relationship infidelity and low agreeableness and low conscientiousness across all regions of the world. Moreover, compliers in Group 1&2 score lower in terms of HIV/AIDS knowledge, although we note that we do not find meaningful differences for the SES and self-reported risky sexual behavior index.[42] These results suggest that characteristics associated with risk tolerance and the aforementioned two personality traits are potentially driving the differences in HSV2 infections between these groups.

5.4.2 Exploring heterogeneity using machine learning methods

Next, we follow a recent machine learning technique proposed by Chernozhukov et al. (2018) to estimate and make inference on core features of heterogeneous treatment effects. We implement this

[41] While one might worry that these outcomes, which are measured in the long-term follow-up survey, might themselves be affected by the intervention, but we do not find empirical support for this possible concern (results are available upon request).
[42] We construct a risky sexual behavior scale using 8 self-reported risky sexual behavior items based on item response theory (Mattson et al. (2008)).

method to explore treatment heterogeneity and relevant characteristics driving heterogeneous effects in the two groups exposed to the direct intervention and the peer effect channels (Group 1&2 versus Group 3). Rather than relying on the standard approach navigating statistically significant interaction effects, Chernozhukov et al. (2018) apply machine learning tools to generate "proxy predictors" for the conditional average treatment effect (CATE) - the difference in the expected potential outcomes between treated and control groups conditional on covariates.

We first verify heterogeneous treatment effects in the male circumcision take-up and illustrate a responsive subgroup of students to the direct intensive free circumcision and transport offer, whose characteristics differ from those students who responded to the peer effect channels. In particular, Table 10 and 11 present three results described by Chernozhukov et al. (2018), which are 1) the Best Linear Predictor (BLP) of CATE (Table 10), 2) the Sorted Group Average Treatment Effects (GATES) or average treatment effect by heterogeneity groups (Panel A of Table 11) and 3) the Classification Analysis (CLAN) or the average characteristics of the most and the least affected subgroups induced by the machine learning proxy predictor (Panel B of Table 11).

Table 10 presents estimation results of the BLP of CATE, presenting the average treatment effects (ATE) and heterogeneity loading (HET) parameters that determines heterogeneity in treatment effect by covariates. Confidence intervals in parentheses are reporting the medians of interval sets resulting from 100 sample splits, and p-values in brackets are also constructed and adjusted in the same sense. We first find that the estimated ATEs of the MC offer (Group 1&2 vs. Group 3) are similar to the unconditional ATE in our main tables when using the Elastic Net and Neural Network methods[43]. Importantly, we detect heterogeneity in the treatment effect by rejecting the null hypothesis that HET is zero at the 10% level for the MC take-up (Round 1) between Group 1&2 vs. Group 3.

In Panel A of Table 11, we first divide the students into five groups based on the machine learning proxy predictors for the expected potential circumcision take-up in Round 1 by covariates, and estimate and compare the average effect for the most and the least affected groups. Both techniques show that the difference of GATES of these two groups is significantly different from zero at the 10% level on circumcision take-up in Round 1 (Columns (3) and (6)). Finally, in Panel B, we ascertain which

[43] Table A8 shows that the Elastic Net and Neural Network outperform the Boosting and Random Forest for our main outcome, male circumcision take-up. Following Chernozhukov et al. (2018), we focus on these two tools for the rest of the analysis.

of the covariates are correlated with the heterogeneity by comparing the average characteristics of the most and the least affected groups. To be consistent with the previous complier characteristics analysis, we focus on four covariates: SES index measured at the baseline, and risk tolerance, agreeableness and conscientiousness measured at the long-term follow-up survey. We find students with higher SES index, higher risk tolerance, less agreeableness and conscientiousness are more likely to respond to the direct male circumcision offer. These differences are statistically significant at least at the 5% level, and the significance is not sensitive to particular machine learning methods used to generate the proxy predictors.

We interpret these results in section 5.4 that both analyses provide evidence consistent with the view that take-up of male circumcision across the different groups is heterogeneous and generated by characteristics such as risk preferences and personality traits, which are consistent with the main finding that the direct free circumcision and transport interventions (Group 1&2) and the indirect peer effects (Group 3) have different long-term impacts on HSV-2 infection rates.

6. Conclusion

This paper examines the demand for male circumcision and its long-term consequences in Malawi, a country that is heavily affected by the HIV epidemic. It is based on a two-stage randomized controlled trial that randomly provided free male circumcision with transport support to 3,970 students across 33 public secondary schools and extensive short and long-term survey data that were complemented with administrative data on clinical circumcisions and HIV and HSV-2 biomarkers.

We view this paper as having made four contributions to the literature. First, we show that the school-based intervention substantially increased the demand for male circumcision by an average of 15.0 percentage points. Secondly, we find empirical evidence for positive peer effects among classroom peers. These peer effects appear to be stronger when two friends are treated together. Third, the results show that the preventive effects of male circumcision against HIV and HSV-2 appear to be mitigated in the long-run through risk compensation. For the group induced to take up circumcision by the direct intervention, we detect a significant *increase* of HSV-2 infection measured by IgG and IgM after four years from the treatment. This effect occurs despite the evidence from the medical literature that male circumcision decreases HSV-2 infection biologically, when holding sexual activities constant. Fourth, we find evidence of risk compensation only among those who were induced to take-up circumcision through the direct incentive intervention, but not among those who were induced through the peer effect.

Our findings based on biomarkers are different from results in several efficacy trials that showed protective effects of 51-60% against HIV acquisition (Auvert et al. (2005), Bailey et al. (2007), Gray et al. (2007)) and protective effects of 28-30% against HSV-2 infection (Sobngwi-Tambekou et al. (2009), Tobian et al. (2009)). Unlike much of the existing evidence on this topic that is based on clinical efficacy trials with shorter follow-up periods, our setting is based on a community based long-term scale-up project implemented by an NGO where the preventive effect of circumcision against HIV is widely known and people can voluntarily accept or decline free male circumcision services. This is important because our setting might more closely resemble what the situation is likely to be under non-study conditions.

A number of implications for public policies related to the scale-up of male circumcision can be drawn from the results of this study. First, while a lack of access to male circumcision is still a major barrier, we show that free male circumcision with well-designed incentives such as transport support can increase demand for male circumcision substantially. Maybe most importantly, we show that male circumcision scale-up projects might be less successful than expected in preventing HIV and HSV-2 infections because those who get circumcised are more likely to engage in risk compensation. These results suggest that male circumcision scale-up projects should be combined with programs to address risk compensation among circumcised men such as intensive public health and education campaigns.

At the same time, our results also have implications for research on peer effects, and more generally about heterogeneous treatment effects coming from differential selection into take-up of alternative development interventions. While this paper sheds light on the important role that peer effects play in the decision to get circumcised, we also show that those who took up circumcision as a result of their peers might display different behavioral responses compared to those who got circumcised through the direct incentive intervention. In our specific case, two separate channels that both increased demand for male circumcision lead to very different rates of sexually transmitted diseases in the longer term. Our analysis suggests that this is because individuals who are getting circumcised because of peer influence have on average risk preferences and personality traits that are less conducive to engaging in risky sexual behaviors. This implies that future work needs to pay added attention to understanding how the heterogeneous take-up of a development intervention affects the desired socio-economic outcomes.

References

Auvert, Bertran, Dirk Taljaard, Emmanuel Lagarde, Joelle Sobngwi-Tambekou, Remi Sitta, and Adrian Puren. 2005. "Randomized, Controlled Intervention Trial of Male Circumcision for Reduction of HIV Infection Risk: The ANRS 1265 Trial." *PLoS Medicine* 2 (11):e298.

Bailey, Robert C, Stephen Moses, Corette B Parker, Kawango Agot, Ian Maclean, John N Krieger, Carolyn FM Williams, Richard T Campbell, and Jeckoniah O Ndinya-Achola. 2007. "Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial." *The Lancet* 369 (9562):643–656.

Baird, Sarah J, Richard S Garfein, Craig T McIntosh, and Berk Ozler. 2012. "Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomised trial." *The Lancet* 379 (9823):1320–1329.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2010. "Social Incentives in the Workplace." *Review of Economic Studies* 77 (2):417–458.

Bengo, Joseph Mfutso, Kondwani Chalulu, Jobiba Chinkhumba, Lawrence Kazembe, Kenneth M Maleta, Francis Masiye, and Don Mathanga. 2010. "Situation Analysis of Male Circumcision in Malawi." College of Medicine, University of Malawi.

Bernstein, David I., et al. "Antibody response, recurrence patterns and subsequent herpes simplex virus type 2 (HSV-2) re-infection following initial HSV-2 infection of guinea-pigs: effects of acyclovir." Journal of general virology 67.8 (1986): 1601-1612.

Bingenheimer, Jeffrey B. and Arline T. Geronimus. 2009. "Behavioral Mechanisms in HIV Epidemiology and Prevention: Past, Present, and Future Roles." *Studies in Family Planning* 40 (3):187–204.

Brooks, Ronald A., Mark Etzel, Lee E. Klosinski, Arleen A. Leibowitz, Sharif Sawires, Greg Szekeres, Mark Weston, and Thomas J. Coates. 2010. "Male Circumcision and HIV Prevention: Looking to the Future." *AIDS and Behavior* 14 (5):1203–1206.

Cassell, Michael M, Daniel T Halperin, James D Shelton, and David Stanton. 2006. "Risk compensation: the Achilles' heel of innovations in HIV prevention?" *BMJ: British Medical Journal* 332 (7541):605–607.

Chernozhukov, Victor, Mert Demirer, Esther Duflo, and Ivan Fernandez-Val. 2018. "Generic Machine Learning Inference on Heterogenous Treatment Effects in Randomized Experiments." w24678, National Bureau of Economic Research, Cambridge, MA.

Chinkhumba, Jobiba, Susan Godlonton, and Rebecca Thornton. 2014. "The Demand for Medical Male Circumcision." *American Economic Journal: Applied Economics* 6 (2):152–177.

Choi, Syngjoo, Raymond Fisman, Douglas Gale, and Shachar Kariv. 2007. "Consistency and Heterogeneity of Individual Behavior under Uncertainty." *American Economic Review* 97 (5):1921–1938.

Chong, Alberto, Marco Gonzalez-Navarro, Dean Karlan, and Martín Valdivia. 2020. "Do Information Technologies Improve Teenagers' Sexual Education? Evidence from a Randomized Evaluation in Colombia." The World Bank Economic Review, 34(2), pp.371-392.

Cialdini, Robert B. 2009. *Influence: science and practice*. Boston: Pearson Education, 5th ed.

Coffman, Katherine B., Lucas C. Coffman, and Keith M. Marzilli Ericson. 2017. "The Size of the LGBT Population and the Magnitude of Antigay Sentiment Are Substantially Underestimated." *Management Science* 63 (10):3168–3186.

Djimeu, E, Brown, A. 2016 "Strength of Evidence on Demand Creation for Voluntary Medical Male Circumcision From 7 Impact Evaluations in Southern and Eastern Africa." J Acquir Immune Defic Syndr. 72(suppl 4): S321-S325.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2015. "Education, HIV, and Early Fertility: Experimental Evidence from Kenya." *American Economic Review* 105 (9):2757–2797.

Duflo, Esther and Emmanuel Saez. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *The Quarterly Journal of Economics* 118 (3):815–842.

Eaton, Lisa and Seth C Kalichman. 2009. "Behavioral Aspects of Male Circumcision for the Prevention of HIV Infection." *Current HIV/AIDS reports* 6 (4):187–193.

Evans, William N. and John D. Graham. 1991. "Risk Reduction or Risk Compensation? The Case of Mandatory Safety-Belt Use Laws." *Journal of Risk and Uncertainty* 4 (1):61–73.

Evens, Emily, Michele Lanham, Catherine Hart, Mores Loolpapit, Isaac Oguma, and Walter Obiero. 2014. "Identifying and Addressing Barriers to Uptake of Voluntary Medical Male Circumcision in Nyanza, Kenya among Men 18-35: A Qualitative Study." *PLoS ONE* 9 (6):e98221.

Fafchamps, Marcel, Ana Vaz, and Pedro C. Vicente. 2020. "Voting and Peer Effects: Experimental Evidence from Mozambique." *Economic Development and Cultural Change* 68 (2): 567–605.

Fafchamps, Marcel and Pedro C. Vicente. 2013. "Political violence and social networks: Experimental evidence from a Nigerian election." *Journal of Development Economics* 101:27–48.

Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, and Oregon Health Study Group. 2012. "The Oregon Health Insurance Experiment: Evidence from the First Year." *The Quarterly Journal of Economics* 127 (3):1057–1106.

Foster, Gigi. 2006. "It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism." *Journal of Public Economics* 90 (8-9):1455–1475.

Gao Y, Yuan T, Zhan Y, et al. 2021. "Association between medical male circumcision and HIV risk compensation among heterosexual men: a systematic review and meta-analysis". *Lancet Glob Health* 9(7): e932-e941.

Godlonton, Susan and Rebecca Thornton. 2012. "Peer effects in learning HIV results." *Journal of Development Economics* 97 (1):118–129.

Gosling, Samuel D, Peter J Rentfrow, and William B Swann. 2003. "A very brief measure of the Big-Five personality domains." *Journal of Research in Personality* 37 (6):504–528.

Gray, Ron, Godfrey Kigozi, Xiangrong Kong, Victor Ssempiija, Frederick Makumbi, Stephen Wattya, David Serwadda, Fred Nalugoda, Nelson K. Sewenkambo, and Maria J. Wawer. 2012. "The effectiveness of male circumcision for HIV prevention and effects on risk behaviors in a posttrial follow-up study." *AIDS* 26 (5):609–615.

