NBER WORKING PAPER SERIES

THE ENTERTAINING WAY TO BEHAVIORAL CHANGE: FIGHTING HIV WITH MTV

Abhijit Banerjee
Eliana La Ferrara
Victor H. Orozco-Olvera

Working Paper 26096
http://www.nber.org/papers/w26096

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2019

This research is part of the entertainment-education program of the World Bank's Development Impact Evaluation Department (DIME), part of the Development Economics Research Group. We thank Oriana Bandiera, Michael Callen, Ruixue Jia, Gaia Narciso, Ricardo Perez-Truglia, Devesh Rustagi and seminar participants at IIES, LSE, Toulouse, Trinity College Dublin, University of Bonn, UCLA, UC San Diego, University of Manheim, University of Oslo, University of Amsterdam, University of Southern California, Yale, UPF, PSE, Wharton and BREAD 2018 conference for helpful comments. Laura Costica and Edwin Ikuhoria did a superb job as research and field coordinators. Tommaso Coen, Viola Corradini, Dante Donati, Francesco Loiacono, Awa Ambra Seck, Sara Spaziani and Silvia Barbareschi provided excellent research assistance. This study was funded by the Bill and Melinda Gates Foundation and the World Bank i2i Trust Fund. La Ferrara acknowledges financial support from ERC Advanced Grant ASNODEV. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Entertaining Way to Behavioral Change: Fighting HIV with MTV
Abhijit Banerjee, Eliana La Ferrara, and Victor H. Orozco-Olvera
NBER Working Paper No. 26096
July 2019
JEL No. D91,I12,O15

ABSTRACT

We test the effectiveness of an entertainment education TV series, MTV Shuga, aimed at providing information and changing attitudes and behaviors related to HIV/AIDS. Using a simple model we show that "edutainment" can work through an individual or a social channel. We conducted a randomized controlled trial in urban Nigeria where young viewers were exposed to MTV Shuga or to a placebo TV series. Among those exposed to MTV Shuga, we created additional variation in the social messages they received and in the people with whom they watched the show. We find significant improvements in knowledge and attitudes towards HIV and risky sexual behavior. Treated subjects are twice as likely to get tested for HIV eight months after the intervention. We also find reductions in STDs among women. These effects are stronger for viewers who report being more involved with the narrative, consistent with the psychological underpinnings of edutainment. Our experimental manipulations of the social norm component did not produce significantly different results from the main treatment. The individual effect of edutainment thus seems to have prevailed in the context of our study.

Abhijit Banerjee
Department of Economics, E52-540
MIT
50 Memorial Drive
Cambridge, MA 02142
and NBER
banerjee@mit.edu

Victor H. Orozco-Olvera
World Bank
Development Impact Evaluation Department (DIME)
1818 H Street
Washington, DC 20433
vorozco@worldbank.org

Eliana La Ferrara
Universita' Bocconi
Dept. of Economics and IGIER
via Roentgen 1
20136 Milano
Italy
eliana.laferrara@unibocconi.it

A randomized controlled trials registry entry is available at AEARCTR0000509

1 Introduction

International organizations and governments in developing countries have invested massively in behavior change campaigns, to improve outcomes in such diverse areas as health, education, financial decision making and governance. The evidence on the effectiveness of such campaigns is mixed at best, especially when we consider campaigns targeting HIV prevention.[1] The general sense seems to be that complex psychological factors are at play and the usual means of public communication fail to touch deep seated preferences in this domain. The policy community is thus increasingly embracing alternative approaches that combine information provision with entertaining content, an agenda which has come to be known as *edutainment* (short for 'entertainment education'). Edutainment consists of media programs, usually radio, television or film, that aim to change attitudes and behaviors by getting the viewer immersed into an entertaining narrative where the educational messages are presented as an integral part of a bigger story. Despite the increased popularity of this approach among policymakers, there is little rigorous evidence on whether edutainment works and, if so, through which mechanisms.

There are two distinct reasons why edutainment might work where ordinary behavior change campaigns fail. Pending more detailed discussion later, one reason is that the appeal of the show makes the individual pay more attention to the message and reduces potential resistance to top-down advice (Bandura, 1976). This is what we will call the 'individual' effect. A second reason, which we refer to as the 'social' effect, comes in because the show portrays alternative lifestyles that viewers could take as a norm. If people conform to what others do or think, then the message on the screen, potentially coupled with the fact that TV shows get seen by a large number of people, could coordinate a shift in the social norm. Given the growing evidence on the importance of conformity (e.g., Bursztyn and Jensen, 2015; Bursztyn, Fujiwara and Pallais, 2017; Perez-Truglia, 2017), it is plausible that changing behaviors might require shifting the norm that everyone conforms to. While any form of public communication can have a similar coordination effect, this is particularly important for popular movies or TV shows, just by virtue of the sheer numbers of people who watch them.

At the same time, there are several reasons why edutainment may *not* work. One is that the educational content may be too 'diluted' in the narrative, and viewers fail to recognize and retain the relevant information. Another is that the fictional nature of the story may lead viewers to trust the quality of the information less than a more 'official' source. Whether edutainment works or not is thus an open question.

This paper proposes a framework for analyzing the impact of edutainment along the channels described above and reports on a field experiment that we carried out to test the impact

[1] See, among others, Krishnaratne et al. (2016). Padian et al. (2010) conducted a systematic review of interventions aimed at preventing sexual transmission of HIV found that only one in seven of these were effective.

of an edutainment TV series. The series was the third season of MTV *Shuga*, a popular serial produced by MTV Staying Alive Foundation to provide information on HIV and change attitudes and behavior relating to HIV and risky sexual behavior more generally. The experiment covered over 5000 young men and women in 80 urban and peri-urban locations in South West Nigeria: 54 locations were randomly assigned to screen MTV Shuga, while the remaining 26 screened a different serial that involved a similar demographic but made no connections to HIV.

The first question we ask is whether MTV Shuga had the desired effects. We find striking effects on knowledge about sources of transmission of HIV and its treatment, on attitudes towards HIV+ people and on a range of behavioral outcomes (both self-reported and objectively measured) eight months or more after the showing. The likelihood of testing for HIV, objectively measured through redemption of a voucher that we distributed at health camps, increased by 3.1 percentage points in the treatment compared to the control group. This corresponds to a 100 percent increase over the control group mean. Analogous effects are estimated for the self-reported measure, where the likelihood of testing increases by 2.5 percentage points. Corresponding to this effect is an improvement in treated individuals' knowledge about HIV, including sources of transmission, awareness of anti-retroviral drugs, and the need to take a second HIV test after at least three months from the first (window period). These are topics specifically covered in MTV Shuga. The effects we estimate are robust to aggregating outcomes into indexes and to correcting for multiple hypothesis testing.

We find more nuanced effects on risky sexual behavior. On the one hand, the acceptability and reported incidence of concurrent sexual partnerships significantly decreased. On the other hand, MTV Shuga did not induce greater condom use, neither as reported by respondents nor as revealed in an experimental game that our subjects played in health camps. Despite the lack of effect on condom use, we do find significant impacts on a biomarker that proxies for unprotected sex with risky partners. The likelihood of testing positive for Chlamydia, a common STD, decreased by 55 percent in response to treatment for women in our sample (the impact on men is in the same direction but statistically insignificant). This is consistent with the reduction in the number of concurrent partners, and possibly with a more general shift away from risky behaviors.

MTV Shuga worked. The next question is why. To explore the importance of emotional involvement in the narrative, we make use of measures of viewer responses from the communication literature (Green and Brock, 2000; Murphy et al., 2011) that, to the best of our knowledge, have not been used in the economics literature before. In our endline survey we asked a battery of questions on how immersed in the story the respondent was while watching the show, and how much he or she identified with the characters. Using these we construct two indexes, '*Transportation*' and '*Identification*', and we show that the treatment effects were stronger for viewers that had higher values of these indexes. While these interaction effects

need to be interpreted with caution since both Transportation and Identification are correlated with other individual characteristics, the results are robust to including the interaction between treatment and a large number of observable controls. The results support the view that edutainment needs to be absorbing in order to work.

The last main question that we seek to address has to do with the role of social effects. The views and behaviors portrayed in MTV Shuga could signal a new norm about how one should interact with HIV positive people, with sexual partners, etc. In addition, our viewers may have expected Shuga to soon be released and have wide viewership, and they may have internalized the shift in norms that would take place as a result. To examine this possibility we take several approaches.

First, in half of the locations where MTV Shuga was shown (randomly selected), viewers were shown statistics on the attitudes of others like them after viewing Shuga: we call this treatment arm T2. As our theoretical analysis makes clear, T2 should have different effects on those viewers who in the baseline already believed that others had the kind of attitudes that MTV Shuga implicitly endorses, compared to viewers who believed that most others had different views from those in Shuga. The latter group is 'surprised' in a more positive direction by the information (where positive is defined to mean 'in the direction endorsed by Shuga') and therefore should respond more strongly.[2] To the best of our knowledge, this strategy of announcing post-viewing beliefs to test for conformity effects is novel.[3] However, we find no evidence of the predicted heterogeneous treatment effect of T2; nor is the mean effect of T2 any different from that of the 'basic' treatment (T1).

A possible reason for the lack of a differential effect of T2 is that exposure to T1 may have already conveyed a precise enough signal about the norm, compared to which T2 adds no new information. If this were the case, we should find that the basic treatment T1 did not change individual priors regarding social norms in their community of origin. We instead find that it did, albeit not in an entirely robust way. However, we find that the observed change in individuals' own attitudes in response to Shuga was not mediated by the perceived change in norms in the way that Bayesian updating would have predicted. This suggests that the prevailing norms in the community are not the main driver of individual choices in our context.

To explore social effects further, our experiment included a third randomly assigned treatment (T3), cross-cut across T1 and T2. In T3 we offered our viewers extra tickets to allow them to bring up to two friends to the screening. This treatment is meant to address two

[2] This argument relies on the decision-maker being Bayesian. We discuss what could happen if this assumption fails in section 2.5.

[3] In subsequent work Bursztyn, Gonzalez and Yanagizawa-Drott (2018) estimate the effect of announcing others' beliefs on married men's attitudes towards female labor force participation in Saudi Arabia. As will be clear from the description of our experiment, we announce the beliefs of people who watched Shuga, while Bursztyn et al. announce pre-intervention beliefs.

distinct concerns. First, if one were concerned that T2 fails to identify a relevant reference group for viewers (we told them that the announcement video and statistics were from "people like them in nearby communities"), T3 by construction allows participants to bring members of their own social network. Second, in general it may be hard to change one's opinions alone, without knowing what one's friends will think. In particular, since MTV Shuga influences behaviors relating to risky sexual behavior and HIV, it could be that the response is larger when people who are potentially sexual partners attend together. We find no evidence that T3 had any differential effect on attitudes and behavior. On the other hand, there is some evidence that when the invited friend was of the opposite sex, and therefore in Nigerian society more likely to be a sexual partner, the viewer is better informed about HIV. This is consistent with the interpretation that participants who attended with potential partners may have discussed the issues more in depth after the viewing.

The differential role of friends of the opposite sex also emerges when we test for spillovers from viewers onto other friends who were *not* invited to the showing (we collected a list of friends at baseline). We find weakly positive spillovers on HIV knowledge on average, stronger when the friend and the treated participant are of the opposite sex. There is no evidence of positive spillovers on attitudes and behavior.

The last strategy we adopt to explore social effects is non-experimental. We collected a range of standard measures of conformity, adherence to norms and independent thinking drawn from the literature in psychology (Schwartz, 2012). We test whether the effect of MTV Shuga differs depending on baseline values of these traits and we find no robust patterns.

Taken together, the above evidence suggests that edutainment works by conveying information while at the same time entertaining viewers. It also suggests that coordination on social norms was not a big part of the effect of MTV Shuga. This could be because viewers do not care about social norms when it comes to very private decisions, like ones relating to HIV and risky sexual behavior. Such interpretation would be consistent with the evidence from Bursztyn and Jensen (2015) and Bursztyn, Fujiwara and Pallais (2017), which are all about public acts of conformity. Compared to existing studies, we consider outcomes that are more sensitive and less discussed in public. Nonetheless, the fact that edutainment can have large individual effects, even in the absence of coordination effects, is quite striking. The enormous potential reach in terms of number of viewers and the low marginal costs of distribution make edutainment communication tools potentially very valuable for development policy.

Our paper contributes to a growing literature on the effects of the media on socioeconomic outcomes (see DellaVigna and La Ferrara, 2015 and La Ferrara, 2016 for a review). Part of this literature exploits non-experimental variation to study the effects of commercially oriented TV programs (e.g., Jensen and Oster, 2009; Chong and La Ferrara, 2009; La Ferrara et al., 2012; Kearney and Levine, 2015, 2019). These evaluations typically use expansion of access to television over time as the main source of variation and do not focus on the mechanism

questions that we are able to investigate. Ours is the first RCT to look at the impact of edutainment on sexual behavior and HIV-related outcomes.

