

Mixed Method Evaluation of a Passive mHealth Sexual Information Texting Service in Uganda

Julian Jamison, Dean Karlan, and Pia Raffler

Abstract

We evaluate the impact of a health information intervention implemented through mobile phones, using a clustered randomized control trial augmented by qualitative interviews. The intervention aimed to improve sexual health knowledge and shift individuals towards safer sexual behavior by providing reliable information about sexual health. The novel technology designed by Google and Grameen Technology Center provided automated searches of an advice database on topics requested by users via SMS. It was offered by MTN Uganda at no cost to users. Quantitative survey results allow us to reject the hypothesis that improving access to information would increase knowledge and shift behavior to less risky sexual activities. In fact, we find that the service led to an increase in promiscuity, and no shift in perception of norms. Qualitative focus groups discussions support the findings of the quantitative survey results. We conclude by discussing a potential mechanism explaining the counterintuitive findings.

JEL Codes: I11, I15

Keywords: mHealth, SMS, text message, health care.

Mixed Method Evaluation of a Passive mHealth Sexual Information Texting Service in Uganda

Julian Jamison
Consumer Financial Protection Bureau and IPA

Dean Karlan
Yale University, IPA, J-PAL, and NBER

Pia Raffler
Yale University and IPA

We thank Google.org and Google.com for funding. The Principal Investigators and Innovations for Poverty Action retained full intellectual freedom throughout this study. We also thank MTN, Grameen Technology Center and Google for participation and collaboration and adherence to experimental protocols. IRB approval for human subjects protocols granted by Yale University 0901004653, Innovations for Poverty Action 08Dec002.No101 and the Uganda National Council of Science and Technology #SS2176. Any views expressed are those of the authors and do not necessarily represent those of the funders or collaborating institutions or firms. Thanks to Lynn Conell-Price, Rebecca Furst-Nichols, Angela Garcia, Kareem Haggag, Priyanka Kanwar, Selvan Kumar, Jennifer Long, and Sara Lowes for research and field assistance. The views expressed are those of the authors and do not necessarily reflect those of the Consumer Financial Protection Bureau or the United States.

CGD is grateful to its funders and board of directors for support of this work.

Julian Jamison, Dean Karlan, and Pia Raffler. 2013. "Mixed Method Evaluation of a Passive mHealth Sexual Information Texting Service in Uganda." CGD Working Paper 332. Washington, DC: Center for Global Development.
<http://www.cgdev.org/publication/mixed-method-evaluation>

Center for Global Development
1800 Massachusetts Ave., NW
Washington, DC 20036

202.416.4000
(f) 202.416.4050

www.cgdev.org

The Center for Global Development is an independent, nonprofit policy research organization dedicated to reducing global poverty and inequality and to making globalization work for the poor. Use and dissemination of this Working Paper is encouraged; however, reproduced copies may not be used for commercial purposes. Further usage is permitted under the terms of the Creative Commons License.

The views expressed in CGD Working Papers are those of the authors and should not be attributed to the board of directors or funders of the Center for Global Development.

1 Introduction

The rapid adoption of mobile phones in developing countries has created new opportunities for dissemination of information to large populations at minimal cost. Many technology interventions aim to address societal or economic problems by passively making improved information available to users - for example, by disseminating price and market information in order to lower transaction costs and allow trades to occur that otherwise would not (e.g. Jensen's [2007] study of market price information for fisherman in India). However, other technology interventions may require more active engagement by service providers and end users in order to achieve their goals, as Toyama (2008) finds in a case study of computer-based educational programs in India. Although the split between passive and active is not always a clear dichotomy, one way to divide interventions is between those where users make queries to get specific information versus those where a central entity disseminates information widely and without prompting to a large number of individuals.

Many public health organizations have designed projects that use mobile technology to support health services and health education - typically called mHealth projects (Vital Wave 2009). MHealth technologies have been adopted in many contexts in developed nations, as the internet has revolutionized people's access to health information (Lim et al. 2008). In less developed countries, where access to mobile phones has become increasingly common even as penetration of internet access and fixed phone lines remains low (Vital Wave 2009), such interventions have been actively promoted. Reviews of mHealth projects in developing countries (Kaplan 2006, Mechael 2009, Deglise et al. 2012, Gorman et al. 2012) have found some positive evidence on the success and cost-effectiveness of using mobile technology. However this literature is young, and mostly comprised of case studies reporting on operational outcomes and technology use rather than the more rigorous randomized trial methodology (Kahn 2010). The mHealth literature has also not explored how variations in the design of technological interventions may affect users' outcomes. In particular, the question remains: in such interventions, is merely passively improving access to information sufficient to further specific social health goals?

Here we use a mixed methods approach to examine a passive, user-initiated, mobile phone-based health intervention in Uganda. The goal of the intervention was to provide improved access to sexual health information, in order to improve users' knowledge of safe and unsafe sexual behaviors and ultimately lead to reductions in risky behavior. The underlying assumption was that lack of information leads individuals to underestimate specific risks that they face, and to engage in riskier behavior than they otherwise would have chosen. Thus better information can make risks more salient and thus more likely to influence behavior.

Uganda has several features that make it a prime candidate for a technological intervention focused on sexual health. HIV prevalence is high, estimated to be 6.5% among adults aged 15-49 in 2009 (UNAIDS 2010). In addition, knowledge regarding sexual health and HIV/AIDS is low, and risky sexual practices

are prevalent (UNAIDS 2010). The rapid increase of teledensity, from under 3% in 2002 to 33.5% in 2010 (UCC 2009, UCC 2011), combined with a total adult literacy rate of 75% (2008) (UNAIDS 2010), allowed this mHealth intervention to reach a large population.

The sexual health intervention we study was implemented by Google, the Grameen Technology Center, and MTN. The technology was developed primarily by Google and Grameen Technology Center. It enabled individuals to send free SMS messages with questions about sexual health, family planning, and local health services. The SMS services were provided by MTN, which is the largest mobile service provider in Uganda, with 60% of the market share in 2008 (Mulira 2010).

Studies of SMS-based mHealth interventions in wealthier countries have found evidence of significant improvements in preventive health behaviors, such as smoking cessation (e.g. Free et al. 2011) and self-management of asthma and diabetes (Galant & Maticka-Tyndale 2004, Istepanian et al. 2009, Quinn et al. 2011). An intervention in Australia with similar aims to the intervention studied here showed that receipt of SMS messages with information on safe sexual practices over a one-year period was associated with improved sexual health knowledge and higher self-reported rates of STI-testing among women (Lim et al. 2012).¹

In the developing world, the existing literature on information interventions has shown that providing individuals with information on new practices or technologies can (but does not always) significantly increase adoption of new agricultural practices (UNESCO, Foster & Rosenzweig 1995, Duflo, Kremer & Robinson 2010), cleaner water sources or methods for water purification (Madajewicz et al. 2007, Jalan & Somanathan 2008), and safer sexual practices (Adetunji & Meekers 2001, Dupas 2011, Chong et al. 2013). SMS-based mHealth interventions more similar to our study have been introduced in developing countries (UNAIDS 2010) and there is suggestive evidence of positive impact on behavior², although not generated through randomized trials. For example, mobile phone-based health projects in Africa with aims such as supporting community health workers (Mahmud, Rodriguez & Nesbit 2010) or improving adherence to antiretroviral treatment (ART) for HIV-positive patients (Haberer et al. 2010, Lester et al. 2010, Pop-Eleches et al. 2011) have found some evidence of significant impacts.

A valuable aspect of mobile technology as a health tool is that it allows two-way communication. Here, we examine one of the first technologies in the

¹However, it should be noted that the sexual health information which was presented to users in the Australian case was chosen by a socially-minded organization, whereas in our setting information was disseminated in response to questions asked by end users, and the answers provided were factual in nature rather than overtly conveying a normative purpose or aim.

²In 2008, the six-week pilot program of Text to Change in Uganda, later also run in Kenya, which administered an incentivized SMS quiz on HIV/AIDS and encouraged STI testing, was associated with a significant increase in the number of people who came to health clinics for STI testing. See <http://www.texttochange.org/content/we-created-sms-campaign-increase-hiv-aids-awareness-uganda>

developing world that exploits that capability, enabling individuals to request information on sexual health from a database of reliable information. To our knowledge, this is the first evaluation of a project enabling individuals to receive SMS responses to direct health questions³ and the first rigorous evaluation of the impacts of any mobile phone based health information intervention in a developing country. The closest projects that we are aware of involve user queries followed by responses (either text or phone) from a trained counselor.⁴

The paper is structured as follows. Section II provides an overview of the intervention, while section III discusses the experimental design and the data. Section IV presents counterintuitive empirical findings the intervention’s effects on knowledge, behavior and outcomes related to sexual and reproductive health. It also discusses potential mechanisms for these effects, informed by focus groups with users. Section V concludes.

2 Intervention

Together with Google and Grameen Technology Center, MTN developed and implemented a novel, interactive text-messaging platform in Uganda for mobile telephones. The service, referred to as 6001 (the phone number to which to send a text message), allows mobile phone users to text questions on sexual and reproductive health to a server. Once the text message is received, an algorithm-based trigger word search matches the query with a pre-prepared piece of advice from a database. The database of responses was jointly compiled by Marie Stopes International Uganda, an international NGO working on reproductive health issues, and Straight Talk Foundation, a Ugandan NGO specializing in communication on adolescence and sexuality. The database content consists of approximately 500 unique messages in categories including HIV/AIDS; other STIs; maternal and neonatal health; body changes and sexuality; and family planning. Messages are about 500 characters long, consist of factual information presented in simple language, and often include an encouragement to use condoms or get tested for HIV. Users can send queries in English or Luganda and receive responses in the respective language. A sample of an SMS exchange in English is as follows:

³A similar technology with an SMS-searchable database of information on prescription medications has been designed but not evaluated by Langer, Kumal and Mittal (2009).

⁴See “Text Me! Flash Me!” in Ghana (<http://www.comminit.com/node/291748>) and “Mobile4Good” in Kenya and Nigeria (<http://oneworldgroup.org/mobile4good>).