Gray, Ronald H, Godfrey Kigozi, David Serwadda, Frederick Makumbi, Stephen Watya, Fred Nalugoda, Noah Kiwanuka, Lawrence H Moulton, Mohammad A Chaudhary, Michael Z Chen, Nelson K Sewankambo, Fred Wabwire-Mangen, Melanie C Bacon, Carolyn FM Williams, Pius Opendi, Steven J Reynolds, Oliver Laeyendecker, Thomas C Quinn, and Maria J Wawer. 2007. "Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial." *The Lancet* 369 (9562):657–666.

Jones, Damon, David Molitor, and Julian Reif. 2019. "What Do Workplace Wellness Programs Do? Evidence from the Illinois Workplace Wellness Study." *The Quarterly Journal of Economics* 134 (4): 1747–1791.

Jung, Jaehyun et al. 2016. "Long-term Effects of Male Circumcision on Risky Sexual Behaviors and STD Infections: Evidence from Malawian Secondary Schools." AEA RCT Registry No.0001335. June 12. https://doi.org/10.1257/rct.1335-2.0

Kalichman, Seth, Lisa Eaton, and Steven Pinkerton. 2007. "Circumcision for HIV Prevention: Failure to Fully Account for Behavioral Risk Compensation." *PLoS Medicine* 4 (3):e138.

Kaufman Z, DeCelles J, Bhauti K, et al. 2016 "A Sport-based intervention to increase uptake of voluntary medical male circumcision among adoles- cent male students: results from the MCUTS 2 cluster-randomized trial in Bulawayo, Zimbabwe." J Acquir Immune Defic Syndr. 72(suppl 4): S297–S303.

Kim, Hyuncheol Bryant, Syngjoo Choi, Booyuel Kim, and Cristian Pop-Eleches. 2018. "The Role of Education Interventions in Improving Economic Rationality." *Science* 362 (6410):83–86.

Kim, Hyuncheol Bryant and Sun-mi Lee. 2017. "When public health intervention is not successful: Cost sharing, crowd-out, and selection in Korea's National Cancer Screening Program." *Journal of Health Economics* 53:100–116.

Kong, X., G. Kigozi, F. Nalugoda, R. Musoke, J. Kagaayi, C. Latkin, R. Ssekubugu, T. Lutalo, B. Nantume, I. Boaz, M. Wawer, D. Serwadda, and R. Gray. 2012. "Assessment of Changes in Risk Behaviors During 3 Years of Posttrial Follow-up of Male Circumcision Trial Participants Uncircumcised at Trial Closure in Rakai, Uganda." *American Journal of Epidemiology* 176 (10):875–885.

Kremer, Michael. 1996. "Integrating Behavioral Choice into Epidemiological Models of AIDS." *The Quarterly Journal of Economics* 111 (2):549–573.

Kremer, Michael and Dan Levy. 2008. "Peer Effects and Alcohol Use among College Students." *Journal of Economic Perspectives* 22 (3):189–206.

Lee, David S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *The Review of Economic Studies* 76 (3):1071–1102.

LeGoff, Jerome, Helene Péré, and Laurent Bélec. 2014. "Diagnosis of genital herpes simplex virus infection in the clinical laboratory." *Virology Journal* 11 (1):83.

Lyle, David S. 2007. "Estimating and Interpreting Peer and Role Model Effects from Randomly Assigned Social Groups at West Point." *Review of Economics and Statistics* 89 (2):289–299.

Magaret, Amalia S., Andrew Mujugira, James P. Hughes, Jairam Lingappa, Elizabeth A. Bukusi, Guy DeBruyn, Sinead Delany-Moretlwe et al. 2016. "Effect of condom use on per-act HSV-2 transmission risk in HIV-1, HSV-2-discordant couples." *Clinical Infectious Diseases*, 62(4), pp.456-461.

Mas, Alexandre and Enrico Moretti. 2009. "Peers at Work." *American Economic Review* 99 (1):112–145.

Mattson, Christine L., Richard T. Campbell, Robert C. Bailey, Kawango Agot, J. O. Ndinya- Achola, and Stephen Moses. 2008. "Risk Compensation Is Not Associated with Male Circumcision in Kisumu, Kenya: A Multi-Faceted Assessment of Men Enrolled in a Randomized Controlled Trial." *PLoS ONE* 3 (6):e2443.

Miguel, Edward and Michael Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72 (1):159–217.

Miller, J. D. 1984. "A New Survey Technique for Studying Deviant Behavior." Ph.D. thesis, The George Washington University.

Minnis, A. M., M. J. Steiner, M. F. Gallo, L. Warner, M. M. Hobbs, A. van der Straten, T. Chipato, M. Macaluso, and N. S. Padian. 2009. "Biomarker Validation of Reports of Recent Sexual Activity: Results of a Randomized Controlled Study in Zimbabwe." *American Journal of Epidemiology* 170 (7):918–924.

Morrow, R., and D. Friedrich. "Performance of a novel test for IgM and IgG antibodies in subjects with culture-documented genital herpes simplex virus-1 or-2-infection." Clinical microbiology and infection 12.5 (2006): 463-469.

NAC. 2012. "Malawi Voluntary Medical Male Circumcision Communication Strategy 2012-2016." National AIDS Commission (NAC) Malawi.

National Statistical Office (NSO). 2017. "Statistical Yearbook 2017." Zomba, Malawi.

Obasi, Angela, Frank Mosha, Maria Quigley, Zebedayo Sekirassa, Tom Gibbs, Katua Munguti, James Todd, Heiner Grosskurth, Philippe Mayaud, John Changalucha, David Brown, David Mabey, and Richard Hayes. 1999. "Antibody to Herpes Simplex Virus Type 2 as a Marker of Sexual Risk Behavior in Rural Tanzania." *The Journal of Infectious Diseases* 179 (1):16–24.

Oster, Emily and Rebecca Thornton. 2012. "Determinants of Technology Adoption: Peer Effects in Menstrual Cup Take-up." *Journal of the European Economic Association* 10 (6):1263–1293.

Padian, Nancy S, Anne Buve, Jennifer Balkus, David Serwadda, and Ward Cates. 2008. "Biomedical interventions to prevent HIV infection: evidence, challenges, and way forward." *The Lancet* 372 (9638):585–599.

Palen, Lori-Ann, Edward A. Smith, Linda L. Caldwell, Alan J. Flisher, Lisa Wegner, and Tania Vergnani. 2008. "Inconsistent Reports of Sexual Intercourse Among South African High School Students." *Journal of Adolescent Health* 42 (3):221–227.

Peltzman, Sam. 1975. "The Effects of Automobile Safety Regulation." *Journal of Political Economy* 83 (4):677–725.

Sacerdote, Bruce. 2001. "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *The Quarterly Journal of Economics* 116 (2):681–704.

Sawires, Sharif R, Shari L Dworkin, Agnes Fiamma, Dean Peacock, Greg Szekeres, and Thomas J Coates. 2007. "Male circumcision and HIV/AIDS: challenges and opportunities." *The Lancet* 369 (9562):708–713.

Schmitt, David P. 2004. "The Big Five related to risky sexual behaviour across 10 world regions: differential personality associations of sexual promiscuity and relationship infidelity." *European Journal of Personality* 18 (4):301–319.

Sgaier S, Ridzon R, Sundaram M, et al. 2016 "Interventions to Drive Uptake of Voluntary Medical Male Circumcision—A Collection of Impact Evaluation Evidence." J Acquir Immune Defic Syndr. 72(suppl 4): S257-S261.

Sobngwi-Tambekou, Joelle, Dirk Taljaard, Pascale Lissouba, Kevin Zarca, Adrian Puren, Emmanuel Lagarde, and Bertran Auvert. 2009. "Effect of HSV-2 Serostatus on Acquisition of HIV by Young Men: Results of a Longitudinal Study in Orange Farm, South Africa." *The Journal of Infectious Diseases* 199 (7):958–964.

Thirumurthy, Harsha, Samuel H. Masters, Samwel Rao, Megan A. Bronson, Michele Lanham, Eunice Omanga, Emily Evens, and Kawango Agot. 2014. "Effect of Providing Conditional Economic Compensation on Uptake of Voluntary Medical Male Circumcision in Kenya: A Randomized Clinical Trial." *JAMA* 312 (7):703.

Thomas, Duncan, Elizabeth Frankenberg, and James P. Smith. 2001. "Lost but Not Forgotten: Attrition and Follow-up in the Indonesia Family Life Survey." *The Journal of Human Resources* 36 (3):556.

Thomas, Duncan, Firman Witoelar, Elizabeth Frankenberg, Bondan Sikoki, John Strauss, Cecep Sumantri, and Wayan Suriastini. 2012. "Cutting the costs of attrition: Results from the Indonesia Family Life Survey." *Journal of Development Economics* 98 (1):108–123.

Thompson, Diane C, Fred Rivara, and Robert Thompson. 1999. "Helmets for preventing head and facial injuries in bicyclists." *Cochrane Database of Systematic Reviews*.

Thornton, Rebecca L. 2008. "The Demand for, and Impact of, Learning HIV Status." *American Economic Review* 98 (5):1829–1863.

Tobian, Aaron A.R., David Serwadda, Thomas C. Quinn, Godfrey Kigozi, Patti E. Gravitt, Oliver Laeyendecker, Blake Charvat, Victor Ssempijja, Melissa Riedesel, Amy E. Oliver, Rebecca G. Nowak, Lawrence H. Moulton, Michael Z. Chen, Steven J. Reynolds, Maria J. Wawer, and Ronald H. Gray. 2009. "Male Circumcision for the Prevention of HSV-2 and HPV Infections and Syphilis." *New England Journal of Medicine* 360 (13):1298–1309.

UNAIDS. 2020. "UNAIDS DATA 2020." Joint United Nations Programme on HIV/AIDS, Geneva, Switzerland.

Weiss, Helen A, Kim E Dickson, Kawango Agot, and Catherine A Hankins. 2010. "Male circumcision for HIV prevention: Current research and programmatic issues." *AIDS (London, England)* 24 (0 4):S61–S69.

Westfall, Peter H. and S. Stanley Young. 1993. *Resampling-based multiple testing: examples and methods for P-value adjustment*. Wiley series in probability and mathematical statistics. New York: Wiley.

WHO/UNAIDS. 2007. "New Data on Male Circumcision and HIV Prevention: Policy and Programme Implications." Montreux, Switzerland.

WHO/UNAIDS. 2008. *Operational guidance for scaling up male circumcision services for HIV prevention*. Geneva, Switzerland: World Health Organization: Joint United Nations Programme on HIV/AIDS.

Wilson, Nicholas L., Wentao Xiong, and Christine L. Mattson. 2014. "Is sex like driving? HIV prevention and risk compensation." *Journal of Development Economics* 106:78–91.

Wilson N, Frade S, Rech D, et al. 2016 "Advertising for demand creation for voluntary medical male circumcision." J Acquir Immune Defic Syndr. 72(suppl 4): S293–S296.

Workowski, Kimberly A. and Gail A. Bolan. 2015. "Sexually Transmitted Diseases Treatment Guidelines, 2015." MMWR Recomm Rep 2015;64(No. RR-3):1-137, Center for Disease Control and Prevention.

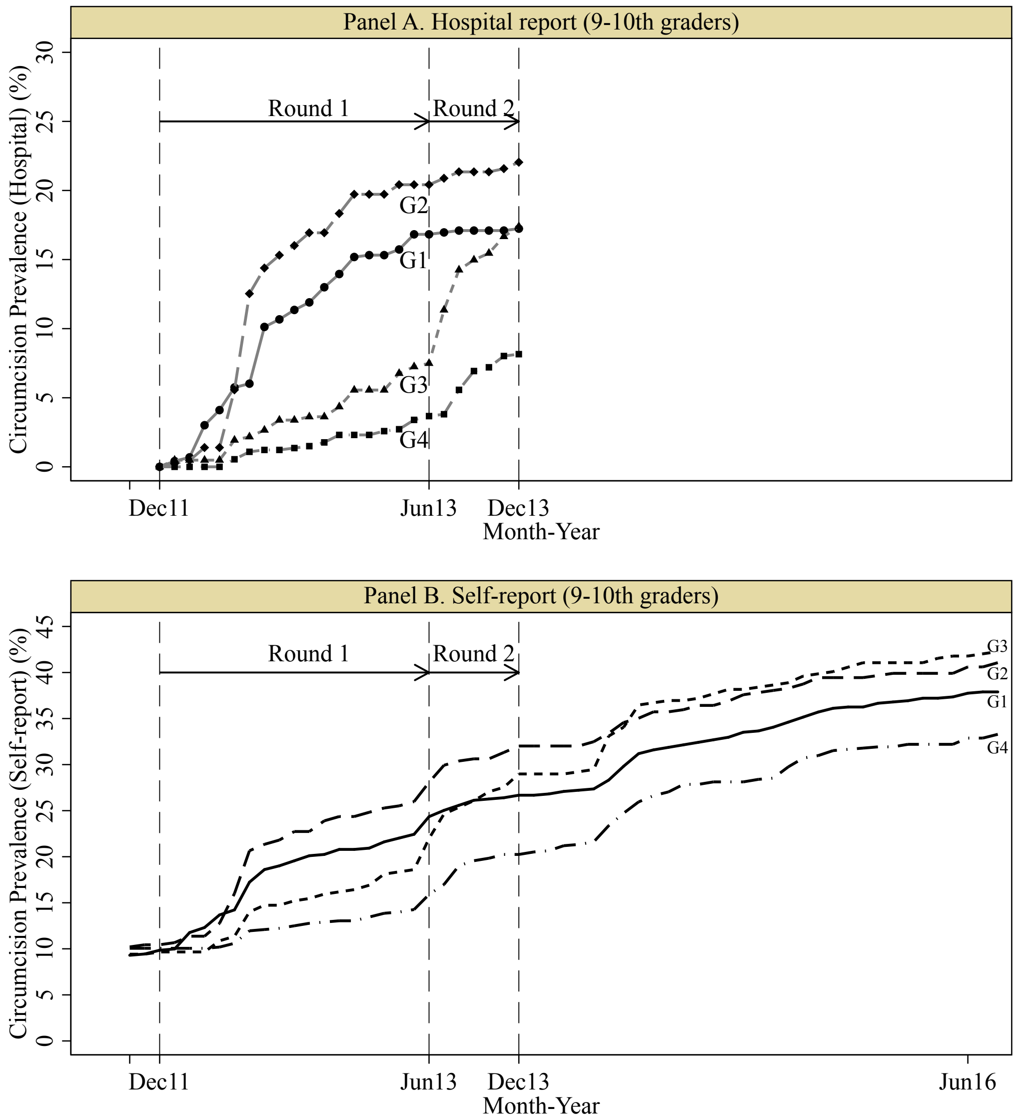
World Health Organization (WHO). 2019. "Voluntary medical male circumcision: remarkable progress in the scale up of VMMC as an HIV prevention intervention in 15 ESA countries (No. WHO/CDS/HIV/19.50)". World Health Organization.