Banerjee, Barnhardt and Duflo (2015), Ravallion et al. (2015), Coville et al. (2014) and Berg and Zia (2017) are examples of RCTs of the effect of edutainment. These authors evaluate interventions to promote, respectively, the consumption of iron-fortified salt, knowledge about a public works program and the latter two on financial literacy. Compared to these studies, our goal is to affect outcomes that are more sensitive and less discussed in public. To the best of our knowledge, we are the first to experimentally evaluate an edutainment TV series designed to change behavior determined by deep seated preferences such as those pertaining to HIV and sexual behavior. Our study also differs because of the emphasis on trying to identify the underlying mechanisms.[4]

There are also a number of recent RCTs that randomize exposure to videos or short documentaries containing information on role models (e.g., Bernard et al., 2014; Bjorvatn et al., 2015). These are not, strictly speaking, examples of edutainment and also focus on changing aspirations, an outcome we do not investigate.

A different strand of literature to which our paper relates is that on social pressure and social image (e.g., DellaVigna et al., 2012; Bursztyn and Jensen, 2015; Bursztyn, Fujiwara and Pallais, 2017). While we share with these authors the interest in how individual beliefs and actions respond to the beliefs and actions of reference groups, we differ in one fundamental dimension. Participants in our experiment remain 'private' in the sense that their choices are not announced to others: the concern that their behavior may be observed and sanctioned by others may emerge in the long run but it is not a direct consequence of our experimental design. What our experiment randomly makes public is information on the beliefs and behaviors of others. The contributions cited above manipulate instead the observability of individual behavior to others and find important effects of 'social image'. We believe that the two approaches are strongly complementary and that much can be learnt in the future by enriching research designs to contemporaneously address all these dimensions. Another fruitful approach may be to endogenize the reference group and the emergence of the social norm as a result of individuals' past choices, as in Benabou and Tirole (2011).

The remainder of the paper is organized as follows. Section 2 presents a stylized model that guides our empirical analysis. Section 3 describes the experimental design and section 4 the empirical strategy and data. In sections 5 and 6 we present results on our basic treatment and on social effects, respectively. Section 7 contains robustness analysis and section 8 concludes.

[4] Our work is also related to that of Paluck (2009), and Paluck and Green (2009), who study the impact of a radio soap opera on post-war reconciliation in Rwanda. These studies report some effects on prescriptive norms but relatively weak impacts on behavior, which may be related to the limitations of radio relative to a more visual medium or to the persistency of the historical legacy of the Rwandan genocide.

2 A simple model of learning and conformity

In this section we present a simple model to highlight potential channels of influence of edutainment and guide our empirical specifications.

2.1 Model idea

Our model is a variant of the standard Bayesian model of learning with quadratic preferences. Agents have a preference for being right as well as a preference for conformity. We model the effect of our interventions as sources of additional signals about the right thing to do as well as possibly about the social norm. The idea is that Shuga generates engagement, and therefore signals that would have otherwise been ignored actually get through to the viewer.

2.2 Preferences

We assume that the respondents in our study want to maximize a utility that depends on three terms: the distance to some objectively correct choice y^*, the difference between the choice they make and their preferred point a and possibly also on the distance between the observed choice and the average choice in the peer population, Y. Formally at time t individual i chooses y_{it} to maximize

$$-E_{it}[\alpha(y_{it}-y^*)^2+\beta(y_{it}-Y_i)^2+(1-\alpha-\beta)(y_{it}-a_{it})^2] \tag{1}$$

where E_t is the expectation operator taken based on the information at time t. This tells us that

$$y_{it}=\alpha E_{it}[y^*]+\beta E_i[Y_i]+(1-\alpha-\beta)a_{it}.$$

Here α captures the importance of information about the 'truth', while β picks up the degree of conformity. The fact that Y_i is assumed not to change over time is based on the idea that while individual choices vary, it all averages out. Allowing Y_i to drift over time would not change anything essential.

The goal of the intervention was to raise the average choice of y_{it}. MTV Shuga was intended to promote a liberal and informed discourse on HIV and risky sexual behavior, so we are effectively assuming that a higher value of y_i represents a more liberal/informed position.

2.3 Information and decisions

In making this decision individual i starts from a prior on y^*, $s_{i0}\sim N(y^*,1/p_y)$ and a prior on Y_i, $r_{i0}\sim N(Y_i,1/p_Y)$, where p indicates the precision of the signal. Therefore the baseline choice y_{i0} , in both treatment and control, is given by

$$y_{i0}=\alpha s_{i0}+\beta r_{i0}+(1-\alpha-\beta)a_{i0}.$$

We assume that a_{it} evolves as a AR(1) but may be shifted up by exposure to Shuga:

$$a_{it} = \rho a_{it-1} + \tau \delta_{T(i)} + \eta_t$$

where $\delta_{T(i)} = 1$ if $T(i) = 1$ (i is treated) and zero otherwise, $\tau \geq 0$ and η_t is distributed as $N(0, 1/p_\eta)$.

We assume that in control the prior signals are all the information that individuals get. Therefore

$$y_{it}^C = \alpha s_{i0} + \beta r_{i0} + (1 - \alpha - \beta) a_{it}.$$

Obviously

$$y_{i1}^C = \alpha s_{i0} + \beta r_{i0} + (1 - \alpha - \beta)\rho a_{i0} + (1 - \alpha - \beta)\eta_1$$

$$= (1 - \rho)(\alpha s_{i0} + \beta r_{i0}) + \rho y_{i0} + (1 - \alpha - \beta)\eta_1. \quad (2)$$

In the different treatments each individual potentially gets a signal S_i about y^*, $S_i \sim N(y^*, 1/p_S)$ as well as a signal about $\bar{Y}_i$, $R_i \sim N(\bar{Y}_i, 1/p_R)$. The two signals are drawn independently, though this assumption is easily relaxed at the cost of some additional notation. We impose no assumptions about correlations of signals across individuals. The updated choice based on the new information is:

$$y_{i1}^T = \alpha \frac{p_y s_{i0} + p_S S_i}{p_y + p_S} + \beta \frac{p_Y r_{i0} + p_R R_i}{p_Y + p_R} + (1 - \alpha - \beta)\rho a_{i0} + (1 - \alpha - \beta)(\tau + \eta_1).$$

We can rewrite this as

$$y_{i1}^T = \alpha \frac{p_S(S_i - s_{i0})}{p_y + p_S} + \beta \frac{p_R(R_i - r_{i0})}{p_Y + p_R} + (1 - \rho)(\alpha s_{i0} + \beta r_{i0}) + \rho y_{i0} + (1 - \alpha - \beta)(\tau + \eta_1). \quad (3)$$

2.4 Empirical approach

2.4.1 Treatment versus control

Differencing equation (2) from equation (3) and taking expectations conditional on y_{i0} and r_{i0} gives us

$$E[y_{i1}^T - y_{i1}^C | y_{i0}, r_{i0}] = \alpha \frac{p_S E[S_i]}{p_y + p_S} + \beta \frac{p_R E[R_i]}{p_Y + p_R} - \frac{p_S}{p_y + p_S} E[\alpha s_{i0} | y_{i0}, r_{i0}] -$$
$$- \beta \frac{p_R}{p_Y + p_R} r_{i0} + (1 - \alpha - \beta)\tau.$$

Note that we are assuming that r_{i0} is known to the econometrician, though in fact we only have proxies for it. This does not make a big difference since we will use the proxy when we

actually estimate the relationship. S_i and R_i are also only in the mind of the decision-maker and will have to be proxied for by treatment dummies. However one can use the fact that

$$y_{i0} - \beta r_{i0} = \alpha s_{i0} + (1 - \alpha - \beta) a_{i0}$$

to come up with the econometrician's expectation of αs_{i0},

$$E[\alpha s_{i0}] = \frac{\Gamma \widetilde{P}_a}{\Gamma \widetilde{P}_a + \widetilde{P}_y}(y_{i0} - \beta r_{i0}) + \frac{\widetilde{P}_y}{\Gamma \widetilde{P}_a + \widetilde{P}_y}\widetilde{y} - \frac{\widetilde{P}_a}{\Gamma \widetilde{P}_a + \widetilde{P}_y}\widetilde{a},$$

under the assumption that the econometrician assigns a distribution $N(\widetilde{y}, 1/\widetilde{P}_y)$ to s_{i0} and a distribution $N(\widetilde{a}, \widetilde{P}_a)$ to a_{i0} and defining $\Gamma = (1 - \alpha - \beta)^{-2}$. Plugging this into the above gives us

$$\begin{aligned} E[y_{i1}^T - y_{i1}^C | y_{i0}, r_{i0}] = \alpha \frac{p_S E[S_i]}{p_y + p_S} + \beta \frac{p_R E[R_i]}{p_Y + p_R} - \frac{p_S}{p_y + p_S}[\frac{\Gamma \widetilde{P}_a}{\Gamma \widetilde{P}_a + \widetilde{P}_y}(y_{i0} - \beta r_{i0}) + \\ + \frac{\widetilde{P}_y}{\Gamma \widetilde{P}_a + \widetilde{P}_y}\widetilde{y} - \frac{\Gamma \widetilde{P}_a}{\Gamma \widetilde{P}_a + \widetilde{P}_y}\widetilde{a}] - \beta \frac{p_R}{p_Y + p_R} r_{i0} + (1 - \alpha - \beta)\tau, \end{aligned}$$

which is more conveniently rewritten as

$$\begin{aligned} E[y_{i1}^T - y_{i1}^C | y_{i0}, r_{i0}] = \{\alpha \frac{p_S E[S_i]}{p_y + p_S} + \beta \frac{p_R E[R_i]}{p_Y + p_R} + (1 - \alpha - \beta)\tau - \frac{p_S}{p_y + p_S}[\frac{\widetilde{P}_y}{\Gamma \widetilde{P}_a + \widetilde{P}_y}\widetilde{y} - \\ - \frac{\Gamma \widetilde{P}_a}{\Gamma \widetilde{P}_a + \widetilde{P}_y}\widetilde{a}]\} - \frac{p_S}{p_y + p_S}\frac{\Gamma \widetilde{P}_a}{\Gamma \widetilde{P}_a + \widetilde{P}_y} y_{i0} - \beta[\frac{p_R}{p_Y + p_R} - \frac{p_S}{p_y + p_S}\frac{\Gamma \widetilde{P}_a}{\Gamma \widetilde{P}_a + \widetilde{P}_y}] r_{i0}. \end{aligned} \tag{4}$$

It is plausible that Treatment 1 provides both a signal about the state of the world (S_i) and one about the social norm (R_i). This suggests an estimating equation:

$$y_{i1} = \pi + \psi \delta_{T1}(i) + \lambda y_{i0} + \varphi r_{i0} + \phi \delta_{T1}(i) y_{i0} + \kappa \delta_{T1}(i) r_{i0} + \varepsilon_i, \tag{5}$$

where $\delta_{T1}(i)$ is a dummy equal to 1 when the status of i is treatment 1 and zero if it is control, $\psi \delta_{T1}(i)$ represents the term in curly brackets, $\phi \delta_{T1}(i) y_{i0}$ picks up the term $\frac{p_S}{p_y + p_S}\frac{\Gamma P_a}{\Gamma P_a + P_y} y_{i0}$ and $\kappa \delta_{T1}(i) r_{i0}$ picks up the term $\beta[\frac{p_R}{p_Y + p_R} - \frac{p_S}{p_y + p_S}\frac{\Gamma \widetilde{P}_a}{\Gamma \widetilde{P}_a + \widetilde{P}_y}] r_{i0}$. We expect ψ to be positive if S_i and/or R_i and/or τ are positive enough; ϕ is clearly negative and κ can go either way.[5] A sufficient condition for $\kappa = 0$ is that $\beta = 0$, but it is not necessary.

For many of the outcomes in our survey we have y_{i0} but not r_{io}. In those cases we can take the expectation with respect y_{i0} alone to get an estimating equation:

[5] The reason why the sign of κ is ambiguous is that a higher r_{i0} directly increases y_{i1} in Treatment but, for any given y_{i0}, it also tells us that s_{i0} must be lower, and this depresses y_{i0} in Treatment.

$$y_{i1} = \kappa + \psi\delta_{T1}(i) + \lambda y_{i0} + \phi\delta_{T1}(i)y_{i0} + \varepsilon_i. \tag{6}$$

Finally we can also take unconditional expectations to get an estimating equation

$$y_{i1} = \kappa + \delta_{T1}(i) + \varepsilon_i. \tag{7}$$

2.4.2 Treatment 2 versus treatment 1

The difference between treatment 1 and treatment 2 was supposed to be that the signal on $\bar{Y}$, R_i,was more precise –formally, the precision of the signal should be $p'_R > p_R$. Pooling the observations from $T1$ and $T2$ into a single treatment category therefore does not change the estimating equation and we can use (5), (6) or (7) to estimate the overall impact of MTV Shuga (with and without the extra announcement).