Sample question 1: What does HIV reinfection mean?

Response: There are different types (strains) of HIV/AIDS so even if you are already infected you can catch another type of the virus which can make you more sick. HIV reinfection is when somebody who already has HIV gets the virus again by having sex without a condom with an infected person or if infected fluid gets into their body through cuts/unsterilized sharp instruments. Use condoms EVERY TIME to protect yourself + others.

Trigger words: HIV, Aids, reinfection

Sample question 2: Does urinating after sex kill sperms?

Response: Pregnancy happens when male sperm meets a female egg. They meet as a result of sex. Sperm have tails to help them swim up to the fallopian tubes + fertilize an egg. Pregnancy can NOT be avoided by having sex standing up or urinating after sex. Family planning is the ONLY way to stop pregnancy. ONLY CONDOMS PREVENT HIV/AIDS+STIs. Reply 1 to learn about birth control, 2-emergency contraception, 3-jumping after sex, 4-asprin/panadol after sex to stop pregnancy, 5-signs of pregnancy.

Trigger words: Urinate, kill, sperm

An additional feature of the service is an interactive clinic directory. Users can query clinic locations by sending in a text message with a key word and the location they are interested in and receive a reply with the names, contact information and services offered by the nearest health facilities.

The service was developed in cooperation between Google.org, Google, Grameen Foundation and MTN independently of the study. The initial idea was to capitalize on the expanding reach of mobile networks and Google's technological know-how in order to improve access to socially beneficial information. After prototyping a number of different ideas, the group settled on an SMS-based interactive advice service on sexual and reproductive health for three reasons. First, misconceptions about sexual and reproductive health are widespread in Uganda, and access to reliable information or sexual education is limited. Second, the team deemed sexual and reproductive health an area of healthcare that lends itself well to giving advice without providing diagnoses. Third, the team thought that confidentiality and anonymity of the mobile phone-based service would be especially valued. The service was piloted in two suburbs of Kampala and launched in June 2009. Use of the service was open to the public and available to both MTN and non-MTN network users, although it was free for MTN clients and cost the price of a regular text message for non-MTN clients.⁵

Figure 1 in the Appendix shows pie charts that summarize categorically the

⁵ Although the original plan by MTN was to charge a nominal fee for the service, as of yet, no fee has been added by MTN nor are we aware of a plan to add one.

type of information disseminated. Not considering ‘uncategorized’ responses,⁶

the modal category of queries was ‘body changes and sexuality,’ with 31% of answers. Within this category, 40% were questions on sex (12% of overall answers) and 31% were questions on genital organs (9% of overall answers). The second largest category is miscellaneous (17%) and the third largest category is HIV (10%). Maternal and neonatal health comprised 8%. Although issues of intellectual property and participant privacy prevented a rigorous test for concordance, in ad-hoc testing by the researchers and more rigorous testing by the developers, the English-language information disseminated matched the information requested well. According to anecdotal evidence, the search algorithm performed less well in Luganda. This notwithstanding, among those respondents who reported having used the service in our endline survey, 73% responded that they found the information they had received from the service very useful, and another 19% said that they found it useful. Responses are very similar when restricting the sample to those respondents who reported having used the service in Luganda (69% and 22% said they found the information received from the service very useful or useful, respectively).⁷

3 Experimental Design and Survey and Qualitative Data

Clustered Encouragement-Design Randomized Evaluation

We conducted a clustered randomized evaluation in 60 villages across the four districts of Masaka, Mpigi, Mityana and Mubende in Central Uganda. Villages were randomly assigned to either the treatment or control group. Treatment villages were exposed to a targeted high-intensity marketing campaign by a professional marketing firm from August through October 2009. Marketing teams visited trading centers three to six times, depending on population size, and each time spent a full day promoting the service through demonstrations, flyers and posters in both English and Luganda. This marketing thus provides the “encouragement” to use the service.⁸The new service was presented as an

⁶Unfortunately, miscellaneous and uncategorized texts make up about one third of the responses. Due to privacy regulation we only have access to the categories of sent responses, and are therefore not able to categorize these texts.

⁷The corresponding survey question read “How useful would you say was the information you received: very useful, useful, somehow useful, or not useful?” The question was answered by 277 respondents, of which 180 said they had used the service primarily in Luganda.

⁸Encouragement designs are often used in social science evaluations where it is difficult or impossible to ensure that all members of the treatment group get fully treated, and no members of the control group receive any treatment (i.e. perfect compliance). Instead, the treatment group is “encouraged” to participate through information or incentives, and the control group is not encouraged. The estimation depends on the strength of the encouragement to create differential usage in treatment and control groups. Note that this approach estimates the impact only on those whose participation is affected by the encouragement. Those who would have participated regardless of the encouragement are treated in both the treatment and control group, and thus differenced out in the treatment effect estimates. Note that the encouragement design in this study does create differential usage initially, but that usage

anonymous, objective source of information to which users could send any question related to sexual and reproductive health. No benefit or incentive accrued to respondents in the treatment group other than being informed about the free service and trained in using it, and the marketing was conducted as part of a natural marketing push by a well-established marketing firm.

The goal of the targeted marketing campaign was to achieve higher uptake in treatment locations than in control locations. The entire study area was largely blocked off from national marketing media campaigns, such as radio and newspaper ads, in order to minimize awareness and use of the service in the control group.⁹ As the differential in uptake between treatment and control group in Figure 1 and Table 2a in the Appendix show, this was effective. Although usage is not substantially different between treatment and control groups after November 2009, most of the overall usage in the study area comes before that point and was heavily concentrated in the treatment communities.¹⁰ This allows us to evaluate rigorously the impact of introducing the service, but note that this is not an estimate of a program that receives ongoing encouragement and marketing, since the program only did one large initial marketing effort. The study was approved by three human subjects committees: Yale University, Innovations for Poverty Action and the Uganda National Council of Science and Technology (#SS2176).

Quantitative Data

The quantitative data come from four sources: an individual survey conducted at baseline, another at endline, a community survey, and metadata on topics queried at 6001 from MTN. All quantitative data were collected using identical procedures in both treatment and control areas.

The baseline survey was conducted prior to the launch of the intervention by enumerators hired by our research team independently of MTN, Grameen Foundation and Google. A total of 1,791 randomly selected individuals from all 60 study villages were interviewed in February 2009. The household selection was conducted by random walk,¹¹ and individual respondents were randomly

eventually converges between treatment and control group, as shown in Figure 1. Thus, our research design estimates the impact of the initial encouragement marketing push, not that of a program that continues to have ongoing marketing.

⁹While the national launch of the service took place about the same time as the high intensity marketing campaign, ads were limited to local newspapers and radio programs which target other parts of the country. MTN marketing trucks did not pass through the study districts.

¹⁰Between December 1, 2009 and April 1, 2010, the average total number of text messages sent per day from all phones associated with treatment villages was 4.0, compared to an average total number of 1.5 messages sent per day from all phones associated with control villages. The difference in means is highly statistically significant

¹¹For random selection of respondents, field teams drew detailed maps of the village and visited every x th household, where x was dependent on the overall population of the village. If no person was eligible at the selected household, enumerators went to the immediate next household. Once they had identified an eligible respondent, they continued to the x th next household.

selected from a household roster. Household members were eligible to participate in the survey if they were 18-35 years old, had a functioning phone with a MTN SIM card in the household, and had completed a minimum of six years of primary school (because literacy was essential to use the 6001 service).

The follow up survey was conducted one year after the baseline in February 2010, with an increased sample of 2,424 respondents. One methodological goal of this study was to determine if being previously surveyed had an effect on mHealth service usage. Therefore a randomly selected 1,200 baseline respondents were interviewed again for the follow-up,¹² and 1,224 new respondents were chosen from among the sample communities according to the same criteria as in the baseline.¹³ Given that the marketing intervention only began in August 2009, the period between the first encouragement to use the service and endline data collection was six months. The plan had been to roll out the intervention immediately following the baseline, but delays occurred on the side of the implementing partners. The follow-up survey could not be extended because of contractual obligations and negotiations with the different parties to this intervention.

The community survey on general community characteristics was conducted with two to four knowledgeable people from each community, including the elected local chairperson, at the same time as the baseline and follow up surveys. Observable characteristics such as mobile network coverage, distance to nearest health center, and distance to tarmac road were verified by enumerators. Finally, we also collected data from MTN and Google on the topics queried on 6001 for the entire study period.

Qualitative Data

In addition to the quantitative data sources described above, we conducted eight focus group discussions and 39 individual in-depth interviews in eight treatment villages after the endline was conducted. Two treatment villages from each of the four study districts were randomly selected for inclusion in the qualitative study. Within the villages, researchers interviewed a convenience sample of people who had used the service. Given the gender sensitive topic, men and women were interviewed separately by an enumerator of their own gender. Topics covered include perception of the service; usage behavior; perceived knowledge, behavior, and attitude changes; and user experience. All interviews, both qualitative and quantitative, were conducted in the local language.

¹²A total of 385 selected respondents had to be replaced with other randomly selected respondents from the baseline because they could not be located for the interview at endline. Attrition is balanced across treatment and control groups.

¹³That is: Aged 18-35, current MTN phone ownership in the household (i.e. at time of the endline), and completed at least six years of primary education. The households were chosen within villages in the same random walk manner.

Quantitative Measures

As main outcome variables from the survey, we use composite indices of HIV knowledge, contraception knowledge, attitudes towards condom use, non-promiscuity, safe sexual behavior, sexual health outcomes, health seeking behavior, and perceived own risk compared to others.¹⁴

We use a list randomization technique (Corstange 2009; Karlan & Zinman 2012) to get more truthful answers on sensitive topics, such as condom use, the number of sexual partners, unfaithfulness, and attitude changes with respect to the social acceptability of these behaviors. A randomly selected half of our sample was asked the sensitive question in the form of a true or false statement (direct elicitation). In addition, we asked them how many of four innocuous statements were true for them, without telling us which ones. The other half of the sample were not asked the sensitive question, but how many of five statements consisting of the same four innocuous ones plus the sensitive statement, and were again asked how many of these were true for them, without telling us which ones (indirect elicitation). Since respondents were randomly assigned to either elicitation method, the number of the four innocuous statements which are true for them is in expectation the same.