———. 2018. *World health statistics 2018: monitoring health for the SDGs*. OCLC: 1040265127.

Zimmerman, David J. 2003. "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment." *Review of Economics and Statistics* 85 (1):9–23.

Figures and Tables

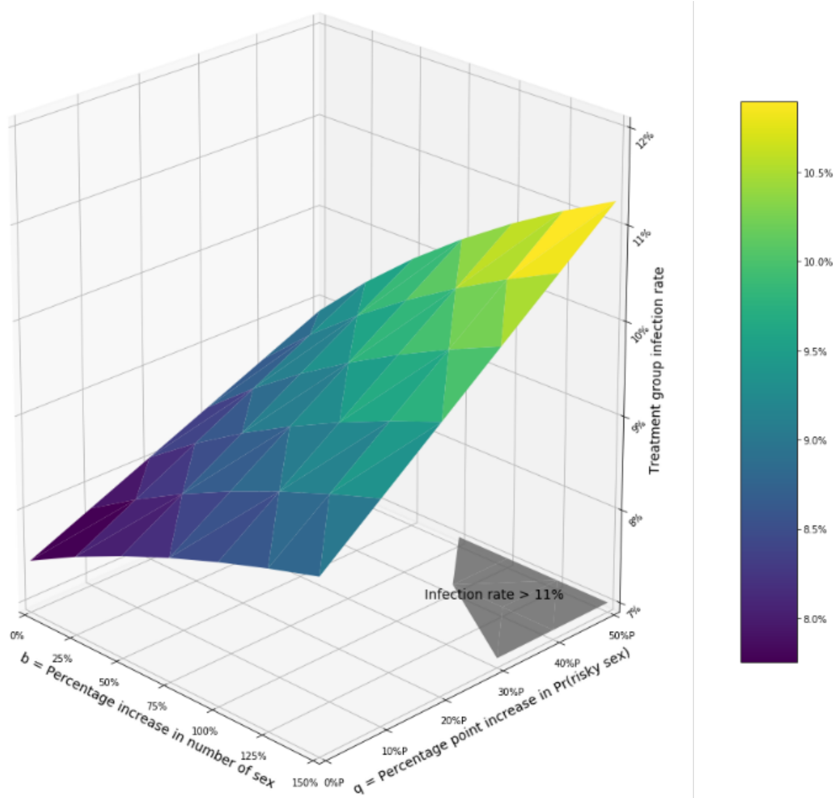
Figure 1. Cumulative Prevalence of Male Circumcision over Time



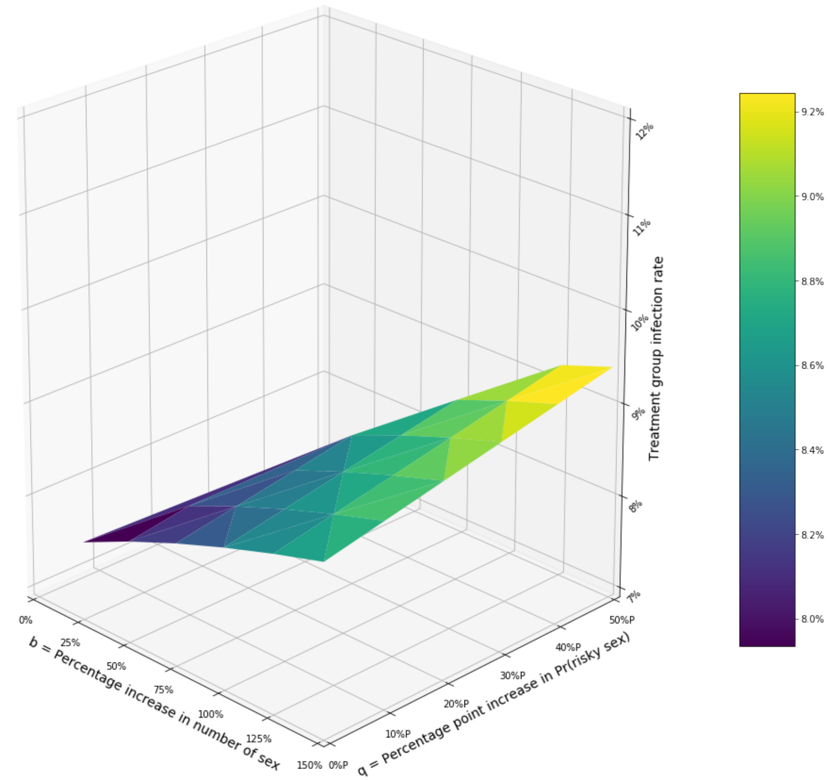
Notes: These figures present cumulative male circumcision prevalence rate over time for the sample of 2,312 9th-10th graders who completed the long-term follow-up survey. Panels A and B show the prevalence based on hospital administration and self-reported data, respectively.

Figure 2. Long-term HSV-2 Infection Rates by Risk Compensation

Panel A. Scenario 1: HSV-2 transmission rate without condom (0.0044) and with a condom (0.0002)



Panel B. Scenario 2: HSV-2 transmission rate without condom (0.0017) and with a condom (0.0006)



Notes: These figures show how the treatment group infection rate changes with the percentage increase in frequency of sex (b) and percentage point increase of probability of risky sex (q) with two different level of HSV-2 transmission rates. We assume that the proportion of those take up circumcision from the risky and safe group is 80% and 20%, respectively.

Table 1: Experimental Design

	Group	Assignment	Full Sample		Baseline 9th-10th grade	
			Classrooms	Students	Classrooms	Students
100% Treatment	G1	Treated in Round 1	41	1,293	27	861
50% Treatment	G2	Treated in Round 1	41	679	28	481
	G3	Treated in Round 2		676		470
No Treatment	G4	Treated in Round 2	42	1,322	28	851
Total			124	3,970	83	2,663

Notes: This table presents two stage randomization design. First, 124 available classrooms within the 33 schools were stratified by grade and randomly assigned into three groups: 100% treatment, 50% treatment, and No treatment. Second, within 50% treatment classrooms, half of the students were randomly assigned to treatment in Round 1 at individual level. Full sample includes all male students from 9th, 10th, and 11th grade at the baseline. The long-term follow-up survey was conducted only for 9th and 10th grade students at baseline.

Table 2: Baseline Statistics and Randomization Balance

	Mean	Difference in Mean (Full Sample)			Mean	Difference in Mean (Baseline 9th-10th graders)		
	G4	(G1vs. G4)	(G2vs. G4)	(G3 vs. G4)	G4	(G1vs. G4)	(G2vs. G4)	(G3 vs. G4)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Socio-demographic Characteristics								
Age (Year)	16.809	-0.219	-0.278	-0.229	16.163	0.070	-0.091	0.026
Circumcising ethnicity	0.158	0.033	0.020	-0.011	0.168	0.022	0.013	-0.019
Muslim	0.050	0.021*	0.021	0.000	0.054	0.017	0.019	-0.005
Orphan	0.068	-0.006	-0.025**	-0.021**	0.060	-0.001	-0.027**	-0.013
Father's tertiary education	0.172	0.007	0.020	0.014	0.166	0.020	0.034	0.020
Mother's tertiary education	0.068	-0.004	0.005	0.003	0.069	-0.009	-0.000	0.005
Father's white-collar job	0.223	0.028	0.029	0.011	0.241	0.006	0.000	-0.002
Mother's white-collar job	0.096	-0.004	0.008	0.005	0.105	-0.008	0.002	-0.003
Household asset count (0-16)	7.313	0.382	-0.192	-0.184	7.424	0.181	-0.408	-0.437
Conventional schools	0.186	-0.010	0.179*	0.169	0.095	0.070	0.337***	0.337***
Panel B. HIV/AIDS Knowledge and Sexual Behavior								
HIV/AIDS knowledge (0-20)	17.371	-0.132	-0.025	-0.052	17.398	-0.272**	-0.059	-0.049
Belief in the efficacy of MC	0.671	-0.045*	-0.001	-0.038	0.680	-0.065*	0.004	-0.025
MC is painful	0.358	0.049**	0.041	0.048*	0.345	0.069**	0.037	0.047
MC is only for Muslim	0.149	0.001	0.005	0.017	0.147	0.010	-0.003	0.012
Ever had sex	0.308	-0.012	0.010	0.002	0.254	0.024	0.038	0.038
Sexually active	0.094	-0.021	0.008	0.007	0.059	0.002	0.043*	0.037**
Multiple partners	0.012	0.004	0.004	0.003	0.011	0.003	0.002	0.002
Inconsistent use of condoms	0.045	-0.014*	-0.005	-0.002	0.041	-0.012	-0.004	0.002
Number of condoms purchased	0.919	-0.193*	0.004	0.051	0.784	-0.029	0.175	0.225*
Already circumcised	0.100	0.011	0.025	-0.014	0.106	-0.001	0.013	-0.012
Observations	1,322	2,615	2,001	1,998	851	1,712	1,332	1,321

Notes: This table reports means of selected baseline variables and shows tests for balance between treatment arms. Panel A summarizes demographic and socioeconomic information, and Panel B summarizes HIV/AIDS knowledge and individual sexual behaviors. Columns (1) and (5) show summary statistics for those initially assigned to G4. Columns (2)-(4) and (6)-(8) report mean differences (and significance levels for difference of mean tests) between groups having different treatment status. Circumcising ethnicity refers to a tribe of which more than 20% population reported being circumcised in 2010 MDHS. HIV/AIDS knowledge is constructed by counting the correct answers from 20 HIV/AIDS related questions. Inconsistent use of condoms is an indicator variable which becomes one if the respondent did not use condoms at least once during the last sexual intercourse with three most recent partners in the past 12 months. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 3: Impacts on Male Circumcision Take-up