As for the difference between $T1$ and $T2$, from equation (3) it is easy to see that

$$E[y_{i1}^{T2} - y_{i1}^{T1}|y_{i0}, r_{i0}] = \beta(\frac{p'_R}{p_Y + p'_R} - \frac{p_R}{p_Y + p_R})(E[R_i] - r_{i0}).$$

The term $(E[R_i] - r_{i0})$ represents the 'surprise' element of our treatment, i.e., the difference between the signal on $\overline{Y}_i$ provided by Shuga and the individual's prior. This suggests an estimating equation

$$y_{i1} = \pi + \psi\delta_{T2}(i) + \lambda y_{i0} + \varphi r_{i0} + \kappa\delta_{T2}(i)r_{i0} + \varepsilon_i. \tag{8}$$

We would expect ψ to be positive and κ to be negative, unless $\beta = 0$ or $p'_R \simeq p_R$. Note that $y_{io}\delta_{T2}(i)$ does not enter this equation. This is because we have assumed that T2 provides no additional information about the correct choice y^*. In fact what is announced in T2 in our experiment is the mean opinion in a similar population after viewing Shuga. Since people know that others watched the same content as they did, they may assume that everyone else got the same signal on y^* as they did. Under this assumption, the announcement in T2 only contains new information on how people reacted to Shuga, thus yielding a more precise signal on $\bar{Y}$.

If we assumed instead that T2 viewers also received a more precise signal on the correct choice y^*, the estimating equation for the extra effect of T2 compared to T1 should also include an interaction term between y_{io} and the T2 dummy, that is

$$y_{i1} = \pi + \psi\delta_{T2}(i) + \lambda y_{i0} + \varphi r_{i0} + \rho\delta_{T2}(i)y_{i0} + \kappa\delta_{T2}(i)r_{i0} + \varepsilon_i. \tag{9}$$

In our empirical analysis, we consider both alternatives and estimate (8) as well as (9).

2.5 Discussion

The key implication of Bayesian learning that the above analysis makes clear is that what matters to the decision-maker is the surprise, the extent to which the signal that he or she gets from the world differs from his/her priors. This is why the interaction of i's priors with treatment enters the estimating equation for the marginal effect of T2 over and above T1 with a negative sign.

Something similar may also hold in non-Bayesian models of learning, though for a slightly different reason. Consider one obvious alternative Bayesian learning–an 'infection' model where the decision-maker, with some fixed probability, forgets his/her prior and adopts the signal he/she is exposed to. Otherwise he continues to hold onto the prior. This model would also generate a similar estimating equation: suppose treatment provides the decision-maker with an addition signal about the state of the world that he does not get in the control group, and T2 is more likely than T1 to provide him/her with a new signal about the social norm. Then the decision-maker's prior interacted with treatment will still enter the estimating equation with a negative sign, because treatment reduces the ex ante probability that he sticks to his prior.

The fact that we assume that Shuga provides a signal to everyone in treatment is a convenient simplification. Suppose instead we assumed that some viewers are engaged by the show but these viewers are all otherwise identical and being engaged is not correlated with their preferences or their prior information. This is equivalent to assuming that a fixed fraction of the treated population actually responds to the intervention. The rest behave as if they were in control, because they are not attentive to the message and effectively assign a precision of zero to it. Because of the linearity of the decision rule and the fact that we are taking averages across the population, this generalization ultimately does not change our estimating equation.

However it may be that those who are more likely to be attentive were also those who were more informed to start with (maybe because they had paid attention to previous messages) or more open to new ideas. In other words, there could be heterogeneity in, say, p_S and this could be correlated with s_{i0} and therefore y_{i0}. In that case the interaction between y_{i0} and treatment would pick up this extra responsiveness to the treatment among those who have a higher initial level of y_i. This positive effect would then counteract the negative interaction effect coming out of our model.

A large number of other assumptions have also gone into making the preceding analysis tractable and the notation less cumbersome than it would otherwise be. For one, as already noted, we assume that we have a proxy for r_{i0} even though in reality it is not observable to the econometrician. Specifically, in our empirical analysis we estimate r_{i0} using a survey question where i reports what is the number of people out of twenty in the community who have a particular view or take a particular action on an issue. This is potentially problematic for two

distinct reasons. First, there is no very good reason why our measure and not some correlate of it –like r_{i0}^2– is the right measure of the perceived norm. Second, from the point of view of conformity, the relevant peer group does not have to be the one that i is reporting about: he or she may be reporting about the objective reality of the local community (which is what our respondents were asked about) but when it comes to conforming, i may only care about his/her friends. We will try to deal with this issue in the empirical work by using alternative definitions of what i's reference group might be. Yet another approach would be to ignore all information about r_{i0} that we collected and use y_{i0} to pick it up (as we do for s_{i0}). This would be exactly like the case where we have no measure of r_{i0}, discussed above, where we estimate (6).

Finally, the simplicity of our estimating equations also derives from the fact that we assume quadratic preferences. Without this assumption how an individual reacts to new information will depend, for example, on how far he or she is from the views of the peer population.

3 Background and experimental design

3.1 Background

The edutainment product we evaluate is a TV series called *Shuga* and produced by MTV Staying Alive Foundation. MTV Shuga is a TV drama designed to raise awareness and change attitudes and behavior related to HIV/AIDS among young people in Africa. It presents young Africans from various socioeconomic strata balancing bright futures with the negative consequences of high-risk behaviors. The third season of Shuga, whose impact we evaluate, was filmed in Nigeria in 2013 and features prominent Nigerian actors and music, making it very appealing for the local public.[6]

In order to have exogenous variation in the exposure to the show, we conducted the study before MTV Shuga was widely distributed in Nigeria and we organized our own screenings in community centers, schools, and other locations that we could rent and that could accommodate about 100 individuals. The series consists of eight 22-minute episodes, which we screened in two blocks of four episodes, for a total duration of about 90 minutes per screening. For the control group we chose another TV series filmed in Nigeria, *Gidi Up*, which portrays a similar setting as MTV Shuga –urban and relatively upscale compared to the average population–but has no educational content. Also Gidi Up was screened in two blocks of comparable duration to Shuga. In all cases, screenings took place on Saturday or Sunday, and were one week apart.

The study sites were 80 urban and peri-urban locations chosen in 7 towns across three

[6] http://www.mtvshuga.com/show/?series=series-3

Note that series 1 and 2 were filmed in Kenya and season 3 was largely unknown during our experiment: at follow-up only 6 percent of respondents in our control group said they had watched Shuga on TV.

states of South-West Nigeria. Locations were defined by drawing a 2-mile radius around each screening center where the intervention was implemented, and randomly selecting households within this radius.[7] We ensured that there were buffer zones between communities to minimize the risk of contamination across study groups. These locations constitute our unit of randomization. Online Appendix A describes the selection of locations and respondents in detail, and Appendix Figure A1 shows the geographic distribution of treatment and control locations.

To identify study participants, we adopted a three-step recruitment strategy. First, enumerators visited a random selection of $200-225$ households per location. Second, the research team randomly selected one person aged 18 to 25 in each household to be invited to an initial film screening, which we denote as Screening 0. This film was different from MTV Shuga as the intent was to recruit amongst people available to attend film screenings during weekends, so as to reduce attrition. The selection was stratified by gender, half males and half females. Out of $17,224$ people invited to Screening 0, $6,613$ attended, with a turnout rate of 38.4 percent. Appendix Table A1 compares those who chose to attend Screening 0 and those who did not, on a number of socioeconomic characteristics. The two samples are well balanced; a few variables show significant differences, yet of extremely small magnitude: in all cases the normalized difference is below the threshold of 0.25 suggested by Imbens and Rubin (2015). We thus conclude that those who attended Screening 0 were a representative subset of the households invited.[8]

The third step was the selection of our baseline sample. In each location enumerators randomly selected 64 individuals among those who attended Screening 0 and visited them at home to administer the baseline survey. From now on, we refer to these 64 individuals as 'main study participants'.[9] At the end of the survey, the main study participants received invitations to attend two other screenings organized in the two weekends following the interview: Screenings 1 and 2. Those in the 'Friends treatment' (to be described in the next section) also received invitations to bring up to two friends of their choice. Note that invitees were not told in advance what they would watch, neither before Screening 0 nor before Screening 1.

Attendance to the screenings was relatively high: 78.4 percent of those invited attended at least one of the two screenings, with the shares being 77 percent in the treatment group and 81 in the control group (significantly different).[10] Note that in all our analysis we report

[7] The condition for a household to be in the study was that at least one of the members should be in the target age range of our intervention, i.e., 18-25 years old.

[8] We also compared our sample to the 2008 Nigeria DHS, restricted to the South-Western region, and found that the two were quite similar in terms of religion (a third Muslim and two thirds Christian), years of education (around 11 years) and television-ownership rates (around 90 percent).

[9] Appendix Table A2 shows balance in observable characteristics between the people who attended Screening 0 and were selected into the baseline sample and those that were not.

[10] 70.5 percent of our sample attended Screening 1, 57.2 attended Screening 2, and 49.4 percent attended both. The likelihood of attending Screening 2 conditional on attending Screening 1 is 70 percent in both treatment and control.

intention-to-treat effects, using the initial assignment as our treatment indicator.

In each location, all baseline survey and screening activities were concluded in four weeks. Implementation was rolled out so that activities in a given location were completed before moving to the next, in order to minimize attrition due to subjects forgetting about the screenings, travelling or relocating. Online Appendix B reports the timeline of activities.

3.2 Experimental design

Since individual-level randomization would run a significant risk of contamination, we implemented a clustered randomized trial where our study locations were randomly assigned to treatment and control groups. The experiment was designed to allow us to study the impact of MTV Shuga screenings alone as well as that of being exposed to Shuga plus information on beliefs and values of peers. We created different treatment arms and stratified the sample so that each town would have an equal number of locations in each arm (where possible).

Prior to the actual intervention, we piloted the Shuga screenings in some urban and peri-urban locations outside our sample frame. We used these pilots to shoot short videos with interviews of participants, and to administer short 'exit surveys' containing selected attitudinal questions.

T1. Treatment T1 consisted only of MTV Shuga screenings and was administered in 27 randomly selected locations. Participants were shown the Shuga TV drama in two screenings of four 22-minutes episodes each. We did not organize any discussion at the end of the screenings, to ensure uniformity of the treatment and to make the experience more comparable –albeit not perfectly– to that of a viewer watching TV at home.

T2. The second treatment (T2) involved another 27 randomly chosen locations and was the same as T1, except that after the MTV Shuga episodes we showed video-clips containing information on beliefs and values of peers in other communities who had watched Shuga. These video-clips were assembled using material from the pilot screenings and included interviews of youth condemning negative behaviors and praising positive ones after watching Shuga, as well as 'smart graphs' with statistics. The intent was to raise awareness about how common certain beliefs and attitudes are among other participants, and how willing they may be to change them. T2 thus embeds a first type of 'social referents', namely young people from other communities who look similar to our respondents and who watched MTV Shuga.

The choice of this way of implementing T2 represented a compromise between a number of different imperatives. Announcing the post-screening average –which is what we opted for– is the right thing to do if the population expects that the rest of their peer group will also see Shuga eventually and they want to be close to that post-Shuga consensus. This is a plausible case since MTV is a well-known brand and our viewers may have correctly expected that Shuga would be shown on television eventually. Consistent with this interpretation, below we show

that T1 participants updated their expectations on the prevailing values in their communities among other people who had not yet seen Shuga.

An alternative possibility for T2 would have been to announce the pre-Shuga average, which would have the advantage that it would be from the same population/community as the one being treated and not from a similar population from elsewhere. On the other hand, our study participants might have discarded this as dated information if they expected their peers to see MTV Shuga and change their mind. Moreover our partners and funders were very worried that we would undermine Shuga's message by reminding participants of their and others' pre-Shuga views. Finally, even if our viewers did not believe that the social norm would shift as a result of Shuga, and therefore wanted to get closer to the pre-Shuga norm, they could back this out of the post-Shuga mean that was announced. They could do so using their prior about the distribution of signals about the state of world provided by Shuga, the signal they themselves got from Shuga and their prior on the distribution of preferred points in the population. This, it is easy to show (available from the authors), will generate the same estimating equation as the one we estimate.