Randomization to the direct or indirect elicitation group was repeated for each survey question asked this way. Subtracting the mean number of true statements in the direct elicitation group from the mean number of true statements in the indirect elicitation group will then give us an unbiased estimate of the percentage of our sample for whom the sensitive statement was true, without however being able to tell for which individual it was true. For an indirect measure of the number of sexual partners, we asked respondents assigned to the indirect elicitation group to add the number of sexual partners in the past three months to the face of a die which they rolled without showing the numerator

¹⁴Index components are equally weighted and had been standardized prior to index construction. Missing values resulted in list-wise deletion. Respondents receive one point for each 'good' response. The indices presented in table 2 consist of the following components: (1) Contains variables on HIV transmission during pregnancy, delivery, breastfeeding, and non-transmission by sharing food or through mosquito bites. (2) Contains variables on named contraception methods (pills, injections, female condoms, foam, IUD, implants, male condoms, emergency pill, female sterilization, male sterilization). (3) Contains variables on agreement to "A male condom should always be put on before intercourse" and "It is not embarrassing to buy or ask for a condom". (4) Contains: Never been unfaithful in past 3 months, number of sexual partners in past 3 months, relationship to last sexual partner was neither casual acquaintance nor commercial sex worker interacted with recent sex, used any type of contraceptives during last sex interacted with recent sex and used condom during last sex, interacted with recent sex. (5) Contains: Never been unfaithful in past 3 months and number of sexual partners in past 3 months. (6) Contains: Relationship to last sexual partner was neither casual acquaintance nor commercial sex worker interacted with recent sex, used any type of contraceptives during last sex interacted with recent sex and used condom during last sex, interacted with recent sex. (7) Contains: Ever had an STD, ever had a bad smelling discharge, ever had genital sore or ulcer, current (or partner's) pregnancy is unwanted and currently pregnant (or partner). (8) Contains: Ever tested for HIV, ever sought treatment for an STD and ever visited health center / worker for an issue related to SRH. (10) Contains variables on perceived risk taking relative to other people and perceived own HIV risk. Average treatment effects on each index component are presented in tables 3-6 in the Appendix.

and to only report the sum of the two values. Subtracting 3.5, the average face value of a fair die, yields the average number of sexual partners reported by individuals in the indirect elicitation group.

One advantage of these methods is that they elicit responses to sensitive questions which are less biased by social desirability. Another advantage is that they allow us to estimate the impact of the intervention on changes in perceptions of social norms, by comparing the difference between direct and indirect responses, i.e. a measure of social desirability, across treatment and control groups, of specific behaviors. To our knowledge, the use of indirect elicitation techniques to estimate the impact of an intervention on attitudes, on perceptions of what one *should* answer, is novel.

Orthogonality and Treatment Compliance

Table 1 shows that the assignment to treatment was orthogonal to key demographic variables, mobile phone usage, community characteristics, and baseline values for key outcomes measures.¹⁵ Of particular relevance for the intervention is the ability to read. Over 80% of respondents are able to read without difficulty in either English or Luganda, as observed by enumerators. Literacy levels are balanced across treatment and control group.

Table 2a in the Appendix shows the usage data, from both self-reports and from MTN data, that confirm the success of the encouragement design in generating large differences in usage between treatment and control areas. 7% of respondents in the control group ever sent a text message to the service’s shortcode 6001, compared to 40% in the treatment group. The difference in the number of text messages sent between treatment and control groups remains significant after the end of the marketing campaign; although we do see a marked drop in use of the service after the marketers left the villages. In qualitative interviews, respondents in treatment villages said they would have liked to be reminded about the existence of the service for a longer time period. Usage trends are displayed in Figure 1. Among those who ever used the service, the median number of text messages sent to 6001 is four. Figure 2 displays the distribution of the number of text messages sent to the service. About 40% of the users sent messages on only one day, 21% on two different days, 23% on three to five, and 16% on six or more different days. The median number of days between first and last use of the service is 6 days. Among those above median in terms of frequency of usage, usage was over an extended period of time: the median number of days between first and last use is 35. Users were more likely to be male, young, married, to own a personal phone, have slightly higher education levels and to be wealthier (proxied by number of meals eaten per day). High frequency users had comparatively low levels of knowledge of sexual and reproductive health at baseline (see table 2c in the Appendix).

¹⁵Balance tests for additional variables are presented in Table 1 in the Appendix.

4 Empirical Results

We use survey data to measure different outcomes Y for estimating intention-to-treat effects with the Ordinary Least Squares specification:

$$Y_i = \beta_0 + \beta_1 * T_i + \beta_2 * BY_i + \beta_3 * Endlineonly_i + \beta_4 * X_i + \epsilon_i$$

where Y_i denotes an outcome variable, T_i is a dummy variable which takes the value 1 if individual i lives in a village assigned to the treatment group, BY_i captures the baseline value of outcome variable Y_i for respondents on which we have both baseline and endline observations, and takes the value 0 otherwise. $Endlineonly_i$ is a dummy variable that takes the value 1 if a respondent was in the endline only. X_i is a vector of control variables including demographics and the stratification variables. The robust error term ϵ_i is calculated allowing for clustering at the village level, which is the unit of randomization. In addition, we analyze whether impact was heterogeneous with respect to gender, as many interventions on sexual health target by gender in particular (although this intervention had no such focus).

Our primary theory was that the intervention effectively improved access to reliable information about sexual health, and that this would lead to higher knowledge. With the assumption that low levels of information lead individuals to underestimate the riskiness of unsafe sex, the theory of change then posits that higher knowledge will change attitudes and then lead to less risky sexual behavior. We now proceed through this theory of change step-by-step:

Hypothesis #1: Access to the texting service increases knowledge on HIV and/or contraception. Although respondents reported that “those who used the service learnt how to protect themselves against dis-eases like AIDS, syphilis and other diseases” in qualitative interviews, in the quantitative survey data we do not find a measurable impact on knowledge relative to the control group. Table 2 Panel A presents results for H1 with respect to knowledge regarding possible modes of HIV transmission and regarding contraception methods and use, respectively. We find no support for H1 in aggregate for either outcome index. We also do not find heterogeneity with respect to gender.

Hypothesis #2: Access to the texting service changes attitudes towards safer sexual practices with respect to condom use. The theory of change predicts that increased knowledge would then shift attitudes towards less risky sexual attitudes towards condom use. However, since knowledge did not increase, we do not expect to see a change in attitudes. Table 2 Panel B presents results for H2, and indeed we find no support for H2 in aggregate or by gender.

Hypothesis #3: Access to the texting service, through an increase in HIV knowledge, leads to more safe sex and less promiscuous sexual activity. The observed impact on sexual behavior is the opposite of what we expected. Our theory of change predicts that increased knowledge would then shift attitudes,

which would then shift behavior. Since H1 and H2 were rejected, we do not expect to find support for H3. Table 2 Panel C presents results for sexual behavior. No support is found for H3, and in fact the results are the opposite: we find the composite index of non-risky behavior¹⁶ decreases (i.e., shifts towards riskier behavior) as does the index of non-promiscuity. We find the overall behavioral index decreases by 0.11 standard deviations (p-value of 0.017). The treatment effect for males is two and three times larger, respectively, than that for women, however this gender difference is not significant statistically (p-value of the test of equality of the coefficients for men versus women is 0.53 for the aggregate index, 0.17 for the non-promiscuity index, and 0.75 for the safe sex index). One possible explanation for this finding could be that people learned from the service how to protect themselves. As one respondent put it in a qualitative interview: “If men learnt through this service that there is a way of minimizing the chances of being infected with an STI, say, by using a condom, they may find the courage to begin having more relationships outside the one they are already having.” However, we believe the evidence suggests people engaging in promiscuous behavior were no more likely to use condoms.¹⁷

Table 3 reports results on direct versus indirect elicitation (i.e., employing a list randomization technique (Corstange 2009)) on three key behaviors: use of condoms during last occurrence of sexual intercourse, unfaithful behavior in last 90 days, and number of sexual partners in last 90 days. Again, we find little support for hypothesis H3. Using indirect elicitation, we find no significant improvements in sexual behavior, but the opposite.

In fact, for two of the three outcome measures, we find the opposite of the intended effect. For infidelity, we find an increase overall from 12% to 27% (p-value of 0.02). Infidelity is defined as ever having been unfaithful to one’s current partner in the past three months. Single respondents are recorded as not having been unfaithful. For number of sexual partners, we find an overall increase from 0.62 to 0.85 (p-value of 0.04). Men increase the number of partners from 0.63 to 1.04 (p-value of 0.02), whereas women increase the number of partners from 0.61 to 0.69 (p-value of 0.59). One might have expected at least women to under-report their number of partners, but given that the averages are below one, our result is less surprising.

¹⁶The composite non-risky behavior index includes (1) never been unfaithful in past 3 months, (2) number of sexual partners in past 3 months, (3) number of sexual partners in past 3 months, (4) relationship to last sexual partner was neither casual acquaintance nor commercial sex worker interacted with recent sex, (5) used any type of contraceptives during last sex interacted with recent sex and (6) used condom during last sex, interacted with recent sex.

¹⁷We do not find any evidence that the service led people to use condoms (the coefficient on condom use during last sex interacted with having had sex in the past year is small, negative and not significant, see Appendix table 6a, column 6). Since we find evidence that the service resulted in some users being more promiscuous, any increase in condom use among people who became more promiscuous would have needed to be accompanied by a decrease in condom use by other respondents in the treatment group in order to result in the overall null finding on condom use. While we cannot rule this out, we believe this is unlikely, implying that there was no increase in condom use of those engaging in more promiscuous behavior.