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample	Full sample (9th-11th graders)						9th and 10th graders					
Timing	Round 1			Round 2			Round 1			Round 2		
Data	Hospital Administration Data											
Panel A												
G1 (100% Treatment)	0.133***	0.138***	0.139***	0.091***	0.096***	0.099***	0.125***	0.127***	0.141***	0.090***	0.088***	0.109***
	(0.027)	(0.026)	(0.022)	(0.029)	(0.028)	(0.024)	(0.024)	(0.026)	(0.017)	(0.026)	(0.029)	(0.020)
G2 (50% Treatment)	0.170***	0.170***	0.177***	0.137***	0.132***	0.136***	0.173***	0.170***	0.170***	0.147***	0.133***	0.121***
	(0.030)	(0.029)	(0.025)	(0.033)	(0.031)	(0.025)	(0.037)	(0.038)	(0.024)	(0.039)	(0.041)	(0.023)
G3 (50% No Treatment)	0.038*	0.040**	0.044**	0.068**	0.065**	0.066***	0.039*	0.040**	0.042*	0.097**	0.085**	0.073***
	(0.021)	(0.020)	(0.022)	(0.033)	(0.029)	(0.025)	(0.023)	(0.019)	(0.022)	(0.040)	(0.035)	(0.023)
F test (Prob >F)												
G1=G2	0.309	0.391	0.226	0.223	0.327	0.223	0.260	0.332	0.271	0.182	0.319	0.609
G1=G3	0.002	0.001	0.000	0.546	0.392	0.245	0.006	0.003	0.000	0.864	0.945	0.153
G2=G3	0.000	0.000	0.000	0.006	0.008	0.005	0.000	0.000	0.001	0.052	0.065	0.064
R-Squared	0.043	0.058	0.104	0.019	0.037	0.091	0.045	0.054	0.107	0.023	0.039	0.106
Panel B												
G1 and G2 combined	0.146***	0.149***	0.153***	0.107***	0.108***	0.113***	0.142***	0.141***	0.152***	0.110***	0.103***	0.113***
	(0.022)	(0.021)	(0.018)	(0.025)	(0.023)	(0.019)	(0.022)	(0.023)	(0.015)	(0.024)	(0.026)	(0.017)
G3	0.038*	0.039**	0.041*	0.068**	0.064**	0.063**	0.039*	0.038*	0.033	0.097**	0.082**	0.070***
	(0.021)	(0.020)	(0.022)	(0.033)	(0.029)	(0.024)	(0.023)	(0.019)	(0.025)	(0.040)	(0.033)	(0.024)
F test (Prob >F)												
G1&G2=G3	0.000	0.000	0.000	0.180	0.107	0.021	0.000	0.000	0.000	0.678	0.452	0.052
R-Squared	0.041	0.057	0.103	0.017	0.036	0.090	0.042	0.053	0.107	0.020	0.037	0.105
Observations	3,970	3,937	3,937	3,970	3,937	3,937	2,663	2,643	2,643	2,663	2,643	2,643
Mean of Dep. Variable from G4		0.048			0.093			0.035			0.075	
Controls												
Grade Fixed Effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Socio-demographic controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
HIV/AIDS education and girl's CCT program	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Notes: The circumcision status is provided by hospital administration data. Robust standard errors clustered by classroom are in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 4: Externalities on Male Circumcision Take-up (50% classrooms)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sample	Full sample (9th-11th graders)						9th and 10th graders					
Timing	Round 1			Round 2			Round 1			Round 2		
Data	Hospital Administration Data											
Panel A												
MC offer	0.132***	0.137***	0.135***	0.069***	0.073***	0.073***	0.133***	0.140***	0.137***	0.050*	0.061**	0.060**
	(0.028)	(0.029)	(0.029)	(0.025)	(0.025)	(0.025)	(0.035)	(0.037)	(0.037)	(0.026)	(0.027)	(0.025)
R-squared	0.034	0.080	0.123	0.007	0.065	0.124	0.037	0.093	0.127	0.004	0.082	0.137
Panel B												
MC offer	0.133***	0.137***	0.136***	0.069***	0.074***	0.074***	0.134***	0.140***	0.138***	0.050*	0.062**	0.061**
	(0.028)	(0.030)	(0.030)	(0.025)	(0.025)	(0.025)	(0.035)	(0.037)	(0.037)	(0.025)	(0.027)	(0.025)
Rate of close friends who got MC offer	0.044	0.035	0.043	0.070	0.049	0.060	0.063	0.050	0.055	0.085*	0.071	0.073
	(0.041)	(0.039)	(0.039)	(0.042)	(0.043)	(0.045)	(0.042)	(0.040)	(0.037)	(0.045)	(0.046)	(0.046)
R-squared	0.035	0.080	0.124	0.010	0.066	0.125	0.039	0.094	0.128	0.008	0.084	0.140
Panel C												
MC offer	0.081***	0.088***	0.090**	0.031	0.036	0.040	0.079**	0.082**	0.085**	0.008	0.010	0.014
	(0.030)	(0.032)	(0.033)	(0.032)	(0.033)	(0.033)	(0.037)	(0.038)	(0.041)	(0.033)	(0.033)	(0.032)
Rate of close friends who got MC offer	-0.027	-0.033	-0.021	0.016	-0.003	0.013	-0.010	-0.028	-0.016	0.029	0.002	0.012
	(0.034)	(0.041)	(0.044)	(0.048)	(0.055)	(0.060)	(0.034)	(0.044)	(0.042)	(0.059)	(0.062)	(0.064)
MC offer x Rate of close friends who got MC offer	0.140**	0.134*	0.126*	0.105	0.102	0.092	0.140*	0.151**	0.138*	0.109	0.134	0.120
	(0.066)	(0.066)	(0.068)	(0.086)	(0.083)	(0.082)	(0.076)	(0.071)	(0.073)	(0.104)	(0.092)	(0.090)
R-squared	0.038	0.083	0.126	0.011	0.067	0.126	0.042	0.098	0.131	0.009	0.086	0.141
Observations	1,355	1,339	1,339	1,355	1,339	1,339	951	942	942	951	942	942
Mean of Dep. Variable from G3		0.152			0.196			0.142			0.198	
Controls												
Grade Fixed Effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Socio-demographic controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
HIV/AIDS education and girl's CCT program	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Notes: This analysis includes only the 50% Treatment classroom sample. Robust standard errors are shown in parentheses clustered at the classroom level. “Rate of close friends who got MC offer” is a variable for male circumcision offer to friends defined as the proportion (rate) of friends who are offered male circumcision. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 5: Impacts on HSV2 and HIV Infections (9th and 10th Graders)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent vars.	**Primary Biomarker**			**Secondary Biomarker**						Standardized Treatment Effect		
	HSV2 IgG Positive			HSV2 IgM Positive			HIV Positive					
Sample	9th and 10th graders											
Data	Hospital Administration Data											
Panel A												
G1 (100% Treatment)	0.024	0.019	0.039***	0.012	0.013*	0.019***	0.005	0.004	-0.002	0.032*	0.029*	0.031***
	(0.021)	(0.017)	(0.012)	(0.008)	(0.008)	(0.006)	(0.004)	(0.004)	(0.003)	(0.018)	(0.015)	(0.011)
G2 (50% Treatment)	0.048*	0.042*	0.035	0.014	0.020*	0.025***	0.002	0.001	-0.002	0.038**	0.039**	0.035***
	(0.029)	(0.024)	(0.022)	(0.010)	(0.010)	(0.007)	(0.004)	(0.004)	(0.003)	(0.018)	(0.017)	(0.014)
G3 (50% No Treatment)	0.018	0.000	-0.001	-0.004	-0.000	0.005	-0.000	-0.001	-0.004	0.002	-0.003	-0.003
	(0.022)	(0.019)	(0.018)	(0.008)	(0.010)	(0.007)	(0.003)	(0.005)	(0.004)	(0.014)	(0.015)	(0.012)
Family-wise adjusted p-value												
	0.472	0.517	0.002	0.394	0.313	0.005	0.472	0.517	0.528			
	0.319	0.086	0.113	0.321	0.057	0.001	0.582	0.790	0.504			
	0.420	0.983	0.944	0.622	0.994	0.479	0.922	0.780	0.327			
F test (Prob >F)												
G1=G2	0.459	0.370	0.832	0.882	0.455	0.477	0.615	0.488	0.985	0.773	0.577	0.745
G1=G3	0.806	0.370	0.032	0.064	0.148	0.193	0.263	0.254	0.659	0.103	0.042	0.014
G2=G3	0.326	0.158	0.217	0.052	0.026	0.044	0.581	0.633	0.686	0.017	0.004	0.011
R-Squared	0.003	0.033	0.062	0.003	0.011	0.039	0.001	0.005	0.021	0.002	0.008	0.021
Panel B												
G1 and G2 combined	0.033*	0.027*	0.038***	0.013*	0.015**	0.021***	0.004	0.003	-0.002	0.034**	0.033**	0.032***
	(0.019)	(0.016)	(0.014)	(0.007)	(0.008)	(0.006)	(0.003)	(0.003)	(0.003)	(0.014)	(0.014)	(0.011)
G3	0.018	-0.001	-0.000	-0.004	-0.001	0.004	-0.000	-0.001	-0.004	0.002	-0.003	-0.004
	(0.022)	(0.019)	(0.021)	(0.008)	(0.010)	(0.007)	(0.003)	(0.005)	(0.004)	(0.014)	(0.015)	(0.013)
Family-wise adjusted p-value												
	0.082	0.091	0.008	0.084	0.047	0.000	0.267	0.325	0.439			
	0.420	0.951	0.997	0.622	0.955	0.629	0.922	0.817	0.356			
F test (Prob >F)												
G1&G2=G3	0.507	0.182	0.129	0.024	0.042	0.057	0.288	0.320	0.657	0.018	0.004	0.006
R-Squared	0.002	0.032	0.062	0.003	0.011	0.039	0.001	0.004	0.021	0.002	0.008	0.021
Observations	2,074	2,058	2,058	2,074	2,058	2,058	2,074	2,058	2,058	6,222	6,174	6,174
Mean of Dep. Variable from G4		0.084			0.012			0.003			0.000	
Controls												
Grade Fixed Effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Socio-demographic controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
HIV/AIDS educ. & girl's CCT	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Notes: This analysis includes only 9th and 10th graders who were surveyed in the long-term follow-up. Dependent variables are the probability of HSV-2 infection measured by IgG and IgM, and HIV infection. Robust standard errors are shown in parentheses clustered at the classroom level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 6: Externalities on HSV2 and HIV Infections (9th and 10th Graders in 50% Classrooms)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent vars.		HSV2 IgG Positive			HSV2 IgM Positive			HIV Positive			Standardized Treatment Effect	
Sample						9th and 10th graders						
Data						Hospital Administration Data						
Panel A												
MC offer	0.030	0.044	0.036	0.018*	0.019*	0.017*	0.003	0.002	0.001	0.039**	0.043***	0.037**
	(0.031)	(0.030)	(0.029)	(0.009)	(0.009)	(0.010)	(0.005)	(0.005)	(0.005)	(0.016)	(0.017)	(0.017)
R-squared	0.002	0.041	0.073	0.005	0.023	0.047	0.000	0.017	0.031	0.002	0.008	0.018
Panel B												
MC offer	0.030	0.044	0.036	0.018*	0.018*	0.017	0.002	0.002	0.001	0.038**	0.043**	0.037**
	(0.031)	(0.030)	(0.029)	(0.009)	(0.009)	(0.010)	(0.005)	(0.005)	(0.005)	(0.016)	(0.017)	(0.017)
Rate of close friends who got MC offer	0.006	0.014	0.009	-0.018	-0.016	-0.012	-0.003**	-0.002	-0.001	-0.026	-0.019	-0.013
	(0.037)	(0.034)	(0.036)	(0.018)	(0.016)	(0.015)	(0.001)	(0.001)	(0.001)	(0.028)	(0.026)	(0.024)
R-squared	0.002	0.041	0.073	0.006	0.024	0.048	0.001	0.017	0.031	0.002	0.008	0.018
Panel C												
MC offer	0.079	0.091*	0.079*	0.008	0.010	0.008	0.003	0.002	0.001	0.046	0.050	0.042
	(0.050)	(0.049)	(0.046)	(0.018)	(0.018)	(0.020)	(0.006)	(0.005)	(0.005)	(0.033)	(0.033)	(0.034)
Rate of close friends who got MC offer	0.071	0.077	0.066	-0.030	-0.027	-0.023	-0.002	-0.002	-0.001	-0.016	-0.009	-0.007
	(0.066)	(0.062)	(0.064)	(0.021)	(0.020)	(0.019)	(0.002)	(0.002)	(0.003)	(0.040)	(0.038)	(0.038)
MC offer x Rate of close friends who got MC offer	-0.125	-0.119	-0.110	0.024	0.021	0.022	-0.002	-0.001	0.000	-0.019	-0.020	-0.012
	(0.092)	(0.088)	(0.091)	(0.033)	(0.033)	(0.036)	(0.003)	(0.004)	(0.004)	(0.058)	(0.054)	(0.057)
Family-wise adjusted p-value (interaction terms)	0.188	0.185	0.236	0.469	0.532	0.542	0.609	0.809	0.976			
R-squared	0.005	0.044	0.075	0.007	0.024	0.048	0.001	0.017	0.031	0.002	0.008	0.018
Observations	757	750	750	757	750	750	757	750	750	2,271	2,250	2,250
Mean of Dep. Variable from G3		0.102			0.008			0.003			0.198	
Controls												
Grade Fixed Effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Socio-demographic controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
HIV/AIDS educ. and girl's CCT	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Notes: This analysis includes only the 50% Treatment classroom sample. Dependent variables are the probability of HSV-2 infection measured by IgG and IgM, and HIV infection. “Rate of close friends who got MC offer” is a variable for male circumcision offer to friends defined as the proportion (rate) of friends who are offered male circumcision. Robust standard errors are shown in parentheses clustered at the classroom level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 7: Impacts on Sexual Behaviors using Item Count Technique Questions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent vars.	Using condoms for casual sex (ICT)			Multiple partners (ICT)			Standardized Treatment Effect		
Sample	9th and 10th graders								
Data	Hospital Administration Data								
Panel A									
G1 (100% Treatment)	0.164*	0.175*	0.173*	0.064	0.077	0.045	-0.023	-0.022	-0.030
	(0.094)	(0.093)	(0.100)	(0.088)	(0.087)	(0.083)	(0.025)	(0.025)	(0.027)
G2 (50% Treatment)	0.218**	0.205**	0.255**	0.035	0.043	0.007	-0.043*	-0.038	-0.058*
	(0.083)	(0.089)	(0.098)	(0.096)	(0.098)	(0.101)	(0.025)	(0.026)	(0.031)
G3 (50% No Treatment)	0.143	0.131	0.175	0.028	0.029	0.003	-0.027	-0.024	-0.040
	(0.108)	(0.114)	(0.119)	(0.125)	(0.119)	(0.123)	(0.027)	(0.027)	(0.030)
Long	0.614***	0.623***	0.618***	0.148*	0.154*	0.150*	-0.108***	-0.108***	-0.108***
	(0.076)	(0.078)	(0.083)	(0.085)	(0.085)	(0.085)	(0.022)	(0.022)	(0.023)
G1 (100% Treatment) x Long	-0.239*	-0.256*	-0.235	-0.091	-0.105	-0.083	0.034	0.035	0.035
	(0.141)	(0.142)	(0.144)	(0.103)	(0.104)	(0.103)	(0.033)	(0.033)	(0.033)
G2 (50% Treatment) x Long	-0.231*	-0.223*	-0.232*	-0.077	-0.065	-0.062	0.035	0.036	0.039
	(0.123)	(0.126)	(0.130)	(0.140)	(0.140)	(0.139)	(0.037)	(0.036)	(0.037)
G3 (50% No Treatment) x Long	-0.107	-0.097	-0.083	-0.103	-0.086	-0.088	0.000	0.002	-0.002
	(0.150)	(0.156)	(0.157)	(0.188)	(0.187)	(0.189)	(0.039)	(0.039)	(0.040)
Family-wise adjusted p-value (interaction terms)									
	0.201	0.167	0.219	0.391	0.323	0.427			
	0.142	0.169	0.159	0.591	0.652	0.660			
	0.711	0.769	0.822	0.711	0.769	0.822			
R-Squared	0.042	0.047	0.062	0.002	0.009	0.031	0.007	0.009	0.015
Panel B									
G1 and G2 combined	0.184**	0.185**	0.201**	0.053	0.065	0.032	-0.030	-0.027	-0.039
	(0.079)	(0.081)	(0.082)	(0.078)	(0.079)	(0.078)	(0.022)	(0.023)	(0.025)
G3	0.143	0.128	0.147	0.028	0.030	0.011	-0.027	-0.023	-0.032
	(0.108)	(0.112)	(0.108)	(0.125)	(0.118)	(0.120)	(0.027)	(0.027)	(0.028)
Long	0.614***	0.622***	0.615***	0.148*	0.154*	0.151*	-0.108***	-0.108***	-0.107***
	(0.076)	(0.078)	(0.082)	(0.085)	(0.085)	(0.085)	(0.022)	(0.022)	(0.023)
MC Treatment (G1 and G2) x Long	-0.236**	-0.243**	-0.230*	-0.086	-0.091	-0.077	0.035	0.035	0.035
	(0.112)	(0.114)	(0.117)	(0.101)	(0.102)	(0.101)	(0.029)	(0.029)	(0.029)
G3 (50% No Treatment) x Long	-0.107	-0.097	-0.078	-0.103	-0.086	-0.090	0.000	0.002	-0.003
	(0.150)	(0.156)	(0.156)	(0.187)	(0.187)	(0.189)	(0.039)	(0.039)	(0.040)
Family-wise adjusted p-value (interaction terms)									
	0.088	0.082	0.114	0.411	0.384	0.453			
	0.711	0.771	0.839	0.711	0.771	0.839			
R-Squared	0.042	0.047	0.062	0.002	0.009	0.031	0.007	0.009	0.015
Observations	2,311	2,294	2,294	2,312	2,295	2,295	4,623	4,589	4,589
Controls									
Grade Fixed Effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Socio-demographic controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
HIV/AIDS education and girl's CCT program	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	Yes	No	No	Yes