T3. To half of the treated individuals, randomly selected, we offered the option of bringing up to two friends to the screenings. The goal of this treatment was to determine whether the effect of Shuga differs when individuals can discuss its content with close peers who also watched the drama. This treatment was randomized at the individual level and cut across the other two treatments. We denote this as the 'Friends treatment' (T3). Compared to T2, T3 includes a different type of 'social referents': friends who are chosen by the individual and thus surely belong to his/her social network. Take-up of T3 was relatively low. Of the 1775 main study participants who received an invitation, only 830 brought at least one friend to at least one screening (47 percent take-up), and 483 brought two friends. In our analysis, we estimate an intention-to-treat effect and focus on an indicator for whether the respondent was given the option to bring a friend (T3).

It is worth discussing why we did not include a 'pure information treatment' in our experiment. A large body of evidence already exists on information campaigns to prevent HIV in low income countries (e.g., Fonner et al. 2014, McCoy et al. 2010), including Sub-Saharan Africa (Harrison et al. 2010, Michielsen et al. 2010). These studies consistently find limited or no impacts on reducing risky sexual behaviors.[11] In a review of systematic reviews, Krishnaratne et al. (2016) concluded that results for HIV behavior change interventions were overall "disappointing". The effective approaches seem to be medical ones or conditional cash transfers (e.g., Bjorkman Nyqvist et al. 2018), which are not based on simple information provision

[11] Duflo, Dupas and Kremer (2015) evaluate the impact of an HIV education program in Kenya in which primary school teachers were trained to deliver the national HIV/AIDS curriculum. They find that, when implemented alone, this program did not succeed in reducing STIs. It is only when combined with an education subsidy that the program was effective.

and may also be quite expensive to scale in a large country like Nigeria. Given budget and logistical constraints, we therefore took stock of the evidence on the lack of effectiveness of pure information campaigns and privileged an experimental design that would allow us to shed light on the channels for the effect of edutainment.

Summary of interventions. To sum up, we have three treatments: T1 and T2, randomized at the cluster level, and the 'Friends' treatment (T3), which is cross-cut across T1 and T2 and is randomized at the individual level. These treatments together cover 54 locations. The remaining 26 locations constitute our control group, where we screened the 'placebo' TV series *Gidi Up*.

We complemented our data collection on the main study participants with a sample of contacts from their social networks that we use for detecting potential spillovers. In all locations, before taking the baseline survey, study participants were asked to list two friends aged 18-25 to whom they regularly talked and who lived in the community. In each location, we administered the baseline and the follow-up surveys to a random sample of 15 contacts among those indicated by respondents who were *not* in the 'Friends' treatment. We refer to these individuals as 'network members' and we will conduct an impact analysis on them too. They should be distinguished from the friends that individuals brought in T3, on whom we have no information other than their gender.[12]

4 Empirical strategy and data

4.1 Empirical strategy

Benchmark specification. To estimate the average treatment effect, we use two specifications. One is the cross-sectional specification corresponding to equation (7) in the model:

$$y_{ilc1} = \beta T_{ilc0} + \mathbf{X}'_{ilc0}\zeta + \delta_c + \varepsilon_{ilc1}. \tag{10}$$

where y_{ilc1} is the outcome of interest for individual i who lives in location l within city c in the follow-up survey (time 1); T_{ilc0} is a dummy equal to 1 if the individual was assigned to treatments T1 or T2; $\mathbf{X}_{ilc0}$ is a vector of controls measured at baseline (time 0) that include age, years of education, and dummies for being enrolled in school, single, Muslim, speaking Yoruba as main language at home, speaking English as main or second language at home, not living with one's parents, household size, a wealth index, homeownership, and two dummies for father and mother having achieved more than secondary education; δ_c denotes town fixed effects. Appendix Tables A3 and A4 report variable definitions and summary statistics.

[12] By construction, contacts of those main respondents who had been given extra tickets for friends (and who could have therefore attended the screenings) are not part of the spillover sample.

The second specification includes the lag of the dependent variable and its interaction with treatment, as in model equation (6):

$$y_{ilc1} = \alpha y_{ilc0} + \beta T_{ilc0} + \gamma(T_{ilc0} * y_{ilc0}) + \mathbf{X}'_{ilc0}\zeta + \delta_c + \varepsilon_{ilc1} \tag{11}$$

where y_{ilc0} is the baseline value of the dependent variable, and the remaining controls are defined as above. In both specifications (10) and (11) we cluster standard errors at the location level, which is our unit of randomization (we have 80 locations).

Effect of announcement. Next we test whether the provision of information on how other people reacted to Shuga (T2) differentially affected viewers compared to the simple screening of the series (T1). In this case we restrict the sample to treated individuals and estimate a specification corresponding to equation (9) in the model:

$$\begin{aligned} y_{ilc1} = {} & \alpha y_{ilc0} + \beta T2_{ilc0} + \gamma(T2_{ilc0} * y_{ilc0}) + \\ & + \lambda(T2_{ilc0} * \widetilde{Y}_{ilc0}) + \mu\widetilde{Y}_{ilc0} + \mathbf{X}'_{ilc0}\zeta + \delta_c + \varepsilon_{ilc1} \end{aligned} \tag{12}$$

where $\widetilde{Y}_{ilc0}$ is individual i's prior (elicited at baseline) on the average realization of outcome y in the community. We also estimate a shorter version of (12) that corresponds to equation (8) in the model and that does not include the interaction between T2 and y_{i0}.

Heterogeneous effects. To shed light on the importance of the edutainment component and of conformism, we exploit individual-level measures that we elicited through our survey. We estimate an augmented version of equation (11) that includes an interaction term between T_{ilc0} and i's involvement with the plot (or i's conformism at baseline), plus of course the standalone variable.

Friends treatment. To test whether viewers who watched Shuga with a friend exhibited different responses, we estimate:

$$y_{ilc1} = \alpha y_{ilc0} + \beta_F Friend_{ilc0} + \mathbf{X}'_{ilc0}\zeta + \delta_c + \varepsilon_{ilc1} \tag{13}$$

where $Friend_{ilc0}$ is a dummy that takes value one for individuals in T3. This regression is estimated on the treatment sample only, because only treated participants received the friends invitation. If the possibility of talking about the show with a friend reinforced the message in Shuga, one would expect $\widehat{\beta}_F > 0$ for outcomes for which the main treatment effect is positive and $\widehat{\beta}_F < 0$ when the main treatment effect is negative. Of course, this need not necessarily be the case, depending on the friend's own preferences.

Spillovers. To estimate spillover effects, we use a different sample of respondents j who were referred to us by our study participants i. We use the notation ji to indicate that j is a member of i's network. We estimate:

$$y_{jilc1} = \alpha y_{jlc0} + \beta_S T_{ilc0} + \gamma_S(T_{ilc0} * y_{jlc0}) + \mathbf{X}'_{jlc0}\zeta + \delta_c + \varepsilon_{jilc1} \tag{14}$$

where the outcome y and the controls $\mathbf{X}$ refer to network member j, but exposure to Shuga is indirect, only through j's friend i. In the presence of spillovers, the estimated coefficient $\widehat{\beta}_S$ should have the same sign as $\widehat{\beta}$ in (11). In other words, if Shuga positively affected i's outcome and i talked about it with his/her friend j, then j's outcome would also respond positively (and vice versa).

We also estimate a variant of (14) that includes an interaction between T_{ilc0} and a dummy for whether i and j have opposite sex. The idea is that discussion around the issues covered in Shuga may be different among people of same vs. opposite sex, and that the latter may be a coarse proxy for whether the pair is potentially a couple. Some behaviors, e.g., condom use, may be easier to adopt if both partners have been exposed to Shuga.

Reporting results. Since we have a large number of outcome variables, we present results in two formats. First, to address the issue of multiple hypothesis testing, we group our original outcomes into indexes. This reduces the number of hypotheses actually tested and increases statistical power by reducing errors due to random variation at the level of the individual variables. We describe the construction of the indexes in section 4.2. Aggregation only partially solves the multiplicity problem, as we still have several hypotheses being tested jointly. To correct for this, we adjust p-values according to the free step-down resampling method (Westfall and Young, 1993) so that they can be used to control the family-wise error rate (FWER), defined as the probability of rejecting at least one true null hypothesis.

Our second way of reporting results focuses on individual outcomes that can be regarded as important on their own. These are selected within the broader list of variables from which we compute the indexes, and we single them out because they are key to the overall message of Shuga (e.g., "you should wear a condom"; "having concurrent partners can be risky"; etc.) or they capture specific messages that are strongly emphasized in certain episodes (e.g., "a young boy should be allowed to play football"; "you have to take a second HIV test after 3 months"; etc.). For our coefficients of interest, we report both 'naive' standard errors corrected for clustering at the location level, and FWER-adjusted p-values –that adjust for multiple hypothesis testing– based on 10,000 replications.

4.2 Data and descriptive statistics

Sample

In total we interviewed 5166 main study participants at baseline and 4986 at follow up.[13] Since our conditional specification includes the lag of the dependent variable, our working sample consists of the individuals for whom we have both rounds of data, i.e., 4986 observations.

[13] We performed power calculations using the Nigeria DHS 2008 and determined that a sample of 64 individuals per location, or 5120 individuals in total, half male and half female, would enable us to detect a change of between $0.15 - 0.20$ standard deviations in our main outcomes of interest. Updated power calculations using our baseline data showed an improved minimum detectable effect of $0.12 - 0.17$ standard deviations.

For some specifications we use smaller samples because of missing data for specific outcomes.

In Appendix Table A5 we regress the decision to participate in the follow up on the treatment dummy and on baseline values of our outcome indexes and of socioeconomic controls. As we can see, attrition is not selective on our main variables of interest.

Indexes

We group our outcome variables into two broad topics: those related to HIV and those related to risky sexual behavior. For each topic we elicit responses on knowledge, attitudes and behavior using several survey questions. The individual questions are then aggregated into indexes following Kling et al. (2007), i.e., we construct equally weighted averages of the z-scores of the variables that enter each index. For robustness, we also use a second method based on principal component analysis. Both methods are described in detail in Online Appendix C, and the list of variables entering each index (with their respective signs and loading factors) are shown in Appendix Tables A6 and A7. Variables are oriented so that the impact of treatment on each component of the index should be positive.

We construct five indexes: *HIV knowledge*, *HIV attitudes*, *HIV testing*, *Attitudes towards risky sex*, and *Risky sexual behavior*. These can be briefly described as follows. *HIV knowledge* measures how aware an individual is of the methods of transmission, the availability of drugs, and the timing of testing for HIV. Higher values of this index correspond to greater awareness. *HIV attitudes* captures respondents' inclination to potentially reveal their HIV+ status, allow HIV+ people to interact with the community, and not hold negative judgements towards them. A higher value of this index denotes more progressive attitudes (consistent with the message of MTV Shuga). The index *HIV testing* measures whether the respondent knows where to get tested, if he/she has been tested and when, if he/she picked up the results and if he/she asked for the test him/herself. Increasing values of the index *HIV testing* correspond to more active testing.

A second family of variables relates to risky sexual behavior. The index *Attitudes towards risky sex* includes opinions regarding multiple concurrent partners, dating sugar daddies, and attitudes towards women who bring a condom. Again, increasing values of this index correspond to attitudes more consistent with the message of Shuga. Finally, *Risky sexual behavior* captures whether the respondent has multiple concurrent sexual partners (and the number), condom use during the last intercourse, and having a main or additional sexual partners. This index is only available for the subset of individuals who are sexually active. Increasing values of the index correspond to *less* risky sexual behavior, so we should expect a positive treatment effect.

Health camps

In order to collect 'objective' outcome measures in addition to survey-based ones, we set up 'health camps' in 80 schools to which survey respondents were invited when they took the follow-up survey. At the health camp participants were informed about testing by counsellors

and were offered the opportunity to test for Chlamydia through urine sample collection. During the same session they also participated in a game that consisted in choosing between a fixed amount of money and a certain number of condom packs. At the end of the session participants received contact details of HIV counselling and testing centres in their town and were given a voucher that would entitle them to free HIV testing at one of these centers. Through the ID number on the voucher we know who took the HIV test, but we do not know the results.[14] A second health camp was set up two weeks after the first, where participants were informed of the results of the Chlamydia test and –if they tested positive– they were prescribed treatment. Online Appendix D contains further details on the health camps.

In our analysis we will use the following outcomes collected at health camps: (i) whether participants attended the health camp and took the Chlamydia test; (ii) whether they tested positive for Chlamydia; (iii) whether they redeemed the voucher to get tested for HIV; and (iv) whether they chose condoms over money when given the choice.

Attendance to health camps was relatively high: 77 percent of the study participants attended, and on average this share was the same in treatment and control locations. Appendix Table A8 shows how baseline characteristics and baseline values of our outcomes correlate with the decision to participate in the health camp. People currently attending school and living outside the family were less likely to attend, possibly due to conflict with school schedules. Treatment status is uncorrelated with the decision to attend, as are baseline values of our outcome indexes.