Hypothesis #4: Individuals are more aware of the risks associated with their behavior. The theory of change starts with improving access to information, leading individuals to be more informed about sexual health. This measure combines knowledge with a potential change in actual behavior, and examines whether individuals change their perception of the riskiness of their behavior. Table 2 Panel D presents the results for H4. We find in aggregate a decrease in perception of one’s own non-riskiness (i.e. safeness) compared to others. This finding is also confirmed by qualitative interviews. Respondents reported that they became more sensitive to the risks associated with different sexual behavior, summarized by one young male as “I felt that unprotected sex was riskier, and protected sex was safer,” and were more likely to advise friends and relatives whom they considered at risk to go for HIV testing.

Hypothesis #5: Access to the texting service changes individual behavior and perception around norms towards safer and less promiscuous sex, i.e., individuals perceive that it is more acceptable socially to engage in safe sex relative to risky sex, and to engage in less promiscuous behavior. Table 3 presents the results for testing H5, and we find support for it. Here we use the comparison between direct elicitation and indirect elicitation (list randomization and dice role) on the survey to draw inference regarding a change in norms. We compare what individuals tell our surveyors during follow up data collection about their behavior in direct elicitation questions (which may suffer from social desirability bias) compared to what the respondents reveal on average to be their true behavior (through the list randomization and dice role techniques). The empirical test is a difference-in-difference estimator, comparing the treatment minus control for direct elicitation compared to indirect elicitation. We have three outcome measures: no condom use at last sexual intercourse, unfaithful in past three months, and number of partners in past three months.

On the self-reported direct elicitation of risky behavior, we find few if any changes. However, as shown above for H3, we find large increases in risky sex and promiscuous behavior when asked indirectly. This difference between direct and indirect elicitation is important, and revealing, about changes in norms amongst our studied population. It implies that either norms for those in our study did not change (i.e., the direct elicitation reveals what people believe they *should* do) or through increased knowledge individuals learned what they are supposed to say, i.e., a shift in norms towards safer behavior (despite a shift towards riskier actual behavior).

Importantly, we are able to reject the opposite of H5: that behavior shifted towards riskier more promiscuous behavior because norms likewise shifted towards riskier and more promiscuous behavior. Had this been true, we would have observed increases in the direct elicitation method as well as the indirect method. This is consistent with the qualitative interviews, in which respondents throughout stressed individual considerations in engaging in safer or more risky sexual practices, but did not refer to norms shifting.

Why did 6001 Not Have the Intended Effects? Qualitative Evidence

Qualitative interviews shed light on the reasons why the 6001 service had such a counterintuitive impact. For information provision to result in behavior changes and improved health outcomes, people must use the service, obtain useful information from it, and be willing and able to act on it. While a number of qualitative respondents say they learned about risk prevention and STI symptoms from the service and changed their behavior in response to it, others mention reasons for not using the service or not acting upon the information received. We summarize the two most common responses below.

As Figure 1 shows, usage dropped markedly after the marketing period. This appears to be driven by two factors. First, many people said they missed being reminded about the service: "...when there is a new thing introduced, we all get excited about it, but after a while, we tend to forget. That is what happened to me, I used the service for a while, then soon forgot about it, and if you hadn't come to ask about it, I don't think I was going to think about it for a long time. But since you have come, I think I will try it again... by the way, is it still working?" Second, a number of respondents said that they were not satisfied with the quality of the search algorithm in Luganda, which was the main language of usage: "... that was the issue, you would ask one question, but receive another answer." One respondent highlights negative social feedback loops amplifying the effect: "... this service confused people [by] giving inappropriate responses to their questions, and yet it was the newest service at the time. You see when something has just come, people try it and if they get dissatisfied with it they leave it very fast. [...] Now the reason why most people stopped using the service was that whenever they would ask their friends whether they received an appropriate response to their query, the answer they would receive was 'No' - which would resonate with their own experience. That is where it would end, and the service died like that." In sum, our qualitative interviews underline the importance of constant reminders, positive social feedback loops, and a positive user experience in the beginning in order to achieve sustained usage of new technology. Recall also that the majority of respondents claimed to have found the service either "very useful" or "useful," so it is difficult to know exactly how much weight to place on stories of initially confusing or inappropriate responses.

Respondents also suggested that they faced constraints on their ability to change their behavior, even after receiving accurate information on sexual and reproductive health, for two main reasons. First, lack of access to resources could pose a stumbling block. One respondent poignantly asked: "Now, you have the information, and you are even told where to get further tests and treatment, but you don't have money for treatment, or even transport to the place you have been referred to: Now have you been helped at all?" Second, since risky sexual behavior inherently involves more than one person, respondents may not have been able to stop engaging in risky practices due to the power balance in their relationship. Specifically, it can be difficult for Ugandan women to stand up to their male partners: "You may tell him that lets start using condoms

to protect ourselves, after getting advice from these messages, [...] because I am worried about our situation. He then asks why you are worried, how come: all along you had never gotten worried. When you tell him the source of the information, trouble then starts. For example saying that MTN does not live in my house so cannot decide for me what to or not to do.” Based on the qualitative interviews, this effect is even more pronounced for younger, unmarried women of lower socioeconomic status.

Considering that a small number of people used the service in a sustained manner, only some of whom received the relevant information and were then willing and able to act on it, it is not surprising that we see many null effects.

5 Discussion and Conclusion

We evaluate a novel mobile phone-based health program in Uganda, developed and implemented by partner organizations, in which treatment communities were randomly assigned to learn about a service providing individually-initiated access to information on sexual and reproductive health via text messages. We find no increase in health knowledge regarding HIV transmission or contraception methods, and no change in attitudes. Rather than seeing reductions in risky sexual behavior, we actually find higher incidence of risky sexual behavior, and more infidelity, although more abstinence as well. Overall, individuals perceive their behavior as being riskier, which could be an indication of better probabilistic assessments but is also likely a result of the riskier (self-reported) behavior and possibly a desire to answer the surveyor in a particular way.

Mechanisms of change are typically difficult to isolate, and this evaluation is no different. This was a fairly dynamic intervention, with different and endogenous intensities of treatment, likely heterogeneous treatment effects, a limited (albeit still large) sample frame for analysis, and highly sensitive outcomes. That said, we posit two mechanisms that may have been influential: a change in norms towards risky behavior, and sexual sorting. Using a novel application of the indirect list randomization methodology in order to estimate a treatment effect on social desirability bias, we find no evidence that individuals in the treatment group perceive risky sexual behavior as more socially acceptable.

We believe the sexual sorting mechanism is more likely, and worthy of further research to isolate and measure it more fully. In particular, qualitative interviews suggest sexual sorting as a possible mechanism for change. Both male and female respondents reported that married women learned from the service about the risks associated with having an unfaithful partner, or used the service to shine a spotlight on what they already knew, and as a result insisted that their husbands be faithful and go for testing with them. Some husbands complied while others did not, leading women to deny them sex and men to seek it from elsewhere. One woman summarized it as “Once he wants sex, he means it, yet the woman is dodging him around (because she became aware of the risks), so he decides to get another one to satisfy his desires.” Sexual sorting is also supported by the finding that there was on average a four percentage point re-

duction in likelihood of recently having had sex (p-value of 0.02). This is largely driven by women, who are 6.3 percentage points less likely to have had sex in the past 12 months (p-value of 0.00), compared to a decrease of 1.4 percentage points for men (p-value of 0.58). Although technically we cannot reject the null hypothesis that treatment effects are the same for men and women, we do have suggestive evidence that the increase in promiscuity was driven by men (p-value of 0.17), while the decrease in sexual activity was driven by women (p-value of 0.14).

Naturally, since the type of information individuals receive depends on the questions they pose, and there is much variety, there should be no surprise that the impacts indicate a multitude of uses and mechanisms. Previous work on sexual health in East Africa (Madajewicz et al. 2007) has found that young women responded to information about HIV risks by switching to lower-risk partners, but only in the context of an NGO-led program with a clear message and face-to-face interaction.

In our setting, individuals must first choose whether to access the service, and then what to ask, and finally what (if anything) to do with the information received. We find that easing access to this information is not sufficient to induce safer behavior, and may directly or indirectly lead to riskier choices. It may be that only in the presence of a guiding hand, such as a local health worker, can this category of information improve outcomes, but our study does not vary that feature so we can only suggest this as a potentially important mechanism. Further work is needed to study whether such first-order interventions work on particular demographics, and what changes to the intervention, for example inclusion of a personalized component, are necessary for other demographics. We also find evidence from a variety of sources suggesting sexual sorting as a potential mechanism, which generates distributional impacts over and above the average impacts, and we believe that further research should focus on understanding such differential responses to information campaigns. Finally, as in any evaluation, the findings may be contingent on the specific setting in which it occurred and the specifics of the partner organization's implementation. Thus one should not draw extreme conclusions from this study about the efficacy of *any* technology-driven information intervention; rather, it would be appropriate to learn from this study that the mere introduction of an information technology, designed by development experts, but left to individuals to self-direct in terms of their usage, does not necessarily lead to the socially desired impacts set out in the original intention of a program.