Notes: This analysis includes only 9th and 10th graders who were surveyed in the long-term follow-up. Long refers to a set of questions that include an extra sensitive item. The extra item for the dependent variable "Using condoms for casual sex" is "I think I have to use a condom in case of sex with somebody that I do not know well.", and that for "Multiple partners" is "I had sex with more than two people in last 12 months." Robust standard errors are shown in parentheses clustered at the classroom level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 8: Impacts on Self-reported Sexual Behaviors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Dependent vars.	No. of condoms purchased	Ever had sex	Age at sexual debut	Sexually active	Multiple partners in past 12 mon.	Multiple partners in lifetime	Inconsistent condom use	Unprotected sex with last partner	No. of sexual partners	No. of non-marital sexual partners	Standardized Treatment Effect
Sample	9th and 10th graders										
Panel A1. 1st Follow-up (9-10th grade sample)											
G1 (100% Treatment)	-0.041	0.030	-0.184	0.001	0.024**	0.007	-0.003	-0.005	-0.016		0.004
	(0.108)	(0.040)	(0.267)	(0.027)	(0.012)	(0.032)	(0.019)	(0.019)	(0.114)		(0.004)
G2 (50% Treatment)	0.353*	0.086*	-0.796***	0.029	-0.011	0.061	-0.015	-0.012	0.218		0.005
	(0.191)	(0.049)	(0.287)	(0.038)	(0.008)	(0.043)	(0.014)	(0.014)	(0.160)		(0.005)
G3 (50% No Treatment)	-0.031	0.084*	-0.269	0.017	0.009	0.045	-0.005	-0.004	0.125		0.001
	(0.160)	(0.045)	(0.374)	(0.027)	(0.017)	(0.041)	(0.024)	(0.024)	(0.149)		(0.005)
R-Squared	0.037	0.108	0.236	0.057	0.026	0.072	0.020	0.021	0.091		0.014
Panel A2. 1st Follow-up (9-10th grade sample)											
G1 and G2 combined	0.086	0.049	-0.393	0.010	0.013*	0.025	-0.007	-0.008	0.060		0.004
	(0.117)	(0.037)	(0.256)	(0.025)	(0.007)	(0.030)	(0.016)	(0.016)	(0.112)		(0.004)
G3 (50% No Treatment)	-0.052	0.081*	-0.254	0.015	0.011	0.042	-0.005	-0.003	0.112		0.001
	(0.163)	(0.044)	(0.386)	(0.027)	(0.017)	(0.040)	(0.024)	(0.024)	(0.145)		(0.005)
R-Squared	0.032	0.107	0.225	0.056	0.018	0.070	0.020	0.021	0.088		0.014
Observations	1,836	1,828	527	1,839	1,829	1,828	1,829	1,829	1,833		15,178
Mean of Dep. Variable from G4	1.013	0.271	15.8	0.107	0.008	0.134	0.037	0.034	0.576		
Panel B1. 2nd Follow-up (9-10th grade sample)											
G1 (100% Treatment)	0.105	0.029	-0.150	0.027	-0.001	0.013	0.023	0.023	0.101	0.052	0.004
	(0.157)	(0.029)	(0.139)	(0.034)	(0.009)	(0.030)	(0.017)	(0.016)	(0.113)	(0.095)	(0.003)
G2 (50% Treatment)	0.436**	0.058*	0.012	0.057	0.026**	-0.008	0.054**	0.046*	0.300	-0.014	0.009**
	(0.198)	(0.034)	(0.152)	(0.037)	(0.011)	(0.040)	(0.025)	(0.025)	(0.200)	(0.079)	(0.004)
G3 (50% No Treatment)	0.222	0.031	0.141	0.031	0.007	0.039	0.022	0.018	0.215	0.228	0.005
	(0.217)	(0.034)	(0.149)	(0.036)	(0.011)	(0.037)	(0.016)	(0.017)	(0.141)	(0.307)	(0.004)
R-Squared	0.007	0.070	0.116	0.067	0.011	0.039	0.051	0.051	0.033	0.094	0.016
Panel B2. 2nd Follow-up (9-10th grade sample)											
G1 and G2 combined	0.218	0.039	-0.093	0.037	0.008	0.006	0.034**	0.031**	0.169	0.022	0.006*
	(0.148)	(0.027)	(0.125)	(0.031)	(0.008)	(0.029)	(0.016)	(0.015)	(0.120)	(0.077)	(0.003)
G3 (50% No Treatment)	0.201	0.029	0.132	0.029	0.006	0.040	0.020	0.017	0.202	0.227	0.005
	(0.219)	(0.034)	(0.146)	(0.035)	(0.012)	(0.036)	(0.016)	(0.016)	(0.136)	(0.302)	(0.004)
R-Squared	0.005	0.070	0.115	0.067	0.008	0.039	0.050	0.050	0.032	0.091	0.016
Observations	2,295	2,289	1,443	2,295	2,286	2,289	2,286	2,286	2,291	127	19,887
Mean of Dep. Variable (G4)	1.397	0.607	17.8	0.549	0.026	0.329	0.071	0.069	1.315	0.079	
Controls											
Grade Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
HIV/AIDS educ. & girl's CCT	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	No	No	No	No	No	No	No	No	No	No	No

Notes: In Panel A1 and A2, we ran a weighted regression because 15 percent of students in the attrition sample were randomly selected for intensive home-visit survey in the 1st follow-up. Column (1) provides the results from an experiment we conducted during survey by giving students 10 kwachas and selling two condoms at subsidized price 5 kwachas (Thornton, 2008). We omit the result for "No. of non-martial sexual partners" in Panel A because there are only 3 married students in the 1st follow-up. Robust standard errors are shown in parentheses clustered at the classroom level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 9: Compliers Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	SES Index (using PCA)	HIV/AIDS Knowledge (0-20)	Sexual Behavior Risk Index	HIV/AIDS Knowledge (0-20)	Sexual Behavior Risk Index	Risk Tolerance	Agree -ableness	Conscien- tiousness
Sample		Baseline				2nd Follow-up		
G1 and G2 combined	0.259	-0.660**	-0.030	0.068	0.030	0.032**	-0.383**	-0.333
	(0.189)	(0.291)	(0.076)	(0.276)	(0.108)	(0.014)	(0.171)	(0.218)
Constant	-0.182	20.078***	-0.980**	19.306***	-1.999***	0.601***	6.774***	8.681***
	(0.285)	(1.333)	(0.418)	(1.307)	(0.593)	(0.061)	(0.957)	(0.844)
Observations	273	270	270	240	240	240	240	240
R-Squared	0.017	0.056	0.076	0.081	0.095	0.090	0.049	0.119
Controls								
Grade Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
HIV/AIDS educ. & girl's CCT	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	No	No	No	No	No	No	No	No

Notes: This table compares compliers characteristics of students in G1 and G2 (G1&2) and those in G3 by restricting the sample to circumcision takers in G1, G2 and G3. Standard errors clustered by classroom are in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table 10: Best Linear Projector (BLP) of Conditional Average Treatment Effect

	(1)	(2)	(3)	(4)
Parameters	ATE	HET	ATE	HET
G1 and G2 combined vs. G3				
Best ML Methods	Elastic Net		Neural Network	
Male Circumcision Take-up (Round 1)				
	0.1	1.329	0.099	0.788
	(0.043,0.162)	(0.259,2.618)	(0.039,0.161)	(0.136,1.385)
	[0.001]	[0.044]	[0.002]	[0.029]
Controls				
Grade Fixed Effects	Yes	Yes	Yes	Yes
Socio-demographic controls	Yes	Yes	Yes	Yes
HIV/AIDS education and girl's CCT program	Yes	Yes	Yes	Yes
School Fixed Effects	No	No	No	No

Notes: This table presents results of the BLP of conditional average treatment effect using the machine learning proxies for the MC take-up. 90% confidence intervals using medians over 100 splits and P-values are presented in parenthesis and brackets, respectively.

Table 11: GATES and CLAN of Male Circumcision Offer

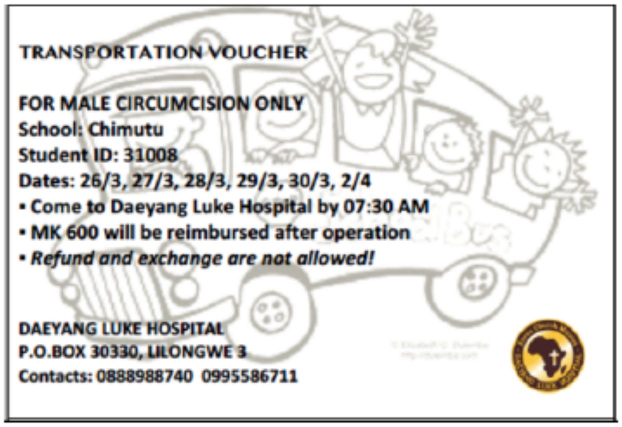
	(1)	(2)	(3)	(4)	(5)	(6)
Parameters	20% Most	20% Least	Difference	20% Most	20% Least	Difference
Panel A: GATES of MC Offer for G1 and G2 combined vs. G3						
Best ML Methods		Elastic Net			Neural Network	
Outcome var.: Male Circumcision Take-up (Round 1)						
	0.178	0.005	0.178	0.182	-0.001	0.176
	(0.081,0.282)	(-0.123,0.128)	(0.016,0.340)	(0.085,0.272)	(-0.128,0.123)	(0.019,0.335)
	[0.001]	[1.000]	[0.062]	[0.001]	[1.000]	[0.054]
Panel B: CLAN of MC Offer for G1 and G2 combined vs. G3						
Best ML Methods		Elastic Net			Neural Network	
Outcome var.: Male Circumcision Take-up (Round 1)						
SES Index	0.62	-0.55	1.215	0.736	-0.627	1.308
	(0.366,0.866)	(-0.775,-0.322)	(0.854,1.593)	(0.503,0.952)	(-0.858,-0.402)	(0.970,1.654)
	-	-	[0.000]	-	-	[0.000]
Risk tolerance	0.641	0.606	0.036	0.652	0.595	0.056
	(0.627,0.655)	(0.592,0.621)	(0.015,0.058)	(0.638,0.667)	(0.581,0.609)	(0.035,0.076)
	-	-	[0.003]	-	-	[0.000]
Agreeableness	6.011	6.422	-0.356	5.981	6.535	-0.585
	(5.844,6.212)	(6.260,6.591)	(-0.641,-0.088)	(5.820,6.146)	(6.363,6.707)	(-0.818,-0.342)
	-	-	[0.017]	-	-	[0.000]
Conscientiousness	5.712	6.36	-0.619	5.804	6.354	-0.538
	(5.527,5.942)	(6.161,6.558)	(-0.894,-0.332)	(5.616,6.000)	(6.173,6.554)	(-0.815,-0.261)
	-	-	[0.000]	-	-	[0.000]
Controls						
Grade Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Socio-demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
HIV/AIDS education and girl's CCT program	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	No	No	No	No	No	No

Notes: Panel A reports the analysis of the GATES (Sorted Group Average Treatment Effects) by comparing the most and the least affected groups. Panel B presents the CLAN (classification analysis or the average characteristics of the most and least affected units defined in terms of the ML proxy predictor) by comparing average characteristics of the two groups to understand what generates heterogeneity in the treatment effects. We focus on four characteristics in this analysis: SES index measured at the baseline, risk tolerance, agreeableness and conscientiousness measured at the long-term follow-up survey. 90% confidence intervals using medians over 100 splits and P-values are presented in parenthesis and brackets, respectively.