Expectations

Among the variables we collected through our survey, it is worth detailing how we elicited expectations regarding community attitudes, because these variables play an important role in our test for social effects. For the main attitudinal outcomes we elicited two types of responses. The first was the individual's own position, for example: "If you had HIV and you had a boyfriend/girlfriend, would you reveal your status to him/her?". This type of variable is used as dependent variable in our analysis.

The second category relates to the position of community members, for example: "If you picked 20 people of your age from your community who had a partner, how many would reveal their status to their partner if they had HIV?". From this type of question we construct the share of community members who would choose a certain action (or support a certain statement) and we employ this variable as the 'prior' about the average choice in the peer population (r_{i0} in the model, $\widetilde{Y}_{ilc0}$ in regression (12)). Due to constraints on the length of the questionnaire, we elicited these priors for some but not all of our outcomes of interest. For this reason, we will be able to estimate specifications (10) and (11) for all outcomes, and specification (12) for a subset of them.

Randomization check

[14] NACA (2015) estimates that in 2012 the South-West region had an adult HIV prevalence rate of 2.8 percent.

Table 1 shows summary statistics for our outcomes of interest (Panel A) and the control variables (Panel B) at baseline. We report the mean in the control and in the treatment group, the p-value for the test that the difference is zero, the normalized difference and the number of observations for each variable.[15]

Our outcome indexes are well balanced: for all five the difference in means is never statistically significant. When we look at individual outcomes, out of 22 variables 3 have p-values of .05 or less (although these p-values are not corrected for multiple hypothesis testing: the p-values would be much higher if we accounted for that). Even so, the normalized difference in means is extremely small, well below the cutoff of 0.25 suggested by Imbens and Rubin (2015).

Turning to control variables, Panel B of Table 1 shows that variables such as gender, age, education, religion, language spoken, having a partner and living outside the family are well balanced. We have some imbalance in household size, wealth and parents' education: on these variables the control group seems to be better off than the treatment. The normalized differences, however, are well below the 0.25 cutoff, so in terms of economic significance of the imbalance we do not find reasons for concern. Furthermore, we control for these variables in all our specifications.

In Appendix Table A9 we perform an alternative test, regressing our treatment dummy on covariates (Panel A) as well as on covariates and outcome indexes (Panel B). None of the regressors is significantly different from zero, except for homeownership and father's education. The F test for joint significance always yields p-values greater than 0.10.

5 Results: average treatment effects

In this section we report our estimates of the impact of Shuga on a variety of outcomes, starting from the average treatment effects and then exploring the role of involvement with the narrative.

5.1 HIV related outcomes

Table 2 reports the average treatment effects for outcomes related to HIV knowledge, attitudes and behavior. Panel A considers aggregate indexes as dependent variables, while Panel B

[15] The normalized difference is a scale-free measure of the difference in distributions, recommended by Imbens and Wooldridge (2009):

$$\Delta = \frac{\overline{X}_A - \overline{X}_B}{\sqrt{S_A^2 + S_B^2}}$$

where $\overline{X}_A$ and $\overline{X}_B$ are the means of covariate X in groups A and B, respectively, and S_A^2 and S_B^2 are the corresponding sample variances of X.

individual outcomes.[16] Columns 1, 3, 5 estimate the cross sectional model (10), while columns 2, 4, 6 include the lag of the dependent variable and its interaction with treatment, as in (11). Appendix Table A11 shows the estimates from a simple ANCOVA model. Individual level controls and city fixed effects are always included but not shown. Standard errors in parenthesis are adjusted for clustering at the screening center level. In square brackets we show p-values corrected for multiple hypothesis testing using FWER. At the bottom of the table we report the mean and the standard deviation of the dependent variable at follow-up in the control group, and the p-value for the test that the sum of the coefficients on $Treated$ and $Treated * Y_{t-1}$ is equal to zero when evaluated at the mean of Y_{t-1}.

The results in Table 2 clearly show that exposure to MTV Shuga significantly improved all HIV-related outcomes. In the top panel, the impact on respondents' knowledge about HIV is positive and significant at the 1 percent level with either specification and either method of correction for the standard errors. The magnitude of the effect in the conditional model (column 2) when evaluated at the mean of Y_{t-1} corresponds to .13 of a standard deviation of this index. Shuga also improved attitudes towards people with HIV. The effect on the aggregate index is again positive and significant, with an effect size of .10 of a standard deviation when evaluated at the mean (column 4). We detect positive and significant impacts also on the *HIV testing* index: based on the estimate in column 6, treatment induced an increase in the aggregate index of .08 of a standard deviation when evaluated at the mean.[17] As for the other coefficients in the table, the lagged dependent variable is always significantly correlated with current outcomes, while the sign on the interaction between treatment and the lagged dependent variables is negative (as predicted by the model) in two out of three cases, though typically insignificant. This may be an indication that there is actually heterogeneity in the response to the intervention along the lines we suggested above, which creates a countervailing effect.

In Panel B of Table 2 we consider three individual outcomes that are included in the indexes but are also of interest in and of themselves, because they are explicitly targeted in the messages of MTV Shuga. First of all, we compare the results on HIV testing obtained when using objective behavior from health camps (columns 1-2) and when relying on respondent's own reports (columns 3-4). Both variables show positive treatment effects. The self-reported probability of testing increases by 2.5 percentage points, over a mean of 8.6 percent in the control group. The impact is even larger when we consider actual testing measured by the redemption of vouchers received at health camps. In this case the probability of testing increases

[16] As a benchmark, we employ the aggregate indexes constructed following Kling et al. (2007). Appendix Table A10 reports estimates for indexes constructed using principal component analysis: the results are qualitatively unchanged.

[17] Appendix Table A12 reports separate estimates by gender. Impacts on knowledge and testing are stronger for women, while impacts on attitudes are more pronounced for men.

by 3.1 percentage points, which is a 100 percent increase over the control group mean.[18]

The remaining columns of Panel B consider outcomes that were explicitly addressed in MTV Shuga. In columns 4-5 we show that treatment improved respondents' knowledge about the need to take a second test and that this test should be after at least three months (window period).[19] Columns 6-7 show improved support among treated respondents for the claim that HIV positive boys should be allowed to play football. This is noteworthy, as Shuga prominently features a sub-plot about a boy who was born with HIV and struggles to remain part of a football team. Appendix Table A13 shows analogous results for a larger number of individual outcomes that are also explicitly featured in the TV series, such as knowledge about transmission during pregnancy, contagion through sexual intercourse, awareness of anti-retroviral (ARV) drugs.

The increase in HIV testing rates is an important result of our intervention and it is worth investigating deeper how it occurs. A first underlying factor is that individual knowledge about what testing is and why one should test improves. Secondly, it is possible that Shuga alerted people to the fact that HIV may be relatively common, thus leading them to update their expected risk of contagion.[20] We test this hypothesis in Appendix Table A14. We find no effects on the expectation that the respondent or the partner have HIV, while we find an increase of about 3 percentage points in the subjective probability that someone of the same age in the community is HIV+, representing an 8 percent increase over the control group mean.[21] Testing may thus at least in part be a response to higher perceived risk.

An additional dimension we investigate is the difference between people who have a stable partner and people who do not. Appendix Table A15 shows that while the nature of the partner does not typically affect the impact of Shuga on other outcomes, it has a strong predictive power when the dependent variable is the index of HIV testing (column 3). The increase in testing in response to treatment is entirely driven by people who do *not* have stable partners, possibly because they realize the risk involved (one of the messages featured in Shuga).

[18] The sample in this regression is smaller because not all respondents attended health camps. Also, given that the option to test for HIV was given at follow up but not at baseline, for this regression we cannot estimate the specification interacted with the baseline value of the outcome.

[19] The discussion on the window period is explicitly featured in a scene where the main female character receives the results of her HIV test and the nurse tells her that she cannot consider herself free from risk until she takes a second test at least three months later.

[20] For an analysis of testing in response to 'external HIV risk', see Godlonton and Thornton (2013).

[21] To elicit subjective probabilities, we followed Delavande and Kohler (2016) and asked respondents to select a number of beads ranging from 0 to 10, with each bead representing a 10 percent increment in probability. As commonly found in the literature, our respondents on average overestimate HIV prevalence rates. Note, however, that 1/4 of the characters in Shuga are HIV+, and in particular the main male character is. This may increase perceptions of how common HIV-AIDS is.

5.2 Risky sexual behavior

In Table 3 we estimate the effect of MTV Shuga on attitudes towards various sexual behaviors and on risky behavior itself as reported by the respondent. The sample for the behavioral outcomes (columns 3-4) is smaller because it is restricted to respondents who are sexually active. Panel A shows that the impact on the two aggregate indexes goes in the expected direction, namely improvement in attitudes and reduction in risky behavior (recall that our outcomes are constructed in a way that the expected treatment effect is positive). However, the effect is not significant –although Appendix Table A12 shows a significant improvement in attitudes for men.

In Panel B we turn to two important variables that enter the index of risky sexual behavior: concurrent partnerships and condom use.[22] Appendix Table A16 reports results for a broader set of individual outcomes. We find that respondents assigned to watch Shuga are less likely to have concurrent sexual partners. Based on the estimates in column 2, the total effect of treatment on the probability of not having concurrent partners is +3.3 percentage points when evaluated at the mean of the dependent variable. Appendix Table A16 also shows a negative effect on the number of concurrent partners is negative when evaluated at the mean, though the effect size depends on the number of partners at baseline.[23] The same table shows that attitudes towards concurrent partnerships also shift in a consistent direction.

In the last two columns of Table 3, Panel B we estimate the effect of treatment on the likelihood of having used a condom the last time the respondent had sex. We find no difference between people who watched MTV Shuga and those who did not. The same holds if we consider alternative self-reported measures of condom use and future intentions to do so (Appendix Table A16). The general lack of an effect is surprising because the importance of wearing condoms is repeatedly stressed in the show. We tested whether the results differed if we distinguished between 'main' and 'secondary' partners (respondents may view the latter as less safe), but found insignificant results in both cases. Also, the result does not seem to be driven by reporting bias, as the next set of results shows that we obtain similar (nil) results when we use behavioral outcomes from health camps.[24]

In Table 4 we consider outcomes related to risky sexual behavior that are 'objectively' measured at our health camps.[25] Panel A shows the results of a game where participants were

[22] These outcomes refer to actual behavior of respondents who are sexually active, hence the smaller number of observations.

[23] We also tested if treatment affected the likelihood of being sexually active and found no effects (results available from the authors).

[24] Note that access to condoms is not an issue in our study sites, as 4 in 5 respondents said that they could get a condom in less than 10 minutes if they wanted.

[25] The sample includes only respondents that attended health camp. Appendix Table A8 shows that the likelihood of attending health is uncorrelated with treatment status and with our index of risky sexual behavior (measured at baseline).

offered a choice between the value of one pack of condoms in cash, and one, two or three packs of condoms. The dependent variable takes value one if the respondent chose the condoms over the monetary amount and zero otherwise. While participants were more likely to choose condoms when the relative price was lower (i.e., when offered a higher number of packs), choice behavior did not differ among those who watched Shuga and those who did not. Results are equally insignificant for women and men (columns 3 to 6). This zero effect of treatment aligns with the results obtained in Table 3 when looking at self-reported condom use.

The absence of an effect on condom use is consistent with different possible explanations. One is that there is strong cultural resistance to condoms in the Nigerian context and Shuga was simply unable to overcome such resistance. Another is an endogenous response to safer sexual behavior by treated individuals. As the incidence of concurrent partnerships is reduced for treated individuals (see Table 3), and possibly the nature of the partners becomes safer, the reduced risk may have induced our treatment group to rely less on condoms.

An additional result consistent with the interpretation that the treatment group adopted a safer behavior comes from the prevalence of sexually transmitted diseases. Panel B of Table 4 shows the effect of treatment on the probability of testing positive for Chlamydia. Results are shown for the full sample, for women and for men. While the estimated effect is negative and comparable in size in all three samples, it is only statistically significant for the female subsample. This is not surprising as Chlamydia is more prevalent among women. Also, the fact that the prevalence is very low in general makes it difficult to detect an effect with high precision. On the other hand, the magnitude of the effect is quite sizeable relative to the baseline prevalence rate: exposure to Shuga leads to a 55 percent decrease in the likelihood that women test positive for Chlamydia. As discussed above, even in the absence of an effect on condom use, this improvement may be generated by more careful behavior on behalf of the respondent, e.g., decreasing the number of sexual partners or choosing 'safer' partners.[26]

5.3 The role of involvement with the narrative

According to the proponents of entertainment education, a key advantage of these programs is that viewers get engaged with the narrative and this leads them to pay more attention to content, learn from characters and be less defensive against external inputs (Singhal and Rogers, 1999). To test the role played by these factors we included in our endline survey a series of questions created by communication experts to measure two key dimensions.