References

- Adetunji, J., & Meekers, D. (2001). Consistency in condom use in the context of HIV/AIDS in Zimbabwe. *Journal of Biosocial science*, 33(1), 121-138.
- Aker, J., & Mbiti, I. (2010). Mobile phones and economic development in Africa. Center for Global Development Working Paper, (211).
- Chong, A., M. Gonzalez-Navarro, D. Karlan, & M. Valdivia (2013). Effectiveness and spillovers of online sex education: evidence from a randomized evaluation in Colombian public schools. National Bureau of Economic Research Working Paper #18776.
- Corstange, D. (2009). Sensitive questions, truthful answers? Modeling the list experiment with LISTIT. *Political Analysis*, 17(1), 45-63.
- Déglise, C., Suggs, L. S., & Odermatt, P. (2012). SMS for disease control in developing countries: A systematic review of mobile health applications. *Journal of Telemedicine and Telecare*, 18(5), 273-281.
- Duflo, E., M. Kremer and J. Robinson (2010). Why don't farmers use fertilizer? Evidence from field experiments in Kenya. Working Paper.
- Dupas, P. (2011). Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya. *American Economic Journal: Applied Economics* 3, 1-34.
- Free, C. et al. (2011). Smoking cessation support delivered via mobile phone text messaging (txt2stop): A single-blind, randomised trial. *The Lancet*, 378 (9785), 49-55
- Foster, A. D., & Rosenzweig, M. R. (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy*, 1176-1209.
- Gallant, M., & Maticka-Tyndale, E. (2004). School-based HIV prevention programmes for African youth. *Social Science and Medicine*, 58(7), 1337-1352.
- Gurman, T. A., Rubin, S. E., & Roess, A. A. (2012). Effectiveness of mHealth behavior change communication interventions in developing countries: A systematic review of the literature. *Journal of Health Communication*, 17(sup1), 82-104.
- Haberer, J. E. et al (2010). Real-time adherence monitoring for HIV antiretroviral therapy. *AIDS and Behavior*, 14(6), 1340-1346.

Istepanian, R. S. et al (2009). Evaluation of a mobile phone telemonitoring system for glycaemic control in patients with diabetes. *Journal of Telemedicine and Telecare*, 15(3), 125-128.

Jalan, J., & Somanathan, E. (2008). The importance of being informed: Experimental evidence on demand for environmental quality. *Journal of Development Economics*, 87(1), 14-28.

Jensen, R. (2007). The digital divide: Information (technology), market performance, and welfare in the South Indian fisheries sector. *The Quarterly Journal of Economics*, 122(3), 879-924.

Kahn, J. G., Yang, J. S., & Kahn, J. S. (2010). 'Mobile'health needs and opportunities in developing countries. *Health Affairs*, 29(2), 252-258.

Kaplan, W. A. (2006). Can the ubiquitous power of mobile phones be used to improve health outcomes in developing countries? *Globalization and Health*, 2(9), 1-14.

Karlan, D. S., & Zinman, J. (2012). List randomization for sensitive behavior: An application for measuring use of loan proceeds. *Journal of Development Economics*, 98(1), 71-75.

Langer, A., Kumar, B., Mittal, A., & Subramaniam, L. V. (2009, March). Mobile medicine: Providing drug related information through natural language queries via SMS. In *Advance Computing Conference, 2009. IACC 2009. IEEE International* (pp. 546-551).

Lester, R. T. et al (2010). Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): A randomised trial. *The Lancet*, 376(9755), 1838-1845.

Lim, M. S. et al (2008). SMS STI: a review of the uses of mobile phone text messaging in sexual health. *International Journal of STD & AIDS*, 19(5), 287-290.

Lim, M. S. et al (2012). Impact of text and email messaging on the sexual health of young people: A randomised controlled trial. *Journal of Epidemiology and Community Health*, 66(1), 69-74.

Madajewicz, M. et al (2007). Can information alone change behavior? Response to arsenic contamination of groundwater in Bangladesh. *Journal of Development Economics*, 84(2), 731-754.

Mahmud, N., Rodriguez, J., & Nesbit, J. (2010). A text message-based intervention to bridge the healthcare communication gap in the rural developing world. *Technology and Health Care*, 18(2), 137-144.

Mechael, P. N. (2009). The case for mHealth in developing countries. *Innovations: Technology, Governance, Globalization*, 4(1), 103-118.

Mulira, N., A. Kyeyune, and A. Ndiwalana (2010). Uganda ICT sector performance review 2009/2010.

Pop-Eleches, C. et al (2011). Mobile phone technologies improve adherence to antiretroviral treatment in resource-limited settings: A randomized controlled trial of text message reminders. *Aids*, 25, 825-834.

Quinn, C. C. et al (2011). Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. *Diabetes Care*, 34(9), 1934-1942.

Toyama, K. (2008). Can technology end poverty? *Boston Review*, 35(6).

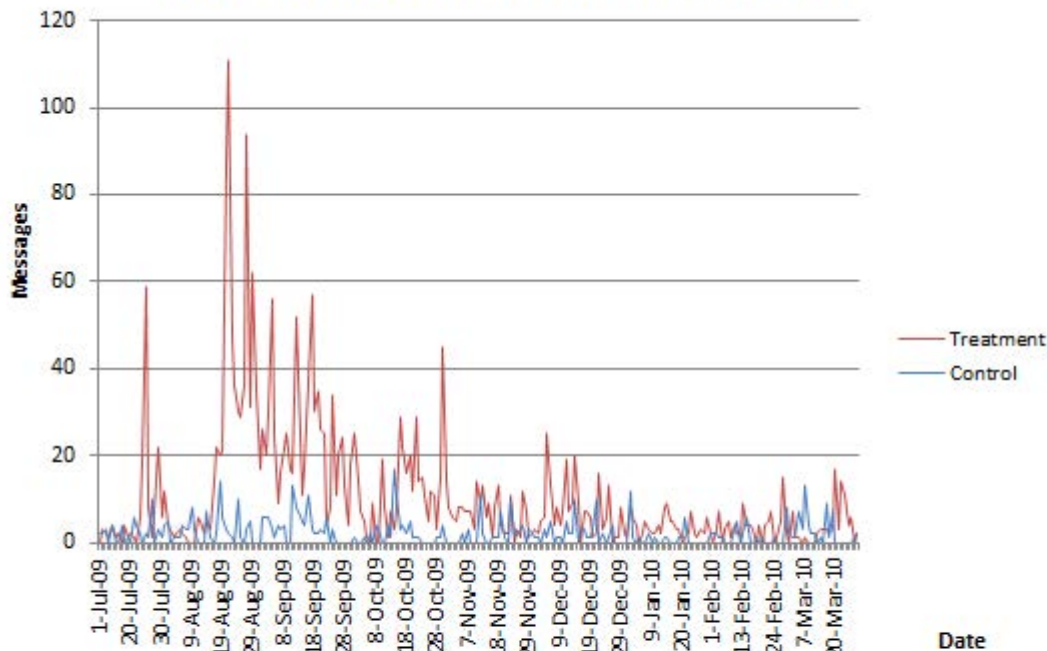
Uganda Communications Commission. Rural Communications Development Policy for Uganda (2009) & (2011).

UNAIDS (2010). 2010 report on the global AIDS epidemic.

United Nations Educational, Scientific and Cultural Organization (UNESCO), Institute for Statistics.

Vital Wave Consulting (2009). The opportunity of mobile technology for healthcare in the developing world (UN Foundation-Vodafone Foundation Partnership, Washington, DC).

Figure 1: Usage Trends by Treatment Status



The graph shows the number of messages sent to 6001 from survey respondents' phone numbers by treatment status, for a total of 3,192 messages from 279 phones, over the course of the study period. Marketing started on August 17, 2010. The usage data displayed here is based on institutional data from MTN. The service was launched prior to the high intensity marketing intervention, explaining the non-zero usage before August 17.

How often have you used 6001? Conditional on having used the service

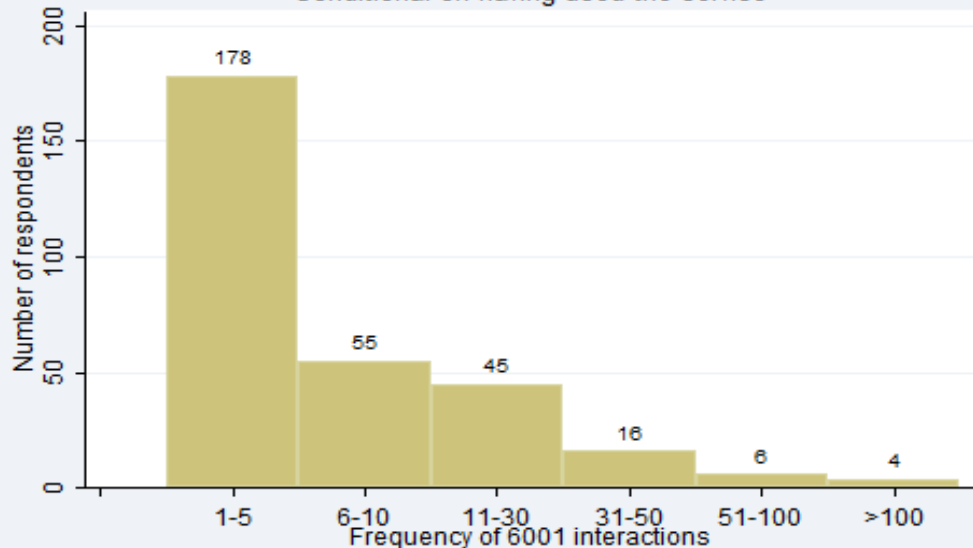


Figure 2: Self-reported number of interactions with service among users

Table 1: Summary Baseline Values and Balance Tests

	Treatment		Control		p-value
	N	Mean	N	Mean	T vs C
	Demographics				
Male	894	0.49	897	0.50	0.59
Age	894	25.37	896	25.26	0.67
Years of education	892	8.14	893	8.11	0.80
Respondent owns a mobile phone (%)	894	0.72	897	0.71	0.67
Ever sent a text message or SMS? (%)	887	0.63	886	0.61	0.30
	Community characteristics				
MTN coverage (1-5)	30	4.07	30	4.00	0.82
Distance to next tarmac road (km)	30	14.79	30	16.09	0.75
	Main outcomes				
HIV knowledge index	894	0.00	895	0.00	1.00
Contraceptive knowledge index	894	0.00	895	0.00	0.96
Condom use attitudes index	891	0.09	887	0.00	0.05
SRH outcome index	894	-0.07	897	0.00	0.12
SRH service seeking index	894	-0.03	897	0.00	0.48
Ever had sex in past 12 months (%)	891	0.81	892	0.80	0.73
Perceived relative non-riskiness index	886	-0.03	891	0.00	0.48