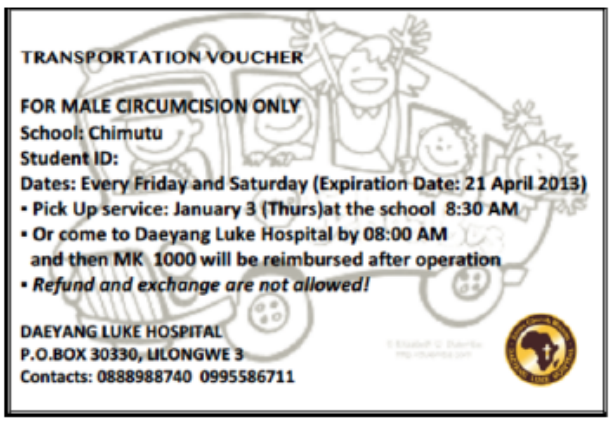
Appendix A: Additional Figures and Tables

Figure A1: Transportation Vouchers

A. 1st Transportation Voucher for Round 1



B. 2nd Transportation Voucher for Round 1



C. Transportation Voucher for Round 2

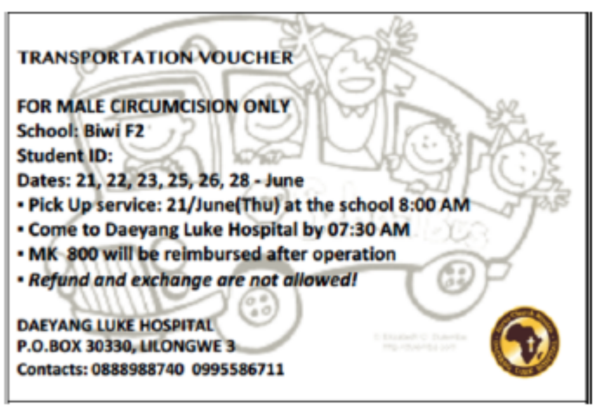


Figure A2: Male Circumcision Brochure

Notes: This brochure was translated in Chichewa and distributed to students for counseling sessions.

Figure A3: Item Count Technique Questions

A. Long-form with a sensitive item

Q231.How many of the statements below apply to you?

(A) I like to play with animals. (B) I have moved house in last three years. (C) I would consider myself a sports fan. (D) I know the full name of the president. (E) I think I have to use condom in case of a sex with somebody that you do not know well.
0 1 2 3 4 5

Q232.How many of the statements below apply to you?

(A) I sleep less than 7 hours per night (B) I have few photos of myself (C) I go to the church or mosque almost every week. (D) I have never been in a traffic accident. (E) I had sex with more than two people in last 12 months.
0 1 2 3 4 5

B. Short-form without a sensitive item

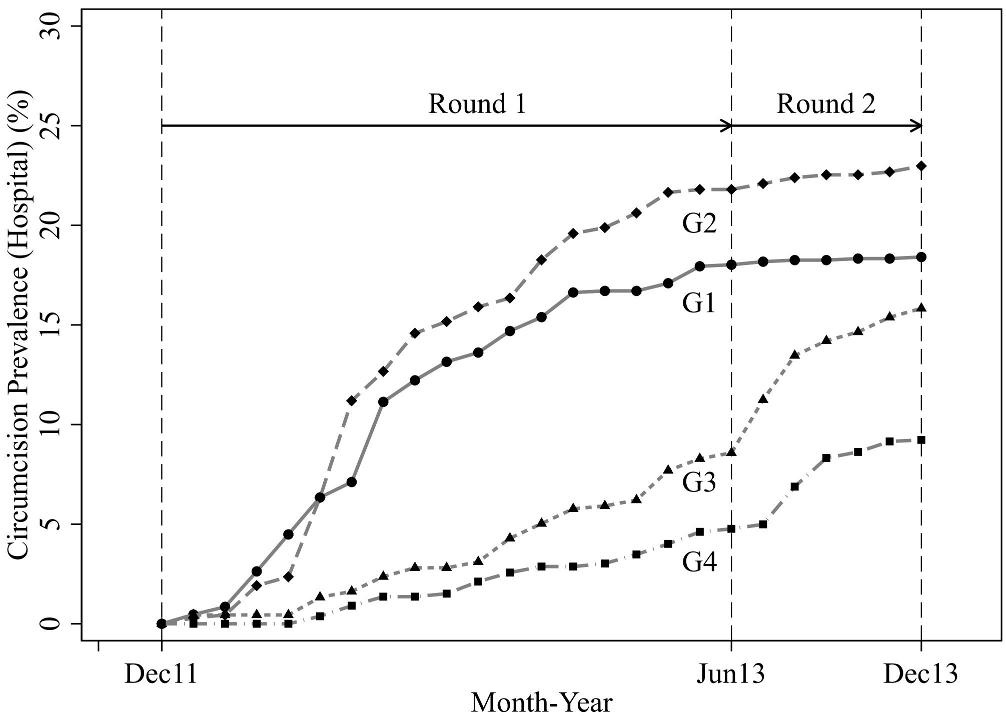
Q231.How many of the statements below apply to you?

(A) I like to play with animals. (B) I have moved house in last three years. (C) I would consider myself a sports fan. (D) I know the full name of the president.
0 1 2 3 4

Q232.How many of the statements below apply to you?

(A) I sleep less than 7 hours per night (B) I have few photos of myself (C) I go to the church or mosque almost every week. (D) I have never been in a traffic accident.
0 1 2 3 4

Figure A4. Cumulative Prevalence of Male Circumcision over Time (9th-11th Graders)

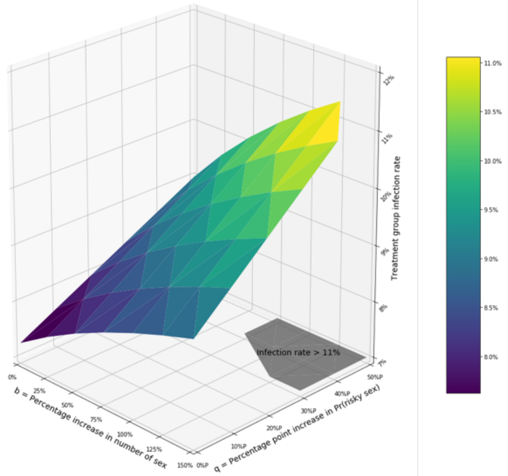


Notes: This figure presents cumulative male circumcision prevalence rate over time for the full sample of 9th-11th graders using hospital administration data.

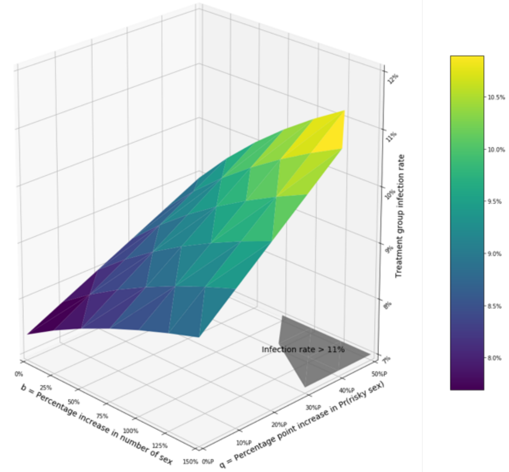
Figure A5: HSV-2 Infection Rate by Risk Compensation and Selection

HSV-2 transmission rate without condom (0.0044) and with a condom (0.0002)

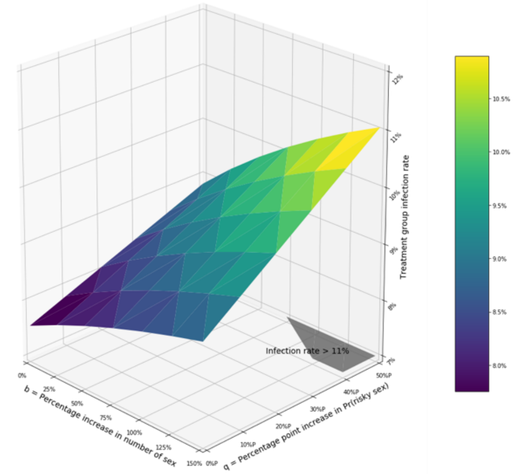
Panel A. Risky group 100%, safe group 0%



Panel B. Risky group 90%, safe group 10%

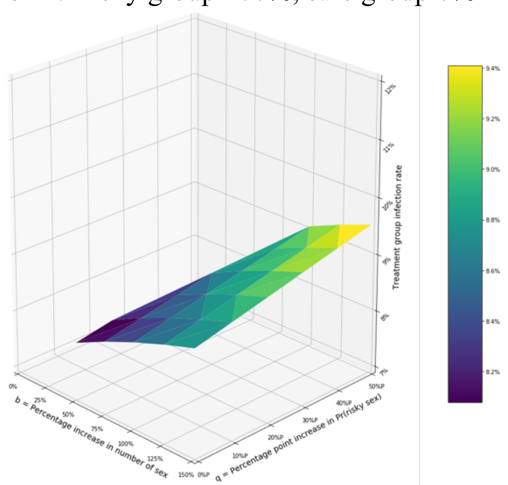


Panel C. Risky group 70%, safe group 30%

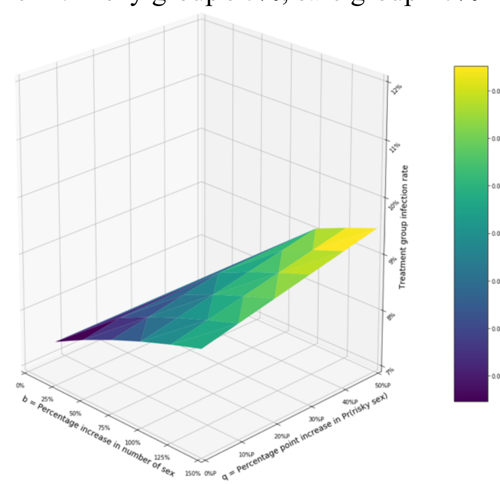


HSV-2 transmission rate without condom (0.0017) and with a condom (0.0006)

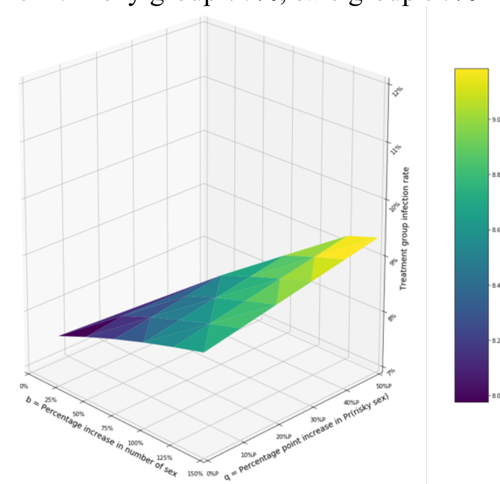
Panel D. Risky group 100%, safe group 0%



Panel E. Risky group 90%, safe group 10%



Panel F. Risky group 70%, safe group 30%



Notes: These figures shows how the treatment group infection rate changes with the percentage increase in frequency of sex (b) and percentage point increase of probability of risky sex (q) with two different level of HSV-2 transmission rates.

Table A1: In-class Friendship Networks: Summary Statistics

Panel A: Friendship Reconstruction			
	Raw count	Eligible male	Reranked eligible male
	(1)	(2)	(3)
First-best friend	3,844	3,102	3,832
Second-best friend	3,839	2,818	3,135
Third-best friend	3,860	2,668	1,621
Panel B: Friendship link treatment status			
	Case	Percentage	
No friend treated	1,697	42.75%	
One friend treated	833	20.98%	
Two friends treated	823	20.73%	
Three friends treated	617	15.54%	

Notes: Panel A Column (1) includes raw friendship data including friends without baseline survey and female friends. Column (2) excludes friends without baseline survey and further excludes female friends from Column (1). Finally, after we exclude friends without baseline survey and female friends, we re-ranked the remaining friendship data from Column (2) as first, second, or third best male friends if available. Panel B present friendship statistics based on re-ranked eligible male best friend data (Panel A, Column (3)).

Table A2: Baseline Statistics and Randomization Balance by Fraction of Treated Friends

	Mean	Difference in Mean	Mean	Difference in Mean
	Full Sample		Baseline 9th-10th grade	
	No friend treated	1 vs. 0 friend Treated	No friend treated	1 vs. 0 friend Treated
	(1)	(2)	(3)	(4)
Panel A. Socio-demographic Characteristics				
Age (Year)	16.760	-0.114	16.151	0.005
Circumcising ethnicity	0.165	0.002	0.172	0.012
Muslim	0.056	0.005	0.058	0.005
Orphan	0.065	-0.014	0.060	-0.024**
Father's tertiary education	0.176	0.004	0.170	0.023
Mother's tertiary education	0.070	0.012	0.070	0.011
Father's white-collar job	0.226	0.022	0.237	0.017
Mother's white-collar job	0.098	0.015	0.104	0.014
Household asset count (0-16)	7.338	0.064	7.361	-0.072
Conventional schools	0.207	0.118*	0.158	0.214***
Panel B. HIV/AIDS Knowledge and Sexual Behavior				
HIV/AIDS knowledge (0-20)	17.354	0.010	17.375	-0.022
Belief in the efficacy of MC	0.674	-0.039	0.686	-0.034
MC is painful	0.364	0.025	0.349	0.027
MC is only for Muslim	0.152	0.007	0.148	0.012
Ever had sex	0.313	-0.003	0.267	0.008
Sexually active	0.100	-0.020	0.071	-0.003
Multiple partners	0.015	-0.001	0.012	-0.003
Inconsistent use of condoms	0.047	-0.008	0.043	-0.010
Number of condoms purchased	0.916	-0.013	0.813	0.131
Already circumcised	0.108	-0.010	0.106	-0.009
Observations	1,697	2,530	1,098	1,655

Notes: This table reports means of selected baseline variables and mean differences (and significance levels for difference of mean tests) between groups having friendship link treatment status as presented in Table A1 Panel B. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level