The first dimension is what Green and Brock (2000) call 'transportation', and which could be alternatively described as 'engagement'. Individuals who are transported into the narrative of a movie tend to focus their cognitive attention on its messages and to have heightened

[26] We tested whether, conditional on showing symptoms, treated respondents were more likely to seek treatment for STDs and found that they were not. The estimated coefficient on *Treated* is insignificant and equal to 0.019, where the mean of the dependent variable is 0.15.

emotions, which reduces counter-arguing. We use twelve questions proposed by Green and Brock (2000), which include statements about things that happen during the screening and ask respondents to agree or disagree on a scale of 1 to 5. Example of these statements include: "You were distracted by activities in the room around you"; "You wanted to learn how the story ended"; "It affected you emotionally"; "You had a clear picture of the characters in the story".

The second dimension we want to explore is the extent to which viewers identify with the characters. Identification is understood to make viewers more receptive to the modeling of behavior and more likely to rehearse the arguments presented (Murphy et al., 2011). We use ten questions proposed by Cohen (2001), also in the form of statements with 5-point scale responses, which include for example: "While viewing the show you felt as if you were part of the action"; "you wanted the characters to succeed in achieving their goals"; "you felt you had experienced the same thing as the character". We aggregate the above questions into a *Transportation* and an *Identification* index, respectively, using principal component analysis. Appendix Table A17 reports the loading factors for the two indexes.

While these indexes cannot be considered as exogenous, in Appendix Table A18 we investigate which observable characteristics correlate with transportation and identification. We find that the only robust correlate of both indexes is wealth. For identification, also gender and speaking English as a primary or secondary language at home matter.[27] In estimating heterogeneous effects according to *Transportation* and *Identification*, we therefore control for the interaction between treatment and observable characteristics $\mathbf{X}_{i0}$.

In Table 5 we estimate the effect of treatment on the five indexes of HIV outcomes and sexual behavior, including an interaction term between *Treated* and *Transportation* (Panel A) or *Treated* and *Identification* (Panel B), plus the standalone variables.[28] Odd-numbered columns do not include the interactions $Treated * X_{i0}$, while even-numbered ones do. *Transportation* and *Identification* enhance the effect of watching Shuga, as shown by the positive coefficient on the interaction term between these indexes and treatment. In the top panel, this coefficient is positive and significant for three out of five indexes; in the bottom panel for two out of five (after correcting for multiple hypothesis testing). In terms of magnitude, for example, based on the estimates in column 1 of Panel A, a one standard deviation increase in *Transportation* is associated with a 0.45 standard deviation increase in *HIV knowledge* for the treatment group compared to the control one. The corresponding effect for a one standard deviation increase in *Identification* (Panel B) is a 0.3 standard deviation increase in *HIV knowledge.*[29]

[27] Women tend to identify less with the characters in Shuga, and people who speak English at home identify more (not surprisingly given that the language of Shuga is English).

[28] Online Appendix Table A19 reports the cross sectional results (without including the lagged dependent variable and its interaction with treatment), which yield a very similar coefficient on the interaction between treatment and *Transportation* (or *Identification*).

[29] Note that the negative coefficient on the standalone indicator *Treated* in the even-numbered columns of

While the results in Table 5 are strongly suggestive of a role for the entertainment component in inducing behavior change, as mentioned above one should be careful in interpreting them causally. However, notice that when we do not control for $Treated * \mathbf{X}_{i0}$ and $Treated * y_{i0}$ (odd-numbered columns), the coefficients on *Treated*∗*Transportation* and *Treated*∗*Identification* are very similar both in terms of magnitude and of significance. This increases our confidence that unobservables may not be driving the correlations we estimate.

6 Results: social effects

An important focus of this paper is whether, in addition to understanding if edutainment interventions are on average successful, we can say anything on the extent to which social effects may reinforce or undermine the impact of edutainment. Since identifying the relevant reference group is not obvious, we employ different strategies to address this question. First, we rely on experimental variation in the announcement of other people's reaction to Shuga (treatment T2). Second, we test if the basic version of MTV Shuga (treatment T1) affected viewers' perceptions about norms in their own community, and if this explains the change in their behavior. Third, we exploit experimental variation in the possibility of bringing friends to the show (treatment T3). Finally, we provide non-experimental evidence on treatment effect heterogeneity according to individuals' self-assessed degree of conformism.

6.1 Announcement treatment

Our first experimental design for testing the importance of social effects relies on complementing the basic treatment with an announcement on how other viewers reacted to Shuga. As explained in section 3.2, in half of the treated locations we complemented MTV Shuga with a short video that included interviews with young people who had watched the show, as well as 'smart graphs' with statistics on their reactions. Our model predicts that, if social effects are important ($\beta > 0$) and if our manipulation increases the precision of the signal about other people's choices ($p'_R > p_R$), then the difference between T2 and T1 should depend on the interaction between the T2 dummy and the individual's prior about the social norm (see equation 8).

Table 6 reports our estimates of equation (12) for the three outcomes for which we announced statistics in T2. The dependent variables are indicators for whether the respondent states that (i) he/she would reveal his/her status to the partner (column 1); (ii) it is not OK to

table 7 is only apparently counter-intuitive: given that those specifications include a full set of interactions between *Treated* and $\mathbf{X}_{i0}$, the coefficient on *Treated* is hard to interpret. As we report in the last row of table 7 (panels A and B), under the most conservative specification the fraction of individuals for whom the overall treatment effect is positive ranges from 66 to 98 percent for HIV-related outcomes and is around 50 percent for outcomes related to risky sexual behavior.

date a sugar daddy in order to finance one's education (column 2); and (iii) men should only have one partner (column 3). The variable *Prior on community* ($r_{i,t-1}$) is the respondent's baseline expectation of the share of community members who agree with each statement. The sample includes only treated individuals as we aim at comparing the effect of T1 and T2.

The results in Table 6 show that T2 did not have a differential effect compared to T1 and that the coefficient on *T2*Prior on community* is never significant. This suggests that learning about the reactions of other viewers did not elicit significant conformity effects on top of what the basic showing of Shuga may have already done. In Appendix Table A20 we also test whether on average T2 had a differential impact on our five outcome indexes compared to T1. We find that it did not.[30]

6.2 Conformity effects of basic treatment

The lack of a significant impact of T2 does not necessarily imply the absence of conformity. A possible reason for failing to find an effect of T2 is that the 'basic' treatment T1 may have already conveyed a signal about the prevailing norm in the reference group, compared to which T2 adds no new information. We next test whether this interpretation is supported in the data.

The first step is to establish whether T1 produced a shift in individual priors regarding social norms in the respondent's community. In Panel A of Table 7 we estimate specification (11) using as outcomes the respondents' expectations of the share of community members who would behave in a certain way or support certain views. The set of dependent variables in this table is the one for which we elicited such expectations, as described in section 4.2. We exclude from the sample people assigned to T2, so the regressor of interest is T1 and we compare the basic screening of MTV Shuga to the control TV series.

The results in Panel A suggest some degree of shift in perceived norms generated by T1. While Shuga did not affect perceptions about how many community members would reveal their HIV+ status or the status of a family member, it did improve attitudes towards HIV+ people (e.g., shopkeepers or boys playing football). The effects on blame also qualitatively point to a reduction in the stigma associated with HIV, although they are only significant for one out of three variables.

In Panel B of Table 7 we test whether the change in individuals' own attitudes was mediated by their perceived change in the prevailing norms. To this end, we turn to model equation (5) and observe that, if people were Bayesian, the coefficient on $T1 * r_{i,t-1}$ should be negative in columns 1, 3 and 4 (where the main treatment effect should be positive) and positive in the remaining ones (where the main effect should be negative). We only find this pattern in one

[30] We also tried to test whether T2 increased the precision of the signal about community norms, using as a proxy for precision respondents's own assessments of how sure they were about their priors. Unfortunately, we have little variation in this as 92 percent of respondents said that they were "sure" about their estimates.

out of seven outcomes (column 7). We thus conclude that changes in perceived norms do not seem to account for the impact of Shuga on individual attitudes and behavior.

6.3 Friends treatment

Our treatment T2 provided information on the reactions of "young people in neighboring communities" who had watched Shuga. These people looked similar to our respondents in age and socioeconomic status, but they did not personally know them. We next consider as an alternative peer group the respondents' own friends. Our treatment arm T3 was designed to test if watching Shuga with a friend led to different outcomes. As described in section 3.2, we randomly provided half of the treated sample with tickets that they could give to up to two friends.

Table 8 reports the estimated impacts on our five outcome indexes. The main regressor of interest is the dummy T3, which takes value 1 for respondents who received tickets to bring friends: its coefficient thus captures an intention-to-treat effect. In columns 1, 4, 7, 10, 13 we estimate an ANCOVA model, while in columns 2, 5, 8, 11 and 14 we also include the interaction between T3 and the lagged dependent variable. The coefficient of T3 is never significant in these specifications.

One possible interpretation is that social effects are absent altogether and individuals do not care about what their friends say. Another possibility is that there are social effects, but half of the sample brought friends who were positively inclined towards the messages of MTV Shuga, while another half brought friends who would 'talk them out' of those messages, generating a zero overall effect.[31]

In columns 3, 6, 9, 12 and 15 we test whether the effect of T3 differs for people who brought at least one friend of the opposite sex, compared to those who brought only friend(s) of the same sex. In the Nigerian context this may be a proxy for the fact that the individual attended the screening with his/her partner. We find a significant effect on HIV knowledge, suggesting that people who watched Shuga with a potential partner learnt more, possibly because the incentives to share knowledge and discuss HIV-related issues were greater. No other outcome, however, shows this effect.

6.4 Spillovers

The differential impact of treatment on knowledge depending on the gender of the friend also emerges from our analysis of spillovers. As described in section 3.2, in each location we interviewed a random sample of 'network members' who were not part of treatment T3 by

[31] As explained in section 3.1, due to logistical constraints we could not collect information at the screening sites on the friends that people brought along, except for their gender. Therefore we do not know the opinions of friends.

design, and thus could not have been directly influenced by Shuga. To test whether people who watched Shuga passed on any of the effects to friends who did not watch it, in Table 9 we estimate model (14).

The observations refer to network members, but the treatment status is that of the main study participant who 'nominated' the respondent. The variable '*Friend of Treated*' is a dummy equal to 1 if the respondent was a friend of a treated individual. The odd-numbered columns present the results of estimating equation (14).[32] We detect positive spillovers on HIV knowledge and negative ones on testing, though neither effect survives the correction for multiple hypothesis testing. There is no significant effect for the other variables.

In columns 2, 4, 6, 8 and 10 we test whether the effects are different for network members who have the opposite sex of the treated individual. Our conjecture is that friends of the opposite sex may include the respondent's partner, and messages like those conveyed by Shuga may be particularly effective if shared between members of a couple. Column 2 shows that indeed the positive effect on knowledge is stronger for friends of the opposite sex. No effect is found for other outcomes.

Overall, the above results uncover the presence of some knowledge spillovers regarding HIV, albeit not significant at conventional levels when correcting for multiple hypothesis testing. People who watched Shuga seem to have passed on some 'factual' information to their friends, but no spillovers are detected on attitudes and behavior. This suggests that, while edutainment programs may have trickle-down effects when it comes to information provision, in order to generate attitudinal and behavioral change direct exposure to the program is needed.[33]

These results can also partly speak to the issue of social effects: if the untreated individuals in our spillover sample conformed to the changed attitudes and behaviors of their friends who were exposed to Shuga, we should observe some impact on their own attitudes and behavior. The fact that we only observe (weak) impacts on knowledge is consistent with a modest role of conformism in the setting we study.

6.5 Self-assessment of conformism

Our final exercise to gauge the importance of social effects tests whether the impact of treatment differs based on respondents' baseline propensity to conform with other people's views. In the model, the parameter β captures how costly it is to deviate from the choices of a reference group. This strategy has the advantage that it would work even if different people conformed

[32] Appendix Table A21 reports the cross-sectional model.

[33] A caveat is that the samples of 'main respondents' and of '*network members*' may be different because the former were (randomly) selected from a population that revealed interest in TV programs by attending screening 0, while the latter were (randomly) selected from lists of contacts provided by the main participants. We tested for balance between the two groups and found that all outcome indexes except *Attitudes towards risky sexual behavior* are balanced (see Appendix Table A22).

with the views of different types of peer groups (e.g., some compare themselves to the average Nigerian while others only care about their friends). The variation we are using here is specific to the individual and tells us how much they care about the particular group that they have chosen to compare themselves to.