Table 2: Main Outcome Variables

Treatment effects on		Overall			By Gender						
					Male		Female		F-test: Male = Female		
A. Knowledge		β	se	n	β	se	β	se	F-stat	p-value	n
(1)	Composite HIV knowledge index	0.0575	0.04	2,274	0.0347	0.05	0.0793	0.06	0.34	0.56	2,274
(2)	Composite contraception knowledge index	0.0196	0.03	2,275	-0.0124	0.05	0.0501	0.04	1.00	0.32	2,275
B. Attitudes											
(3)	Condoms use attitudes index	0.0064	0.04	2,266	0.0465	0.04	-0.0318	0.05	1.91	0.17	2,266
C. Behavior											
(4)	Composite current behavior index*	-0.1128**	0.05	2275	-0.1552**	0.06	-0.0723	0.06	1.03	0.32	2,275
(5)	Non-Promiscuity index*	-0.1096**	0.04	2266	-0.1679***	0.06	-0.0543	0.06	1.90	0.17	2266
(6)	Safe Sex behavior index*	-0.0521	0.04	2270	-0.0797	0.06	-0.0258	0.08	0.25	0.62	2270
(7)	Composite outcome index	-0.0400	0.04	2,275	-0.0425	0.05	-0.0377	0.06	0.00	0.96	2,275
(8)	Health services index	-0.0210	0.04	2,275	-0.0369	0.06	-0.0059	0.05	0.21	0.65	2,275
(9)	Ever had sex in past 12 months	-0.0393**	0.02	2,264	-0.0142	0.03	-0.0631***	0.02	2.25	0.14	2,264
D. Risk Perception											
(10)	Perceived relative non-riskiness index	-0.1010**	0.04	2,267	-0.1754***	0.06	-0.0300	0.06	2.99	0.09	2,267

Note: All regressions include controls for baseline values unless indicated with *. Robust standard errors are clustered at the village level. Controls for stratification variables are included. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Index composition:

(1) Contains variables on HIV transmission during pregnancy, delivery, breastfeeding, and non-transmission by sharing food or through mosquito bites.

(2) Contains variables on named contraception methods (pills, injections, female condoms, foam, IUD, implants, male condoms, emergency pill, female or male sterilization).

(3) Contains variables on agreement to "A male condom should always be put on before intercourse" and "It is not embarrassing to buy or ask for a condom".

(4) Contains: Never been unfaithful in past 3 months, number of sexual partners in past 3 months, relationship to last sexual partner was neither casual acquaintance nor commercial sex worker interacted with recent sex, used any type of contraceptives during last sex interacted with recent sex and used condom during last sex, interacted with recent sex.

(5) Contains: Never been unfaithful in past 3 months and number of sexual partners in past 3 months.

(6) Contains: Relationship to last sexual partner was neither casual acquaintance nor commercial sex worker interacted with recent sex, used any type of contraceptives during last sex interacted with recent sex and used condom during last sex, interacted with recent sex.

(7) Contains: Ever had an STD, ever had a bad smelling discharge, ever had genital sore or ulcer, current (or partner's) pregnancy is unwanted and currently pregnant (or partner).

(8) Contains: Ever tested for HIV, ever sought treatment for an STD and ever visited health center / worker for an issue related to SRH.

(10) Contains variables on perceived risk taking relative to other people and perceived own HIV risk.

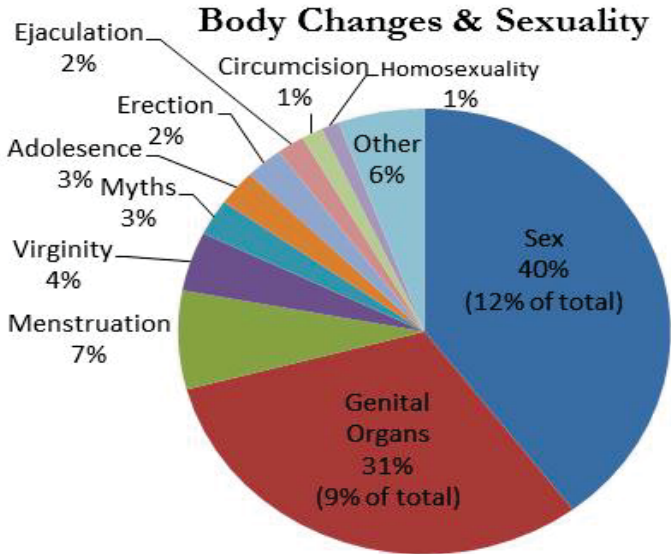
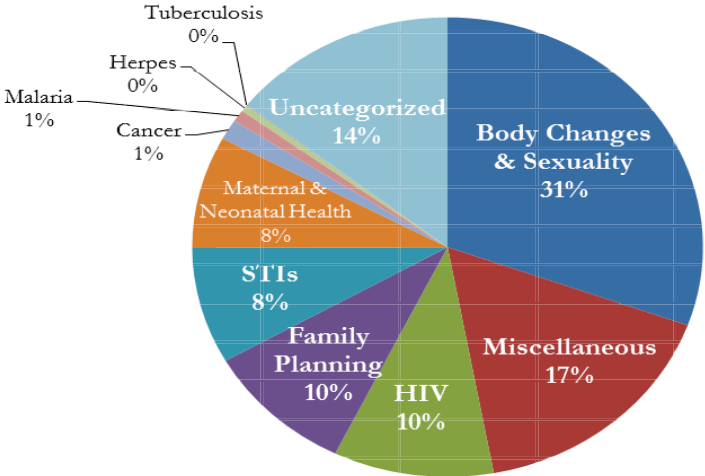
Table 3: Difference in Difference between Direct and Indirect Responses

	NO CONDOM AT LAST SEX				UNFAITHFUL				NUMBER OF PARTNERS			
	Direct proportion	Indirect proportion	Diff: Direct-Indirect	Diff in Diff	Direct proportion	Indirect proportion	Diff: Direct-Indirect	Diff in Diff	Direct proportion	Indirect proportion	Diff: Direct-Indirect	Diff in Diff
Overall												
Treatment	0.649 (0.02)	0.681 (0.04)	-0.032 (0.05)	-0.063 (0.06)	0.129 (0.01)	0.268 (0.05)	-0.139*** (0.04)	-0.157 *** (0.06)	1.195 (0.04)	0.854 (0.08)	0.341*** (0.09)	-0.171 (0.12)
Control	0.667 (0.02)	0.635 (0.04)	0.031 (0.04)		0.137 (0.01)	0.119 (0.05)	0.018 (0.05)		1.131 (0.03)	0.619 (0.08)	0.512*** (0.09)	
Diff: T-C	-0.018 (0.03)	0.046 (0.06)	0.063* (0.05)		-0.008 (0.02)	0.149** (0.06)	-0.157*** (0.05)		0.064 (0.05)	0.235** (0.12)	-0.171* (0.09)	
N	1,220	1,176	2,396		1,242	1,123	2,365		1,241	1,154	2,395	
Male												
Treatment	0.578 (0.03)	0.715 (0.06)	-0.137** (0.07)	-0.146 ** (0.08)	0.197 (0.02)	0.270 (0.07)	-0.073 (0.06)	-0.091 (0.09)	1.318 (0.07)	1.037 (0.12)	0.281* (0.14)	-0.263 (0.18)
Control	0.586 (0.03)	0.576 (0.06)	0.009 (0.06)		0.196 (0.02)	0.178 (0.07)	0.018 (0.07)		1.172 (0.05)	0.628 (0.13)	0.544*** (0.14)	
Diff: T-C	-0.008 (0.04)	0.138** (0.08)	-0.146** (0.07)		0.001 (0.03)	0.092 (0.09)	-0.091 (0.07)		0.146* (0.09)	0.409** (0.18)	-0.263* (0.14)	
N	581	573	1,154		610	524	1,134		608	544	1,152	
Female												
Treatment	0.711 (0.03)	0.649 (0.06)	0.062 (0.06)	0.007 (0.08)	0.061 (0.01)	0.266 (0.06)	-0.205*** (0.06)	-0.219 ** (0.09)	1.076 (0.05)	0.691 (0.10)	0.385*** (0.11)	-0.096 (0.15)
Control	0.743 (0.02)	0.688 (0.06)	0.055 (0.06)		0.081 (0.02)	0.067 (0.06)	0.014 (0.06)		1.092 (0.05)	0.611 (0.10)	0.481*** (0.11)	
Diff: T-C	-0.032 (0.04)	-0.039 (0.08)	0.007 (0.06)		-0.020 (0.02)	0.199** (0.09)	-0.219*** (0.06)		-0.016 (0.06)	0.080 (0.15)	-0.096 (0.11)	
N	639	603	1,242		632	599	1,231		633	610	1,243	