Table A3: Attrition

Dependent vars.	Surveyed in short-term follow-up		Surveyed in long-term follow-up		Biomarker Testing (HIV and HSV2)	
Sample			9th and 10th graders			
	(1)	(2)	(3)	(4)	(5)	(6)
100% Treatment (G1)	-0.056*	-0.030	-0.022	-0.014	-0.024	-0.003
	(0.031)	(0.024)	(0.023)	(0.019)	(0.025)	(0.024)
50% Treatment (G2)	-0.104**	-0.057**	0.013	0.019	0.009	0.026
	(0.043)	(0.027)	(0.022)	(0.023)	(0.030)	(0.029)
50% No Treatment (G3)	-0.057	-0.013	-0.003	0.004	-0.004	0.014
	(0.038)	(0.026)	(0.021)	(0.023)	(0.027)	(0.025)
F test (Prob >F)						
G1=G2	0.301	0.391	0.136	0.136	0.282	0.292
G1=G3	0.970	0.565	0.374	0.457	0.487	0.503
G2=G3	0.100	0.115	0.435	0.461	0.638	0.689
R-squared	0.021	0.091	0.015	0.035	0.019	0.040
Observations	2,643	2,643	2,643	2,643	2,643	2,643
Mean of Dep. Variable from G4	0.940		0.865		0.779	
Controls						
Grade Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Socio-demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
HIV/AIDS education and girl's CCT program	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	No	Yes	No	Yes	No	Yes

Notes: This table is to test for systemic attrition by regressing a dummy of the survey or testing completion on a set of indicators for treatment arms. Two mean dependent variables from G4 in Columns (1)-(4) refer to the effective survey rate. The effective survey rate (ESR) a function of the regular school follow-up rate (RFR) and intensive home visit follow-up rate (HFR) after we random selected 15% students from the attrition sample (Baird et al. 2012): ESR = RFR + (1-RFR) × HFR. The controls include grade fixed effects, socio-demographic characteristics and AFF's HIV/AIDS education and girls' CCT program interventions. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table A4: Impacts on Self-reported Male Circumcision

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	9th and 10th graders					
Data	Hospital Administration Data			Self-report		
Timing	Round 2			Long-term Follow-up Survey		
Panel A						
G1 (100% Treatment)	0.090***	0.088***	0.109***	0.057	0.044	0.119**
	(0.026)	(0.029)	(0.020)	(0.050)	(0.036)	(0.053)
G2 (50% Treatment)	0.147***	0.133***	0.121***	0.280	0.313	0.348*
	(0.039)	(0.041)	(0.023)	(0.190)	(0.215)	(0.201)
G3 (50% No Treatment)	0.097**	0.085**	0.073***	0.085*	0.124*	0.141
	(0.040)	(0.035)	(0.023)	(0.051)	(0.064)	(0.107)
F test (Prob >F)						
G1=G2	0.182	0.319	0.609	0.249	0.219	0.179
G1=G3	0.864	0.945	0.153	0.618	0.189	0.792
G2=G3	0.052	0.065	0.064	0.289	0.307	0.293
R-Squared	0.023	0.039	0.106	0.003	0.024	0.071
Panel B						
G1 and G2 combined	0.110***	0.103***	0.113***	0.140*	0.136*	0.203**
	(0.024)	(0.026)	(0.017)	(0.081)	(0.078)	(0.100)
G3	0.097**	0.082**	0.070***	0.085*	0.106*	0.073
	(0.040)	(0.033)	(0.024)	(0.051)	(0.057)	(0.109)
F test (Prob >F)						
G1&G2=G3	0.678	0.452	0.052	0.483	0.661	0.372
R-Squared	0.020	0.037	0.105	0.001	0.021	0.070
Observations	2,663	2,643	2,643	2,312	2,295	2,295
Mean of Dep. Variable from G4		0.075			0.330	
Controls						
Grade Fixed Effects	No	Yes	Yes	No	Yes	Yes
Socio-demographic controls	No	Yes	Yes	No	Yes	Yes
HIV/AIDS education and girl's CCT program	No	Yes	Yes	No	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	Yes

Notes: The circumcision status is provided by hospital administration data and self-reports from the long-term follow-up survey. Columns (1)-(3) repeat Columns (10)-(12) from Table 3 to compare circumcision rates in the long run. Robust standard errors clustered by classroom are in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table A5: Heterogeneous Effects by Prior Knowledge

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Full sample (9th-11th graders)			9th and 10th graders		
	G1 vs. G4	G2 vs. G4	G3 vs. G4	G1 vs. G4	G2 vs. G4	G3 vs. G4
Timing	Round 1					
Data	Hospital Administration Data					
Male circumcision offer	0.153***	0.208***	0.043	0.157***	0.209***	0.051
	(0.031)	(0.040)	(0.028)	(0.032)	(0.055)	(0.033)
Knowledge of benefits of circumcision for HIV	0.011	0.013	0.010	0.015	0.016	0.014
	(0.013)	(0.012)	(0.012)	(0.015)	(0.013)	(0.013)
Male circumcision offer	0.002	-0.029	-0.014	0.002	-0.031	-0.033
x Knowledge of benefits of circumcision for HIV	(0.025)	(0.034)	(0.027)	(0.030)	(0.044)	(0.028)
Circumcision is painful	-0.014	-0.016	-0.013	-0.011	-0.015	-0.011
	(0.012)	(0.012)	(0.011)	(0.013)	(0.013)	(0.012)
Male circumcision offer	-0.030	-0.057	-0.011	-0.047	-0.053	0.003
x Circumcision is painful	(0.025)	(0.036)	(0.022)	(0.028)	(0.035)	(0.027)
Circumcision is only for Muslim	-0.033**	-0.025*	-0.026*	-0.015	-0.006	-0.001
	(0.015)	(0.014)	(0.015)	(0.019)	(0.017)	(0.019)
Male circumcision offer	-0.013	0.008	0.051	-0.034	-0.048	0.035
x Circumcision is only for Muslim	(0.030)	(0.055)	(0.037)	(0.032)	(0.047)	(0.049)
R-Squared	0.070	0.105	0.033	0.075	0.113	0.030
Observations	2,591	1,975	1,977	1,697	1,315	1,305
Mean of Dep. Variable from G4		0.048			0.035	

Notes: This table shows the heterogeneous effects on take-up of male circumcision. Male circumcision offer variable equals 1 when students get male circumcision offer either from 100% Treatment classrooms or from 50% Treatment classrooms. These regression results use grade fixed effects, socio-demographic controls and controls for AFF's HIV/AIDS education and girls' CCT programs. Robust standard errors clustered by classroom are in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table A6: Impacts of HIV/AIDS Education on Male Circumcision Take-up

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Sample	Full sample (9th-11th graders)								9th and 10th graders							
Timing	Round 1				Round 2				Round 1				Round 2			
Data	Hospital Administration Data															
Panel A																
G1 (100% Treatment)	0.133***	0.138***	0.131***	0.138***	0.091***	0.096***	0.084**	0.091**	0.130***	0.127***	0.114***	0.113***	0.094***	0.088***	0.075**	0.072*
	(0.026)	(0.026)	(0.035)	(0.035)	(0.028)	(0.028)	(0.038)	(0.039)	(0.024)	(0.026)	(0.028)	(0.030)	(0.027)	(0.029)	(0.032)	(0.036)
G2 (50% Treatment)	0.171***	0.170***	0.169***	0.169***	0.137***	0.132***	0.126***	0.125***	0.175***	0.170***	0.161***	0.165***	0.150***	0.133***	0.124**	0.127**
	(0.030)	(0.029)	(0.040)	(0.038)	(0.032)	(0.031)	(0.043)	(0.040)	(0.036)	(0.038)	(0.048)	(0.049)	(0.038)	(0.041)	(0.050)	(0.051)
G3 (50% No Treatment)	0.039*	0.040**	0.061**	0.063**	0.068**	0.065**	0.086**	0.082**	0.045**	0.040**	0.051*	0.052**	0.103**	0.085**	0.108**	0.099**
	(0.021)	(0.020)	(0.028)	(0.027)	(0.033)	(0.029)	(0.043)	(0.039)	(0.022)	(0.019)	(0.027)	(0.025)	(0.039)	(0.035)	(0.050)	(0.043)
HIV/AIDS Education	0.005	0.005	0.011	0.013	0.000	-0.002	-0.001	-0.001	0.025	0.021	0.012	0.015	0.022	0.014	0.004	0.007
	(0.017)	(0.017)	(0.020)	(0.020)	(0.019)	(0.019)	(0.030)	(0.027)	(0.018)	(0.019)	(0.018)	(0.020)	(0.021)	(0.022)	(0.027)	(0.028)
G1 (100% Treatment) x HIV/AIDS Education			0.008	0.002			0.016	0.011			0.036	0.034			0.041	0.039
			(0.043)	(0.044)			(0.048)	(0.050)			(0.043)	(0.046)			(0.048)	(0.053)
G2 (50% Treatment) x HIV/AIDS Education			0.004	0.003			0.023	0.016			0.028	0.010			0.050	0.013
			(0.043)	(0.040)			(0.050)	(0.045)			(0.044)	(0.045)			(0.051)	(0.048)
G3 (50% No Treatment) x HIV/AIDS Education			-0.055	-0.057*			-0.047	-0.042			-0.027	-0.037			-0.030	-0.046
			(0.034)	(0.034)			(0.048)	(0.045)			(0.037)	(0.039)			(0.053)	(0.051)
R-Squared	0.043	0.058	0.044	0.059	0.019	0.037	0.020	0.038	0.046	0.054	0.048	0.056	0.024	0.039	0.025	0.041
Panel B																
G1 and G2 combined	0.146***	0.149***	0.143***	0.147***	0.107***	0.108***	0.097***	0.101***	0.147***	0.141***	0.129***	0.129***	0.114***	0.103***	0.091***	0.089***
	(0.021)	(0.021)	(0.029)	(0.029)	(0.024)	(0.023)	(0.033)	(0.032)	(0.021)	(0.023)	(0.026)	(0.028)	(0.024)	(0.026)	(0.031)	(0.032)
G3	0.039*	0.039**	0.061**	0.062**	0.069**	0.064**	0.086**	0.080**	0.046**	0.038*	0.051*	0.048**	0.103**	0.082**	0.108**	0.096**
	(0.021)	(0.020)	(0.028)	(0.027)	(0.033)	(0.029)	(0.042)	(0.038)	(0.023)	(0.019)	(0.027)	(0.024)	(0.040)	(0.033)	(0.050)	(0.042)
HIV/AIDS Education	0.006	0.006	0.011	0.013	0.002	-0.000	-0.001	-0.001	0.027	0.023	0.012	0.015	0.024	0.016	0.004	0.008
	(0.017)	(0.016)	(0.020)	(0.020)	(0.019)	(0.019)	(0.030)	(0.027)	(0.017)	(0.018)	(0.018)	(0.020)	(0.020)	(0.022)	(0.027)	(0.028)
G1 and G2 combined x HIV/AIDS Education			0.009	0.005			0.022	0.015			0.037	0.027			0.049	0.032
			(0.035)	(0.034)			(0.041)	(0.040)			(0.033)	(0.036)			(0.040)	(0.043)
G3 x HIV/AIDS Education			-0.055	-0.056			-0.047	-0.041			-0.027	-0.036			-0.030	-0.045
			(0.034)	(0.034)			(0.048)	(0.045)			(0.037)	(0.039)			(0.053)	(0.051)
R-Squared	0.041	0.057	0.043	0.058	0.017	0.036	0.018	0.037	0.044	0.053	0.046	0.054	0.021	0.037	0.023	0.039
Observations	3,970	3,937	3,970	3,937	3,970	3,937	3,970	3,937	2,663	2,643	2,663	2,643	2,663	2,643	2,663	2,643
Mean of Dep. Variable from Non-assignment	0.042				0.094				0.028				0.073			
Controls																
Grade Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Socio-demographic controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Girl's CCT program	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes: The circumcision status is provided by hospital administration data. Robust standard errors clustered by classroom are in parentheses. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table A7: Impacts on HSV2 Infections (9th and 10th Graders)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent vars.	HSV2 IgG Positive & IgM Negative			HSV2 IgG Negative & IgM Positive			HSV2 IgG Positive & IgM Positive		
Sample	9th and 10th graders								
Data	Hospital Administration Data								
Panel A									
G1 (100% Treatment)	0.012	0.006	0.019*	0.000	-0.000	-0.002	0.012*	0.013**	0.020***
	(0.018)	(0.015)	(0.011)	(0.005)	(0.005)	(0.004)	(0.007)	(0.006)	(0.007)
G2 (50% Treatment)	0.032	0.021	0.010	-0.002	-0.002	-0.000	0.016**	0.022***	0.025***
	(0.025)	(0.022)	(0.021)	(0.005)	(0.006)	(0.003)	(0.008)	(0.008)	(0.007)
G3 (50% No Treatment)	0.017	-0.004	-0.010	-0.005	-0.005	-0.004	0.001	0.005	0.009
	(0.022)	(0.017)	(0.017)	(0.004)	(0.005)	(0.003)	(0.007)	(0.008)	(0.008)
F test (Prob >F)									
G1=G2	0.456	0.523	0.640	0.583	0.685	0.731	0.694	0.270	0.586
G1=G3	0.824	0.593	0.095	0.071	0.085	0.541	0.158	0.314	0.322
G2=G3	0.597	0.359	0.452	0.321	0.242	0.312	0.088	0.053	0.069
R-Squared	0.002	0.027	0.052	0.001	0.005	0.057	0.003	0.012	0.026
Panel B									
G1 and G2 combined	0.019	0.011	0.015	-0.001	-0.001	-0.001	0.014**	0.016***	0.022***
	(0.016)	(0.014)	(0.013)	(0.005)	(0.005)	(0.003)	(0.006)	(0.006)	(0.006)
G3	0.017	-0.005	-0.008	-0.005	-0.005	-0.004	0.001	0.004	0.008
	(0.022)	(0.018)	(0.018)	(0.004)	(0.005)	(0.003)	(0.007)	(0.008)	(0.007)
F test (Prob >F)									
G1&G2=G3	0.922	0.406	0.305	0.041	0.032	0.310	0.079	0.124	0.106
R-Squared	0.001	0.027	0.052	0.001	0.005	0.057	0.003	0.012	0.026
Observations	2,074	2,058	2,058	2,074	2,058	2,058	2,074	2,058	2,058
Mean of Dep. Variable from G4		0.077			0.005			0.008	
Controls									
Grade Fixed Effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Socio-demographic controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
HIV/AIDS education and girl's CCT program	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	Yes	No	No	Yes

Notes: This analysis includes only 9th and 10th graders who were surveyed in the long-term follow-up. Dependent variables are the probability of HSV-2 infection measured by IgG and IgM. Robust standard errors are shown in parentheses clustered at the classroom level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table A8: Comparison of ML Methods

	(1)	(2)	(3)	(4)
ML Methods	Elastic Net	Boosting	Neural Network	Random Forest
G1 and G2 combined vs. G3				
Outcome var.: Male Circumcision Take-up (Round 1)				
Best BLP	0.017	0.013	0.015	0.013
Best GATES	0.061	0.031	0.063	0.039

Notes: Following Chernozhukov et al. (2018), this shows that the Elastic Net and Neural Network outperform the Boosting and Random Forest for our main outcome, male circumcision take-up. These ML results use grade fixed effects, socio-demographic controls and controls for AFF's HIV/AIDS education and girls' CCT programs.