We use a series of questions aimed at measuring how strongly individuals identified with three of the values categorized by Schwartz (2012): conformity, tradition and self-direction. Online Appendix C describes the construction of the indexes in detail and Appendix Table A23 reports the list of variables and their loading factors from principal component analysis. We construct the following three indexes: (i) *Conformity* captures how inclined an individual is to restrain his/her own choices if these were to upset others or violate social norms; (ii) *Tradition* captures individuals' acceptance and commitment to the values that their culture or religion promote; (iii) *Self-direction*: captures how inclined an individual is to think and act independently. '*Tradition*' and '*Conformity*' are similar in that they capture individuals' willingness to subordinate themselves to what is expected from them, but they differ in the group respondents subordinate themselves to: in the case of conformity it is mainly people (e.g., parents or peers), while in the case of tradition it is religious and cultural customs.

In Table 10 we estimate a series of regressions having as outcomes our five outcome indexes, and as main regressors of interest *Treated* and the interaction of *Treated* with *Conformity* (Panel A), *Tradition* (Panel B) and *Self-direction* (Panel C), plus the lagged dependent variable and its interaction with treatment.[34] The goal is to test if exposure to MTV Shuga had differential effects depending on viewers' degree of conformity or independent judgement. Save a couple of exceptions, the interaction of treatment with these indexes is never significant at conventional levels. Clearly we cannot interpret these results in a causal sense, as conformism may be correlated with unobservables, but it is instructive that we fail to uncover significant correlations.

7 Robustness

In this section we discuss additional results and robustness checks to deal with some potential concerns.

Social desirability bias. Some of our dependent variables (notably biomarkers, HIV testing and condom choice) are objectively observed in health camps. Others are elicited through a survey but are factual (e.g., knowledge about transmission and treatment of HIV). Another set of outcomes, however, are self-reported and possibly subject to social desirability bias (e.g., attitudes, number of partners, etc.). Note that if reporting bias were similar across treated and control subjects, this would not be a problem: the concern only arises if treated individuals are more likely to misreport in a direction consistent with the message of Shuga.

[34] Appendix Table A24 reports results for the cross-sectional model.

Three pieces of evidence lead us to believe that our results are not driven by experimenter demand effects. First, the point estimates of our treatment effect on objective and subjective indicators for HIV testing are very similar at 0.031 and 0.025, respectively (see Table 2). In the presence of experimenter demand effects, one would expect the latter coefficient to be larger.

Second, many relevant self-reported outcomes do not display significant treatment effects. Notably, we find no effects on self-reported condom use (Panel B of Table 3) and on five other condom-related variables (Appendix Table A16).

Third, if treated individuals were differentially affected by reporting bias, we should find that when they have a chance of pleasing the research team by choosing condoms in the experimental game played in health camps, they should be more likely to do so. Panel A of Table 4 shows that this was not the case.

Heterogeneous effects by education and language. An interesting question is whether our effects differ depending on viewers' level of education or ability to understand English (as MTV Shuga was in English). In Appendix Table A25 we interact the treatment dummy with the respondent's years of education (odd-numbered columns) and with a dummy for whether respondent speaks English as the primary or secondary language at home (even-numbered columns). The average respondent in our sample has 11.6 years of education, and 96 percent of the respondents speak English at home as either primary or secondary language.[35]

We find that Shuga had a stronger effect on HIV knowledge and attitudes of relatively more educated respondents, although the coefficient on the interaction term with education is not significant when adjusting for multiple hypothesis testing. This may seem surprising as one may expect television to be a particularly effective means of communication for audiences with low literacy, but it should be noted that ours is a uniformly highly educated sample, as is the region.[36] So it is plausible that the education effect we uncover reflects the fact that the message of Shuga was relatively progressive.[37] No robust effects are found for the interaction between treatment and speaking English at home, possibly because the latter variable has very little variation in our sample.

8 Conclusions

In this paper we have tested the effectiveness of an entertainment education TV series, *MTV Shuga*, aimed at providing information and changing attitudes and behaviors related to HIV.

[35] Note that English is the language of instruction in schools and is the only official language in the country.

[36] The high level of education is not an artifact of our sample: the average respondent in the 2008 Nigeria DHS (of similar age and living in the South West) had 11 years of education.

[37] Note that despite the negative coefficient on the standalone treatment dummy, the overall effect of treatment is abundantly positive when evaluated at the mean years of education. For example, *ceteris paribus* the impact on knowledge is positive for any respondent with at least 9.5 years of education, which corresponds to the 4th percentile in the distribution of education.

The simple model we set up to motivate the analysis captures the idea that edutainment can work through an 'individual' or through a 'social' channel. We conducted a randomized controlled trial in urban Nigeria where young viewers were exposed to MTV Shuga or to a non-educational TV series. Among those who watched Shuga, we created additional variation in the 'social messages' they received and in the people with whom they watched the show.

We found that MTV Shuga led to significant improvements in knowledge about and attitudes towards HIV and to less risky sexual behavior. Treated subjects were twice as likely to get tested for HIV 8 to 9 months after the intervention. We also found reductions in STDs among women. Our experimental manipulations of the social norm component, on the other hand, did not produce significantly different results from the main treatment. Finally, we detected (weak) spillovers on friends who did not watch Shuga in terms of HIV knowledge, but not on attitudes and behavior.

While it is virtually impossible to embed every possible type of social interaction in a single experiment, our experimental design allowed for several different types of social effects, and we found no evidence that any of those explains the impact of Shuga. We thus learnt two lessons. First, the 'individual' effect of edutainment seems to have prevailed in the context of our study, and this is remarkable because it suggests that -at least in the context of HIV-AIDS- people react to the messages they see on TV regardless of what others say. Second, if policymakers wanted to leverage social effects in edutainment, they should experiment with different (and potentially larger) reference groups, e.g., schools, villages, etc. Finally, more research is needed to assess the potential role of conformity when the private versus public nature of the message is varied. This seems especially relevant for the edutainment agenda given the growing importance of social networks in today's society.

References

[1] Bandura, A. (1976), *Social Learning Theory*. Englewood Cliffs, NJ: Prentice-Hall.

[2] Banerjee, A., S. Barnhardt and E. Duflo (2015), "Movies, Margins and Marketing: Encouraging the Adoption of Iron-Fortified Salt", in D. Wise (ed.), *Insights in the Economics of Aging*, NBER.

[3] Benabou, R. and J. Tirole (2011), "Laws and Norms", NBER Working Paper 17579.

[4] Berg, G., and Zia, B. (2017). "Harnessing Emotional Connections to Improve Financial Decisions: Evaluating the Impact of Financial Education in Mainstream Media." *Journal of the European Economic Association*, 15(5), 1025-55.

[5] Bernard, T., S. Dercon, K. Orkin, and A. S. Taffesse (2014), "The Future in Mind: Aspirations and Forward-Looking Behaviour in Rural Ethiopia", CSAE Working Paper WPS/2014-16.

[6] Bjorkman Nyqvist, M., L. Corno, D. de Walque and J. Svensson (2018). "Incentivizing Safer Sexual Behaviour: Evidence from a Lottery Experiment on HIV Prevention", *American Economic Journal: Applied Economics*, 10(3): 287-314.

[7] Bjorvatn, K., A. Cappelen, L. Helgesson Sekeiz, E. Sørensen, B. Tungodden (2015), "Teaching through television: Experimental evidence on entrepreneurship education in Tanzania", mimeo, NHH Norwegian School of Economics.

[8] Bursztyn, L. and R. Jensen (2015), "How Does Peer Pressure Affect Educational Investments?", *Quarterly Journal of Economics*, 130(3), 1329-1367.

[9] Bursztyn, L., T. Fujiwara and A. Pallais (2017), "Acting Wife: Marriage Market Incentives and Labor Market Investments," *American Economic Review*, 107(11), 3288-3319.

[10] Bursztyn, L., A. Gonzalez and D. Yanagizawa-Drott (2018), "Misperceived Social Norms: Female Labor Force Participation in Saudi Arabia," NBER Working Paper 24736.

[11] Chong, A. and La Ferrara, E. (2009). "Television and Divorce: Evidence from Brazilian Novelas" *Journal of the European Economic Association: Papers & Proceedings* 7(2-3), 458-468.

[12] Cohen, J. (2001). "Defining identification: a theoretical look at the identification of audiences with media characters." *Mass communication and society,* 4(3), 245-64.

[13] Coville, A., Di Maro, V., Zottel, S., and Dunsch, FA (2014). "Nigeria's Nollywood Nudge. An entertaining approach to savings", in *Enhancing financial capability and behavior in low- and middle-income countries*, ch. 10, The World Bank.

[14] Delavande, A. and H.-P. Kohler (2016), "HIV-related Expectations and Risky Sexual Behavior in Malawi," *Review of Economic Studies*, 83(1), 118-164.

[15] DellaVigna, S., J. List and U. Malmendier (2012), "Testing for Altruism and Social Pressure in Charitable Giving," *The Quarterly Journal of Economics*, 127 (1), 1-56.

[16] DellaVigna, S., and E. La Ferrara (2015). "Economic and Social Impacts of the Media", in Anderson, S., Strömberg, D., and Waldfogel, J. (eds.), *Handbook of Media Economics*, vol. 1. North-Holland, Amsterdam.

[17] Duflo, E. P. Dupas and M. Kremer (2015), "Education, HIV, and Early Fertility: Experimental Evidence from Kenya", *American Economic Review*, 105(9), 2757-97.

[18] Fonner, V., K. Armstrong, C. Kennedy, K. O'Reilly and M. Sweat (2014). "School based sex education and HIV prevention in low- and middle-income countries: a systematic review and meta-analysis," *PLoS One*.

[19] Godlonton, S. and R. Thornton (2013), "Learning from Others' HIV Testing: Updating Beliefs and Responding to Risk," *American Economic Review: Papers and Proceedings*, 103(3), 439-444.

[20] Green, M. C., & Brock, T. C. (2000). "The role of transportation in the persuasiveness of public narratives." *Journal of Personality and Social Psychology*, 79, 701–721.

[21] Harrison, A., M. Newell, J. Imrie, G. Hoddinott (2010). "HIV prevention for South African youth: which interventions work? A systematic review of current evidence," *BMC Public Health*.

[22] Imbens, G. and J. Wooldridge (2009), "Recent Developments in the Econometrics of Program Evaluation", *Journal of Economic Literature*, 47(1), 5-86.

[23] Imbens, G. and D. Rubin (2015), *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge University Press.

[24] Jensen, R., and Oster, E. (2009). "The Power of TV: Cable Television and Women's Status in India." *The Quarterly Journal of Economics* 124(3), 1057-1094.

[25] Kearney, M.S., and Levine, P.B. (2015). "Media Influences on Social Outcomes: The Impact of MTV's 16 and Pregnant on Teen Childbearing." *American Economic Review*, 105(12) , 3597-3632.

[26] Kearney, M.S., and Levine, P.B. (2019). "Early Childhood Education by Television: Lessons from Sesame Street". *American Economic Journal: Applied Economics* 11(1), 318-50.

[27] Kling, J.R., J.B. Liebman and L.F. Katz (2007), "Experimental Analysis of Neighborhood Effects." *Econometrica*, 75(1), 83-119.

[28] Krishnaratne, S., B. Hensen, J. Cordes, J. Enstone and J. Hargreaves (2016). "Interventions to strengthen the HIV prevention cascade: a systematic review of reviews," *Lancet HIV*, 3(e307–17).

[29] La Ferrara, E. (2016). "Mass Media and Social Change: Can We Use Television to Fight Poverty?", *Journal of the European Economic Association*, 14(4), 791-827.

[30] La Ferrara, E., Chong, A., and Duryea, S. (2012). "Soap Operas and Fertility: Evidence from Brazil." *American Economic Journal: Applied Economics* 4(4), 1-31.

[31] McCoy, S., R. Kangwende and N. Padian (2010). "Behavior change interventions to prevent HIV infection among women living in low and middle income countries: a systematic review." *AIDS Behavior*.

[32] Michielsen, K., M. Chersich, S. Luchters, P. De Koker, R. Van Rossem and M. Temmerman (2010). "Effectiveness of HIV prevention for youth in sub-Saharan Africa: systematic review and meta-analysis of randomized and nonrandomized trials." *AIDS*.

[33] Murphy, S. T., L. B. Frank, M. B. Moran, and P. Patnoe-Woodley (2011), "Involved, Transported, or Emotional? Exploring the Determinants of Change in Knowledge, Attitudes, and Behavior in Entertainment-Education", *Journal of Communication*, 61, 407-31.

[34] NACA - National Agency for the Control of AIDS (2015), *Global AIDS Response Country Progress Report*, Abuja, Nigeria.

[35] Padian, N., S. McCoy, J. Balkus and J. Wasserheit (2010), "Weighing the gold in the gold standard: challenges in HIV prevention research", *AIDS*, 24(5), 621-35.