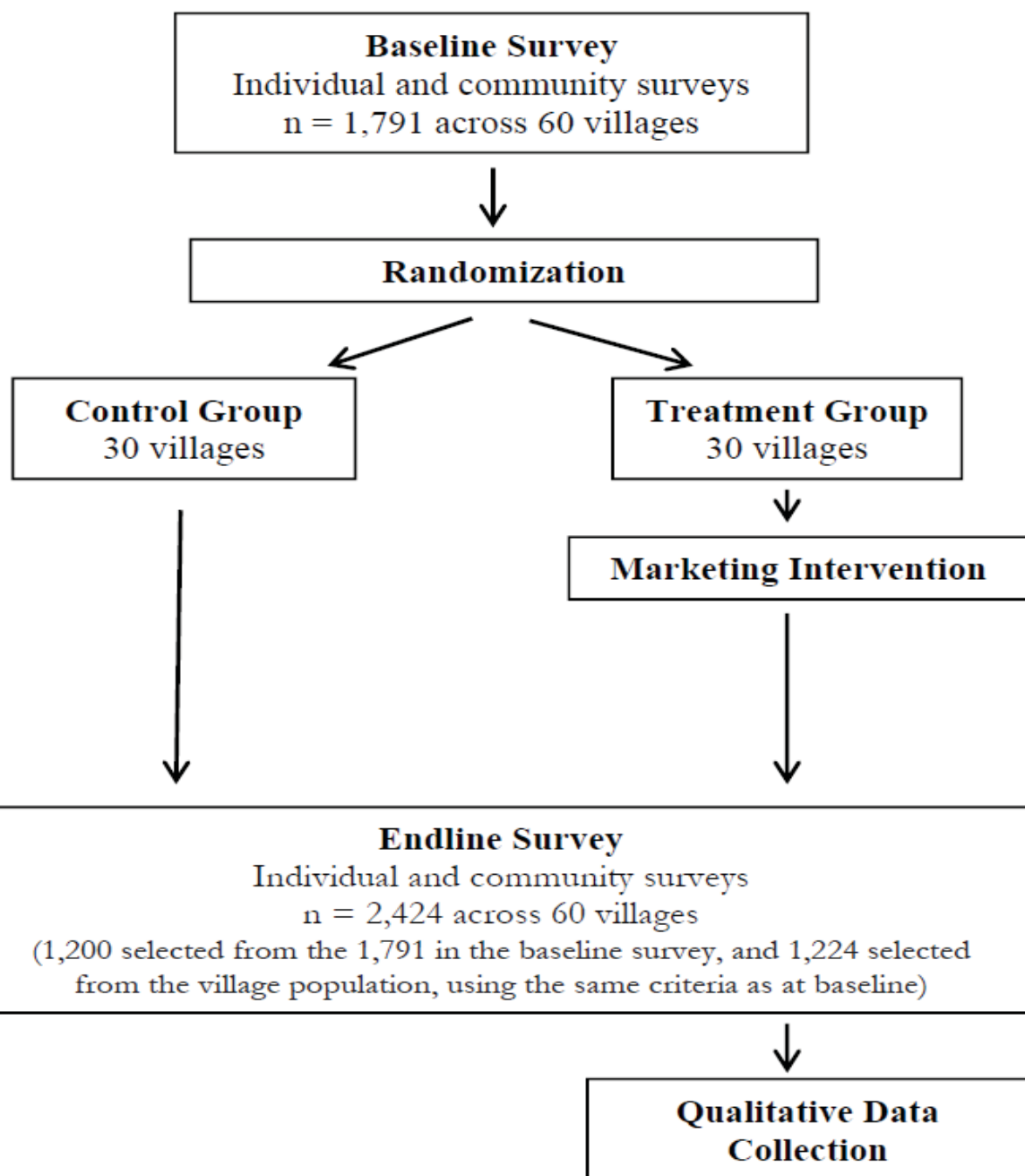
Simple difference in means comparison. The third column compares respondents who were directly or indirectly (using ListIt methods) asked, within a treatment group. The third row compares respondents between treatment and control group, within an interview method. The fourth column presents the difference in difference, i.e. the difference between the two differences (Direct-Indirect and Treatment-Control). Standard errors are presented in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Figure 1: 6001 Hits by Topic According to Metadata, Nationwide

6001 Health Hits by Category



Appendix figure 2: Evaluation flow chart



Appendix Table 1: Additional Baseline Values and Balance Tests

	Treatment		Control		p-value
	N	Mean	N	Mean	T vs C
Demographics					
Married or Cohabiting (%)	891	0.59	895	0.57	0.41
Has partner, but not cohabiting (%)	894	0.18	897	0.17	0.78
Low perceived relative HIV risk	848	0.49	858	0.49	0.81
Can read without difficulty (%)	893	0.84	894	0.81	0.22
Total number of mobile phones in household	886	1.51	883	1.51	0.84
Some difficulty charging in trading center (%)	886	0.33	887	0.31	0.28
Community characteristics					
Total number of households	30	381.50	30	370.30	0.89
Remote	30	0.20	30	0.27	0.55
Rural-Urban scale (1-5)	28	3.04	25	2.84	0.49
Average percentage of Muslims	30	0.18	30	0.20	0.63
Distance to next periodic market (km)	30	2.10	30	3.36	0.24
Outcomes					
Never had STD (%)	885	0.65	881	0.64	0.60
Partner or respondent are currently pregnant (%)	894	0.12	897	0.12	0.79
Partner or respondent are currently unwantedly pregnant (%)	894	0.03	897	0.03	0.58
Has ever tested for HIV (%)	894	0.57	897	0.59	0.36
Any ANC visits if currently pregnant, females only (%)	357	0.20	372	0.17	0.25
Number of ANC visits if currently pregnant, females only	327	0.78	357	0.69	0.50

Stratification variables include number of households, distance to tarmac road, distance to next periodic market, remoteness by category, urbanity by category, MTN network coverage, average education level, average SRH knowledge level, average SRH behavior score and predominant religion.

Appendix Table 2b: Determinants of Use of 6001
Dependent Variable: Ever Used 6001 (Observed and Self-reported)

	Sample frame:	Full sample	Males only	Females only	Full Sample
	Coefficient reported below for:	Covariate	Covariate	Covariate	Covariate interacted with treatment
Male		0.0561*** [0.02]			-0.0186 [0.03]
Own HIV risk perceived as low at baseline		0.0419 [0.03]	0.0580 [0.04]	0.0305 [0.04]	-0.0487 [0.03]
Married or cohabiting at baseline		0.0502** [0.02]	0.0618* [0.04]	0.0348 [0.03]	0.0258 [0.05]
Age		-0.0038* [0.00]	-0.0070** [0.00]	-0.0011 [0.00]	-0.0019 [0.00]
Education		0.0069* [0.00]	0.0112** [0.01]	-0.0002 [0.01]	-0.0131* [0.01]
Number of siblings		0.0002 [0.00]	0.0006 [0.00]	-0.0002 [0.00]	0.0004 [0.00]
Father's education		0.0003 [0.00]	0.0075 [0.01]	-0.0041 [0.01]	-0.0046 [0.01]
Meals eaten per day		0.0244* [0.01]	0.0253 [0.02]	0.0158 [0.02]	0.0331 [0.02]
Condition of footwear		-0.0108* [0.01]	-0.0177 [0.01]	-0.0087 [0.01]	0.0151 [0.01]
Ability to read		-0.0156 [0.01]	-0.0275 [0.02]	-0.0049 [0.02]	0.0468 [0.03]
Owns personal phone		0.0656*** [0.02]	0.1874*** [0.03]	0.0089 [0.03]	-0.0983* [0.05]
Frequency of using text messages		0.0067 [0.01]	0.0071 [0.01]	0.0100 [0.01]	0.0413*** [0.01]
Remoteness (community)		-0.0196 [0.07]	-0.0478 [0.08]	-0.0016 [0.07]	0.0140 [0.06]
MTN coverage (community)		0.0254 [0.03]	0.0359 [0.03]	0.0153 [0.02]	0.0313* [0.02]
Average SRH knowledge at baseline (community)		0.0805 [0.11]	0.0743 [0.14]	0.0818 [0.09]	0.0974 [0.07]
Average SRH behavior at baseline (community)		-0.0981 [0.15]	-0.0837 [0.19]	-0.1152 [0.14]	-0.0409 [0.11]
Percentage of Muslims (community)		-0.2171 [0.18]	-0.0406 [0.19]	-0.3444** [0.17]	0.1281 [0.13]
Number of households (community)		-0.0000 [0.00]	-0.0000 [0.00]	-0.0000 [0.00]	-0.0001** [0.00]
Distance to next tarmac road (community)		0.0016 [0.00]	0.0027* [0.00]	0.0008 [0.00]	-0.0004 [0.00]
Distance to next periodic market (community)		-0.0081 [0.01]	-0.0104 [0.01]	-0.0054 [0.00]	0.0013 [0.01]
Observations		2,276	1,099	1,167	2,276

Notes: Probit regression reporting marginal effects, robust standard errors clustered at the village level are reported in brackets. * p<0.10, ** p<0.05, *** p<0.01.

Appendix Table 2a: 6001 Awareness and Usage

Question		Control	Treatment	n	P-value t-test
<i>Respondent has ever heard of Google SMS</i>	Yes	46%	57%	2410	0.00
<i>Respondent has ever heard of 6001 service</i>	Yes	24%	48%	2405	0.00
<i>Respondent has heard of 6001 /Google SMS through...</i>	Radio	14%	13%	2405	0.51
	Vantage	4%	27%	2405	0.00
	Friends	2%	2%	2405	0.40
	Family	0.1%	0.4%	2405	0.10
	Champion	0%	1%	2405	0.00
	Teacher	0%	0.3%	2405	0.04
	Other	4%	4%	2405	0.61
	Had not heard of it	76%	52%	2405	0.00
<i>Respondent has ever sent a message to 6001</i>	Self-reported	3%	22%	2424	0.00
	Self-reported & observed	7%	40%	2424	0.00
<i>Median # of messages sent to 6001 if ever used</i>	Self-reported & observed	3	4	574	
<i>Median # of different days on which service was used if ever used</i>	Observed	5	4	236	
<i>Median # of days between first and last use if ever used</i>	Observed	3	6	236	
<i>Median # of days between first and last use if used on 4 or more diff days</i>	Observed	47	35	236	
<i>Respondent has used 6001 in...</i>	English	66%	26%	303	0.00
<i>(conditional on having used 6001)</i>	Luganda	34%	68%	303	0.00
	Both	0%	5%	303	0.15
<i>Most frequently mentioned reasons for not using 6001 service</i>	Not interested	8%	12%	2424	0.00
<i>(conditional on having heard of 6001)</i>	Don't know how to text	2%	4%	2424	0.02
	Don't have phone	1%	2%	2424	0.02
	Afraid of fees	2%	1%	2424	0.16
	Know how to text, but instructions unclear	6%	6%	2424	0.40
	Don't speak English or Luganda	0%	0%	2424	0.32

Appendix Table 3: Components of Knowledge Indices

HIV Knowledge		Respondents knows HIV can be transmitted...			Respondents knows HIV cannot be transmitted...	
		...during pregnancy	...during delivery	...during breastfeeding	...through sharing food	...through mosquito bites
		(1)	(2)	(3)	(4)	(5)
Full sample from follow-up survey						
A. Overall						
	Full sample	0.0025	0.0094	0.0044	0.0328	0.0290
		[0.03]	[0.01]	[0.02]	[0.02]	[0.03]
	# of observations (# of individuals)	1,092	1,092	1,092	1,092	1,091
B. By Gender						
	Male	-0.0508	0.0150	0.0102	0.0458	0.0203
		[0.04]	[0.01]	[0.02]	[0.03]	[0.03]
	Female	0.0576	0.0036	-0.0016	0.0192	0.0379
		[0.04]	[0.01]	[0.02]	[0.02]	[0.05]
	F-test: Male = Female	3.00	0.58	0.11	0.62	0.09
	F-Test: p-value	0.09	0.45	0.74	0.43	0.77
	# of observations (# of individuals)	1,092	1,092	1,092	1,092	1,091

Note: All regressions include controls for baseline values of the dependent variable and for stratification variables. Robust standard errors clustered at the village level are reported in brackets. * p<0.10, ** p<0.05, *** p<0.01.