Table A9: Impacts on HSV2 and HIV Infections (9th and 10th Graders who were not Circumcised at Baseline)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent vars.	**Primary Biomarker**			**Secondary Biomarker**					
	HSV2 IgG Positive			HSV2 IgM Positive			HIV Positive		
Sample	9th and 10th graders								
Data	Hospital Administration Data								
Panel A									
G1 (100% Treatment)	0.028	0.022	0.044***	0.013	0.014*	0.018**	0.005	0.005	-0.002
	(0.022)	(0.018)	(0.013)	(0.009)	(0.009)	(0.007)	(0.005)	(0.004)	(0.003)
G2 (50% Treatment)	0.057*	0.048*	0.037	0.017	0.024**	0.028***	-0.000	-0.001	-0.003
	(0.029)	(0.025)	(0.023)	(0.010)	(0.011)	(0.008)	(0.004)	(0.004)	(0.003)
G3 (50% No Treatment)	0.022	0.004	-0.002	-0.003	0.003	0.006	-0.000	-0.000	-0.002
	(0.022)	(0.020)	(0.020)	(0.009)	(0.011)	(0.008)	(0.004)	(0.005)	(0.004)
R-Squared	0.004	0.032	0.060	0.004	0.012	0.044	0.001	0.006	0.023
Panel B									
G1 and G2 combined	0.039**	0.030*	0.042***	0.015*	0.017**	0.021***	0.003	0.003	-0.002
	(0.019)	(0.016)	(0.014)	(0.008)	(0.008)	(0.006)	(0.004)	(0.004)	(0.003)
G3	0.022	0.002	-0.000	-0.003	0.002	0.003	-0.000	0.000	-0.002
	(0.022)	(0.020)	(0.022)	(0.009)	(0.011)	(0.008)	(0.004)	(0.005)	(0.004)
R-Squared	0.003	0.031	0.060	0.003	0.012	0.044	0.001	0.005	0.023
Observations	1,881	1,866	1,866	1,881	1,866	1,866	1,881	1,866	1,866
Mean of Dep. Variable from G4		0.087			0.012			0.003	
Controls									
Grade Fixed Effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Socio-demographic controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
HIV/AIDS educ. & girl's CCT	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	Yes	No	No	Yes

Notes: This analysis includes only 9th and 10th graders who were not circumcised at baseline and surveyed in the long-term follow-up. Dependent variables are the probability of HSV-2 infection measured by IgG and IgM, and HIV infection. Robust standard errors are shown in parentheses clustered at the classroom level. *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Table A10: Lee's bounds for Impacts on HSV2 and HIV infections (9th and 10th Grade)

	(1)	(2)	(3)
Dependent vars.	**Primary Biomarker**	**Secondary Biomarker**	
	HSV2 IgG Positive	HSV2 IgM Positive	HIV Positive
Sample	9th and 10th graders		
Data	Hospital Administration Data		
G1 and G2 combined	0.0330*	0.0130*	0.0037
	(0.0187)	(0.0074)	(0.0033)
R-Squared	0.0028	0.0020	0.0006
Lee lower bound	0.0323**	0.0129**	0.0037
	(0.0148)	(0.0065)	(0.0033)
Lee upper bound	0.0398*	0.0203	0.0067***
	(0.0208)	(0.0171)	(0.0025)
Observations	1,702	1,702	1,702
Mean of Dep. Variable from G4	0.0845	0.0121	0.0030

Notes: Following Lee (2009), this analysis provides confidence intervals for the true value of the treatment effect (G1 and G2 combined vs. G4) in the presence of nonrandom sample attrition in biomarker testing. Dependent variables are the probability of HSV-2 infection measured by IgG and IgM, and HIV infection.

Appendix B: Simulation of STD Infection with Risk Compensation

In this section we explain in greater detail the simulation exercise summarized in Section 5.3.3 of the paper.

Section B1: How risk compensation can affect the HSV infection rate

We define the HSV Infection rate as follows:

$$Infection\ rat = Frequency\ of\ sexual\ activity * \text{Probability}(risky\ sex) * Transmission\ rate$$

The infection rate is the product of the frequency of sexual activity during the study period, the probability of having sexual intercourse without protection and the HSV transmission rate. Male circumcision is known to decrease the HSV transmission rate by half and thus lowers the overall HSV infection level (Sobngwi-Tambekou et al., 2009). However, there are behavioral adjustments that might be associated with male circumcision. First, male circumcision could *ex ante* appeal to individuals with an inclination towards risky sex. This implies that there is a selection effect, namely, those who are more likely to engage in risky sex have a higher tendency to decide to get the circumcision procedure. Second, male circumcision induces *ex post* risk compensating behavior. With a lower probability of infection, the circumcised could be more likely to engage in risky sexual behaviors or they could increase the frequency of sex, or both.

The aim of this simulation exercise is to examine the effect of male circumcision on the infection rate taking into account selection effects and risk compensation. The latter will be modeled in two dimensions, i.e., as an increase in the frequency of sexual activity and an increase in the probability of having risky sex.

Section B2: Model

Because we have an RCT in our setting, the frequency of sexual activity (*n)* and the probability of unprotected risky sex (*p)* is the same between the treatment and control group at baseline. We assume that both groups consist of a mix of risky and safe individuals that differ in the frequency of sexual activity and the probability of risky sex. The share of each group is shown in Table B1. The frequency of sexual activity for the risky group is larger by *a* compared to the group mean, n, and it is smaller by a in the safe group. Similarly, the probability of risky sex is larger by r in the risky group compared to the group mean, while it is smaller by r in the safe group. Hence, regardless of the value of a and r, the average frequency of sexual activity and the probability of risky sex of both groups stay constant at n and p. We also assume that the risky group is more likely to take-up male circumcision than the safe group. We will present four scenarios, where the proportion of individuals who take-up circumcision from the risky and safe groups is (100%, 0%), (90%, 10%), (80%, 20%), and (70%, 30%), but (80%, 20%) will be our preferred scenario.

Assume that after male circumcision, a subset of the risky group engage in even riskier sexual practices, which we model as an additional increase in the frequency of sexual activity (b) and the probability of unprotected sex (q). Therefore, the frequency of sexual activity rises from $n(1+a)$ to $n(1+a)(1+b)$ and the probability of risky sex increases from *par* to $p+r+q$ among risky group members that take up circumcision. This is shown in Table B1.

Table B1: Change in Frequency of Sex and the Probability of Risky Sex after Male Circumcision

	Male circumcision	Before MC	After MC
Risky Group	Yes	$n(1+a)$	$n(1+a)(1+b)$
		$p+r$	$p+r+q$
	No	$n(1+a)$	$n(1+a)$
		$p+r$	$p+r$
Safe Group	Yes	$n(1-a)$	$n(1-a)(1+b)$
		$p-r$	$p-r+q$
	No	$n(1-a)$	$n(1-a)$
		$p-r$	$p-r$

The transmission rates for those who are circumcised and uncircumcised are shown in Table B2. The medical impact of male circumcision is that it reduces the transmission rate by half. However, the rate also differs by whether one uses a condom or not. The transmission rate without condom is x and the one with a condom is y among the uncircumcised. In one scenario, we use the point estimate of 0.0017 for the transmission rate x and 0.0006 for the transmission rate y, based on Magaret et al. (2016). The 95% confidence interval for these estimates is as follow.

$$x \in CI[0.0006, 0.0044]$$

$$y \in CI[0.0002, 0.0017]$$

We parametrize the transmission rates given they are point estimates with large confidence interval. Both confidence intervals are partitioned into 5 equidistant intervals and we use the resulting 6 endpoints as plausible values of the transmission rates (see Table B4). For our second scenario, which is the most likely to find large treatment effects we choose the largest possible transmission rate without protection (0.0044) and the smallest transmission rate with a condom (0.0002).

Table B2: Transmission Rate

	With circumcision	Without circumcision
Without condom	*0.5*x*	*X*
With condom	*0.5*y*	*Y*

Section B3: Population HSV infection rate

We compute the population HSV infection rate using the parameters described above.

The first step is to calculate the infection rate of each group. Then, the population infection rate is simply the weighted sum of the groups with weights determined by the take-up parameter, *t*. The infection rate for each group is the following:

(1) Circumcised among risky group

$$n(1+a)(1+b)\{(p+r+q)*0.5x+(1-p-r-q)*0.5y\}$$

(2) Uncircumcised among risky group

$$n(1+a)\{(p+r)*x+(1-p-r)*y\}$$

(3) Uncircumcised among safe group

$$n(1-a)\{(p-r)*x+(1-p+r)*y\}$$

Thus, the population infection rate is simply the weighted sum:

$$i = tn(1+a)(1+b)\{(p+r+q)*0.5x+(1-p-r-q)*0.5y\}$$

$$+(0.5-t)n(1+a)\{(p+r)*x+(1-p-r)*y\}$$
$$+0.5n(1-a)\{(p-r)*x+(1-p+r)*y\}$$

Note that the population infection rate equation above can be rearranged to express the probability of risky sex, p, in terms of other variables. This will be used to fit the model using the control group condition.

$$p=[-tn(1+a)(1+b)\{0.5x(r+q)+0.5y(1-r-q)\}$$
$$-(0.5-t)n(1+a)\{xr+y(1-r)\}-0.5n(1-a)\{-xr+y+yr\}]/$$
$$[(x-y)\{0.5tn(1+a)(1+b)+(0.5-t)n(1+a)+0.5n(1-a)\}]$$

Section B4: Simulation

We run simulations with Python to find a set of parameters that generate results consistent with the findings from our RCT. We proceed in two steps. First, we assume plausible values of the 7 unknown parameters, *(n, a, b, r, q, x, y)* and use the control group condition shown in Table 3 to pin down one parameter, the probability of risky sex, *p*(n, a, b, r, q, x, y)*. Thus, all the pairs of parameter values, *(n, a, b, p*(n, a, b, r, q, x, y), r, q, x, y)*, are the ones that yield a control group infection rate *i=0.084* given a take up rate *t=0.33*. Next, plugging in the same set of parameter values with take-up rate, *t = 0.466*, we compute the treatment group infection rates. The ultimate goal is to find a set of parameter values $(\bar{n},\bar{a},\bar{b}\,p^*(\bar{n},\bar{a},\bar{b}\,\bar{r},\bar{q}\,\bar{x},\bar{y}),\bar{r},\bar{q}\,\bar{x},\bar{y})$ that generate treatment group infection rate, *i = 0.111*.

Table B3: Infection Rate and Take-up Rate of Control and Treatment Group

	Infection rate (*i*)	Take-up rate (*t*)
Control group	0.084	0.33
Treatment group	0.111	0.466

The values used for each parameter are the following.

Table B4: Parameter Values Used for Simulation

Variables	Values
Percentage increase in the frequency of sexual activity (*a*)	{0, 0.1, … 0.5}
Percentage increase in the frequency of sexual activity (*b*)	{0, 0.25, … 1.25, 1.5}
Percentage point increase in the probability of risky sex (*r, q*)	{0, 0.1 … 0.5}
Transmission rate without protection (*x*)	{0.0006, 0.00136, … 0.00364, 0.0044}
Transmission rate with protection (*y*)	{0.0002, 0.0005, … 0.0014, 0.0017}

Section B5: Simulation results

We fix the average frequency of sexual activity at 60 (n=60) during the study period. As discussed previously, for the first scenario, we use the transmission rate $x = 0.0044$ and $y = 0.0002$ and for the second scenario we use $x = 0.0017$ and $y = 0.0006$ which is the point estimate of Magaret et al. (2016). The main results of this simulation exercise are best shown graphically. In Panels A and B of Figures 2, we show for our two scenarios how the treatment group infection rate varies with the percentage increase in the probability of having a risky sexual encounter and the increase in the percentage increase in the frequency of sexual encounters. In Panel A, the areas that are more yellow correspond to treatment group infection rates around 11%. One can observe that such infection rates are achieved when there is a doubling in the frequency of sexual encounters and an increase in the probability of risky sexual behavior of about 35%.[44] In Panel B, we show the results for the second scenario using the point estimates for infection for infection using and not using condom protection. In this scenario, the maximum infection rate that we can reach with our simulation is 9.5%. While this result is smaller than the point estimate of our treatment effect (11.3%), we note that our treatment effect is estimated with noise, and therefore the confidence interval contains the treatment effects from our second scenario as well as other scenarios that require much smaller selection and risk compensation effects. In sum, this simulation exercise results suggest that the relatively large treatment effects that we estimate could be explained by changes driven by a combination of a selection effect as well as large but reasonable changes in risk compensation.

[44] In Figure A5, we also show the corresponding results for the (100%, 0%), (90%, 10%), and (70%, 30%).