Appendix Table 4: Components of Contraceptive Knowledge Index

		Respondents knows...									
		... the pill	...injections	...female condoms	...foam	...IUDs	...implants	...male condoms	...emergency pill	...female sterilization	...male sterilization
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Full sample from follow-up survey											
A. Overall											
	Full sample	0.0045	0.0046	-0.0151	-0.0019	-0.0310	0.0231	-0.0161	-0.0059	-0.0089	-0.0300**
		[0.01]	[0.03]	[0.02]	[0.01]	[0.03]	[0.03]	[0.03]	[0.01]	[0.02]	[0.01]
	# of observations (# of individuals)	1,092	1,092	1,081	1,080	1,084	1,082	1,090	1,082	1,082	1,081
B. By Gender											
	Male	-0.0036	-0.0094	-0.0000	0.0077	0.0112	0.0228	-0.0183	-0.0042	-0.0219	-0.0317
		[0.02]	[0.04]	[0.03]	[0.01]	[0.04]	[0.03]	[0.04]	[0.01]	[0.02]	[0.02]
	Female	0.0128	0.0192	-0.0308	-0.0120*	-0.0746*	0.0234	-0.0137	-0.0076	0.0045	-0.0282
		[0.02]	[0.03]	[0.03]	[0.01]	[0.04]	[0.05]	[0.04]	[0.01]	[0.03]	[0.02]
	F-test: Male = Female	0.36	0.47	0.66	5.11	2.13	0.00	0.01	0.05	0.49	0.02
	F-Test: p-value	0.55	0.50	0.42	0.03	0.15	0.99	0.92	0.82	0.49	0.89
	# of observations (# of individuals)	1,092	1,092	1,081	1,080	1,084	1,082	1,090	1,082	1,082	1,081

Note: All regressions include controls for baseline values of the dependent variable and for stratification variables. Robust standard errors clustered at the village level are reported in brackets.

Appendix Table 5: Components of Attitudes Indices

		Perceived rel. non-riskiness		Condom use attitudes	
		Relative to others, I do <i>not</i> take risks in my life	Perceived own HIV risk compared to a typical person (1 larger, 3 smaller)	Respondent agrees to the statement that....	
				A male condom should always be put on before intercourse	It is <i>not</i> embarrassing to buy or ask for a condom
		(1)	(2)	(3)	(4)
Full sample from follow-up survey					
A. Overall					
	Full sample	-0.0933**	-0.0225	0.0186*	-0.0240
		[0.04]	[0.04]	[0.01]	[0.02]
	Observations	2,267	2,209	2,206	2,258
B. By Gender					
	Male	-0.1621***	-0.0247	0.0314**	-0.0173
		[0.05]	[0.05]	[0.01]	[0.02]
	Female	-0.0278	-0.0203	0.0061	-0.0305
		[0.06]	[0.06]	[0.01]	[0.02]
	F-test: Male = Female	2.99	0.00	2.06	0.22
	F-Test: p-value	0.09	0.95	0.16	0.64
	Observations	2,267	2,209	2,206	2,258

Note: The higher the value of the dependent variable in (1), (2) and (3), the more comfortable disagreeing. All regressions include controls for baseline values of the dependent variable and for stratification variables. Robust standard errors clustered at the village level are reported in brackets.

* p<0.10, ** p<0.05, *** p<0.01.

Appendix Table 6a: Components of Behavior Indices and Related Variables

		Ever had sex with 2 or more different partners in same week during past 3 months (negative sign)	Never been unfaithful in past 3 months (No control on baseline values)	Number of sexual partners in past 3 months (No control on baseline values) (negative sign, i.e. a negative effect means more partners)	Relationship to last sexual partner was <i>neither</i> casual acquaintance nor commercial sex worker, interacted with recent sex	Used any type of contraceptives during last sex, interacted with recent sex	Used condom during last sex, interacted with recent sex (No control on baseline values)	Currently married or cohabiting^	Currently has partner, but not married/cohabiting^
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Full sample from follow-up survey									
A. Overall									
	Full sample	-0.0195	-0.0658*	-0.1345*	-0.0053	-0.0985*	-0.0232	0.0064	-0.0338**
		[0.02]	[0.03]	[0.07]	[0.466]	[0.05]	[0.03]	[0.01]	[0.01]
	Observations	998	2,199	2,247	2,270	461	2,218	2,270	2,275
B. By Gender									
	Male	-0.0331	-0.0527	-0.2827***	-0.0075	-0.0587	-0.0511	0.0231	-0.0480**
		[0.03]	[0.05]	[0.10]	[0.012]	[0.07]	[0.04]	[0.02]	[0.02]
	Female	-0.0062	-0.0781	0.0059	-0.0033	-0.1336**	0.0030	-0.0095	-0.0202
		[0.02]	[0.05]	[0.09]	[0.008]	[0.07]	[0.04]	[0.02]	[0.01]
	F-test: Male = Female	0.60	0.14	5.28	0.09	0.61	0.75	1.96	1.34
	F-Test: p-value	0.44	0.71	0.03	0.76	0.44	0.39	0.17	0.25
	Observations	998	2,199	2,247	2,270	461	2,218	2,270	2,275

Note: All regressions include controls for baseline values of the dependent variable and for stratification variables. Robust standard errors clustered at the village level are reported in brackets.

^ not an index component. * p<0.10, ** p<0.05, *** p<0.01.

Appendix Table 6b: Components of Outcome Index

		Ever had STD	Ever had bad smelling discharge	Ever had genital sore or ulcer	Current (or partner's) pregnancy is unwanted	Currently pregnant (or partner)
		(1)	(2)	(3)	(4)	(5)
Full sample from follow-up survey						
A. Overall						
	Full sample	0.0076	-0.0086	0.0366**	0.0085	0.0051
		[0.02]	[0.01]	[0.02]	[0.01]	[0.01]
	Observations	2,228	2,267	2,259	1,093	1,093
B. By Gender						
	Male	-0.0073	0.0010	0.0183	0.0219	0.0055
		[0.03]	[0.02]	[0.02]	[0.01]	[0.02]
	Female	0.0218	-0.0178	0.0540**	-0.0053	0.0048
		[0.03]	[0.02]	[0.03]	[0.02]	[0.02]
	F-test: Male = Female	0.60	0.40	0.96	1.09	0.00
	F-Test: p-value	0.44	0.53	0.33	0.30	0.99
	Observations	2,228	2,267	2,259	1,093	1,093

Note: All regressions include controls for baseline values of the dependent variable and for stratification variables. Robust standard errors clustered at the village level are reported in brackets. * p<0.10, ** p<0.05, *** p<0.01.

Appendix Table 6c: Components of Treatment Seeking Index and Related Variables

		Ever tested for HIV	Ever sought treatment for an STD	Ever visited health center / worker for issue related to SRH	Ever had STD (symptoms) and did not seek treatment	Any ANC visit during current pregnancy (females only)	Number of ANC visits during current pregnancy (females only)
		(1)	(2)	(3)	(4)	(5)	(6)
Full sample from follow-up survey							
A. Overall							
	Full sample	-0.0136	-0.0202	0.0122	-0.0048	-0.0330	-0.0204
		[0.02]	[0.01]	[0.02]	[0.01]	[0.031]	[0.106]
	Observations	2,275	2,275	2,265	2,275	1,090	863
B. By Gender							
	Male	-0.0322	0.0009	-0.0063	0.0066		
		[0.03]	[0.02]	[0.04]	[0.02]		
	Female	0.0042	-0.0403*	0.0298	-0.0157		
		[0.03]	[0.02]	[0.02]	[0.01]		
	F-test: Male = Femal	1.13	1.26	0.86	1.24		
	F-Test: p-value	0.29	0.27	0.36	0.27		
	Observations	2,275	2,275	2,265	2,275		

Note: All regressions include controls for baseline values of the dependent variable and for stratification variables. Robust standard errors clustered at the village level are reported in brackets. * p<0.10, ** p<0.05, *** p<0.01.

Appendix Table 7: Information on Health Services

		Knows place to get an HIV test	Knows place to get free HIV test	Knows place to get STD test	Knows place to get free STD test
		(1)	(2)	(3)	(4)
Full sample from follow-up survey					
A. Overall					
	Full sample	-0.0076	0.0192	-0.0014	-0.0343
		[0.01]	[0.03]	[0.02]	[0.03]
	Observations	1,090	1,093	1,091	1,093
B. By Gender					
	Male	-0.0104	0.0423	0.0059	0.0081
		[0.01]	[0.04]	[0.02]	[0.04]
	Female	-0.0047	-0.0047	-0.0091	-0.0786*
		[0.01]	[0.04]	[0.02]	[0.04]
	F-test: Male = Female	0.10	0.85	0.19	2.92
	F-Test: p-value	0.76	0.36	0.67	0.09
	Observations	1,090	1,093	1,091	1,093

Note: All regressions include controls for baseline values of the dependent variable and for stratification variables. Robust standard errors clustered at the village level are reported in brackets. * p<0.10, ** p<0.05, *** p<0.01.