[36] Paluck, E.L. (2009), "Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda", *Journal of Personality and Social Psychology*, 96, 574-587.

[37] Paluck, E.L., and Green, D.P. (2009). "Deference, Dissent, and Dispute Resolution: A Field Experiment on a Mass Media Intervention in Rwanda." *American Political Science Review* 103(4), 622-644.

[38] Perez-Truglia, R. (2017), "Political Conformity: Event-Study Evidence from the United States." *Review of Economics and Statistics*, forthcoming.

[39] Ravallion, M., D. van de Walle, P. Dutta and R. Murgai (2015). "Empowering Poor People through Public Information? Lessons from a Movie in Rural India", *Journal of Public Economics*, 132, 13–22.

[40] Schwartz, Shalom H. (2012). "An Overview of the Schwartz Theory of Basic Values". *Online Readings in Psychology and Culture*, 2 (1), 1-20.

[41] Singhal, A., and E. M. Rogers (1999). *Entertainment education: A communication strategy for social change*. Mahwah, NJ: Erlbaum.

[42] Westfall, P. and S. Young (1993), *Resampling-based Multiple Testing: Examples and Methods for P-value Adjustment*, vol. 279, John Wiley & Sons.

Table 1: Average individual characteristics, pre-treatment

	Mean Control (1)	*Mean Treated* (2)	*Diff=0 (p-value)* (3)	*Normalized Diff.*[(a)] (4)	*No. Obs.* (5)
Panel A: Outcomes					
Indexes					
HIV knowledge	-0.049	0.205	0.177	0.029	5166
HIV attitudes	0.036	-0.016	0.625	-0.010	5166
HIV testing	-0.048	-0.035	0.916	0.002	5166
Attitudes towards risky sexual behavior	0.006	-0.041	0.638	-0.010	5166
Risky sexual behavior (conditional on sexually active)	-0.018	0.079	0.458	0.020	3246
Individual variables					
HIV transmitted during pregnancy	0.612	0.611	0.962	-0.001	5166
Mentions ARV drugs spontaneously	0.020	0.024	0.330	0.021	5166
Mentions drugs to live longer with HIV	0.619	0.634	0.310	0.021	5166
Recognizes ARV when mentioned by enumerator	0.187	0.213	0.030	0.046	5166
Knows that second test is necessary	0.277	0.287	0.450	0.016	5166
Knows about 3-months window period	0.074	0.089	0.078	0.038	5166
Can get HIV through intercourse	0.947	0.948	0.897	0.003	5166
Would buy from an HIV+ shopkeeper	0.415	0.427	0.398	0.018	5166
An HIV+ boy should play footbal	0.579	0.571	0.571	-0.012	5166
People HIV+ should not be blamed	0.652	0.632	0.165	-0.029	5166
HIV is not punishment for sleeping around	0.433	0.465	0.031	0.046	5166
Would reveal HIV status	0.707	0.694	0.365	-0.019	5166
Tested last 6 months (self-reported)	0.053	0.055	0.675	0.009	5166
Men should have one partner only	0.842	0.860	0.106	0.034	5163
Women should have one partner only	0.880	0.898	0.056	0.040	5166
Not ok date sugardaddy to finance educ	0.760	0.745	0.265	-0.024	5166
Not ok date sugardaddy for money	0.670	0.675	0.713	0.008	5166
Not ok date sugardaddy to go out	0.886	0.866	0.050	-0.042	5166
If a woman brings a condom does not mean she's not serious	0.579	0.603	0.111	0.034	5166
Has not had multiple concurrent sexual partners	0.776	0.786	0.521	0.018	2933
Number of current sexual partners if sexually active	1.329	1.276	0.119	-0.048	2349
Used condom the last time he/she had sex	0.520	0.495	0.219	-0.036	2690
Panel B: Controls					
Female	0.473	0.474	0.943	0.002	5166
Age	20.618	20.614	0.962	-0.001	5166
Currently attending school	0.342	0.350	0.565	0.012	5166
Years of education	11.598	11.596	0.950	-0.001	5166
Speaks English	0.958	0.962	0.484	0.015	5166
Single	0.232	0.224	0.509	-0.014	5166
Does not live with the family	0.229	0.246	0.186	0.028	5166
Household size	4.482	4.257	0.001	-0.067	5166
Wealth index	1.781	1.736	0.010	-0.054	5150
Home owner	0.448	0.355	0.000	-0.135	5165
Father's education > secondary	0.376	0.314	0.000	-0.092	3928
Mother's education > secondary	0.252	0.214	0.004	-0.065	4393
Muslim	0.370	0.374	0.799	0.005	5166
Yoruba	0.920	0.921	0.894	0.003	5166

Notes: (a) Normalized difference is the difference in the sample means of treatment and control groups divided by the square root of the sum of the sample variances.

Table 2: Impact on HIV outcomes

Panel A: HIV indexes						
Dep. Var. (Y_t):	*HIV knowledge*		*HIV attitudes*		*HIV testing*	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.898***	0.782***	0.344**	0.339***	0.353**	0.335**
	(0.244)	(0.215)	(0.135)	(0.103)	(0.148)	(0.128)
	[0.001]	[0.001]	[0.022]	[0.003]	[0.022]	[0.012]
Treated*Y_{t-1}		-0.065*		0.008		-0.033
		(0.035)		(0.037)		(0.032)
		[0.201]		[0.826]		[0.523]
Y_{t-1}		0.390***		0.360***		0.472***
		(0.029)		(0.030)		(0.027)
R-squared	0.085	0.198	0.050	0.179	0.092	0.261
P-value test joint sig[(a)]		0.000		0.001		0.011
Observations	4986	4986	4986	4986	4986	4986
Mean Dep. Var. (Control)	0.0400	0.0400	0.0439	0.0439	-0.139	-0.139
Std Dev Dep. Var. (Control)	6.106	6.106	3.474	3.474	4.145	4.145

Panel B: HIV individual outcomes							
	HIV testing					*HIV+ boy should be*	
Dep. Var. (Y_t):	*Objective*[(b)]	*Self-reported*		*Knows window period*		*allowed to play football*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated	0.031**	0.025***	0.026***	0.049***	0.045***	0.051***	0.080***
	(0.013)	(0.009)	(0.009)	(0.012)	(0.012)	(0.016)	(0.028)
	[0.014]	[0.014]	[0.004]	[0.001]	[0.003]	[0.012]	[0.022]
Treated*Y_{t-1}			-0.038		-0.007		-0.046
			(0.063)		(0.062)		(0.036)
			[0.550]		[0.955]		[0.581]
Y_{t-1}			0.253***		0.286***		0.300***
			(0.054)		(0.048)		(0.030)
R-squared	0.023	0.031	0.059	0.028	0.072	0.044	0.123
P-value test joint sig[(a)]			0.005		0.000		0.006
Observations	3828	4982	4971	4986	4986	4986	4986
Mean Dep. Var. (Control)	0.033	0.086	0.086	0.129	0.129	0.662	0.662
Std Dev Dep. Var. (Control)	0.180	0.280	0.280	0.335	0.335	0.473	0.473
Controls[(c)]	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. FWER adjusted p-values in square brackets, based on 10,000 replications. Increasing values of the dependent variable correspond to outcomes more consistent with the messages of Shuga.

(a) P-value test of joint significance: p-value for the null that the sum of the coefficient on Treated and Treated*Y_{t-1} (evaluated at the mean) is zero.

(b) For this outcome we cannot estimate the model including the lag of the dependent variable because testing through health-camps was only offered at follow-up.

(c) Controls in each regression include: female, age, currently attending school, years of education, english spoken at home, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table 3: Impact on Risky Sexual Behavior

Panel A: Indexes				
	Attitudes towards risky sexual behavior		*Risky sexual behavior*	
Dep. Var. (Y_t):	(1)	(2)	(3)	(4)
Treated	0.149	0.148	0.175	0.148
	(0.091)	(0.090)	(0.151)	(0.145)
	[0.205]	[0.196]	[0.337]	[0.317]
Treated*Y_{t-1}		0.001		-0.117**
		(0.034)		(0.056)
		[0.982]		[0.085]
Y_{t-1}		0.292***		0.322***
		(0.028)		(0.046)
R-squared	0.021	0.099	0.100	0.152
P-value test joint sig$^{(a)}$		0.104		0.318
Observations	4986	4986	3070	3070
Sample	Full Sample		Sexually Active	
Mean Dep. Var. (Control)	0.00186		-0.0631	
Std Dev Dep. Var. (Control)	3.452		3.625	
Panel B: Individual outcomes				
Dep. Var. (Y_t):	*Has NOT had multiple concurrent partners*		*Used condom last time he/she had sex*	
Treated	0.027*	0.113***	-0.003	-0.021
	(0.015)	(0.041)	(0.019)	(0.032)
	[0.232]	[0.018]	[0.844]	[0.506]
Treated*Y_{t-1}		-0.103**		0.038
		(0.041)		(0.053)
		[0.028]		[0.574]
Y_{t-1}		0.288***		0.319***
		(0.035)		(0.047)
R-squared	0.119	0.172	0.073	0.194
P-value test joint sig$^{(a)}$		0.007		0.513
Observations	3339	2623	3084	2308
Mean Dep. Var. (Control)	0.780		0.494	
Controls$^{(b)}$	Yes	Yes	Yes	Yes

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. FWER adjusted p-values in square brackets, based on 10,000 replications. Increasing values of the dependent variable correspond to outcomes more consistent with the messages of Shuga.

(a) P-value test of joint significance: p-value for the null that the sum of the coefficient on Treated and Treated*Y_{t-1} (evaluated at the mean) is zero.

(b) Controls in each regression include: female, age, currently attending school, years of education, english spoken at home, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table 4: Risky sexual behavior outcomes measured at health camps

Panel A: Demand for condoms

	Full Sample		*Females*		*Males*	
Dep. Var. =1 if chose condoms over N50	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.013	-0.024	0.012	-0.006	0.014	-0.045
	(0.021)	(0.043)	(0.026)	(0.054)	(0.028)	(0.066)
Treated * # packs offered		0.018		0.009		0.028
		(0.019)		(0.028)		(0.030)
# packs offered	0.059***	0.046***	0.049***	0.043*	0.068***	0.049*
	(0.009)	(0.016)	(0.012)	(0.023)	(0.014)	(0.025)
Constant	-0.061	-0.033	-0.304**	-0.292*	-0.085	-0.039
	(0.116)	(0.118)	(0.148)	(0.160)	(0.187)	(0.194)
Observations	3,827	3,827	1,844	1,844	1,983	1,983
R-squared	0.137	0.137	0.054	0.054	0.063	0.063
Controls[(a)]	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var. (Control)	0.383	0.383	0.229	0.229	0.520	0.520

Panel B: STD biomarkers

	Full Sample	*Females*	*Males*
Dep. Var. =1 if tested positive for Chlamydia	(1)	(2)	(3)
Treated	-0.014	-0.017*	-0.013
	(0.012)	(0.010)	(0.015)
Observations	3,820	1,839	1,981
R-squared	0.010	0.024	0.014
Controls[(a)]	Yes	Yes	Yes
Mean Dep. Var. (Control)	0.029	0.031	0.013

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively.

(a) Controls in each regression include: female, age, currently attending school, years of education, english spoken, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table 5: Involvement with the narrative

Dependent variable:	*HIV knowledge*		*HIV attitudes*		*HIV testing*		*Attitudes towards risky sexual behavior*		*Risky sexual behavior (for sexually active)*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Transportation Index										
Treated*Transportation	0.428***	0.436***	0.141**	0.146**	0.120	0.117	0.311***	0.318***	0.091	0.092
	(0.108)	(0.107)	(0.067)	(0.066)	(0.076)	(0.076)	(0.064)	(0.065)	(0.066)	(0.065)
	[0.001]	[0.000]	[0.075]	[0.061]	[0.123]	[0.133]	[0.000]	[0.000]	[0.169]	[0.167]
Treated	1.035***	-7.329**	0.333***	-1.970	0.281*	-2.116	-0.009	-1.813	0.085	-3.607*
	(0.221)	(3.561)	(0.125)	(1.721)	(0.161)	(2.021)	(0.115)	(1.905)	(0.149)	(2.080)
	[0.000]	[0.122]	[0.020]	[0.439]	[0.086]	[0.439]	[0.936]	[0.345]	[0.809]	[0.168]
Treated*Y_{t-1}	-0.053	-0.075*	-0.008	-0.019	-0.055	-0.073	-0.027	-0.038	-0.131**	-0.123*
	(0.038)	(0.042)	(0.038)	(0.039)	(0.043)	(0.045)	(0.038)	(0.038)	(0.065)	(0.066)
	[0.430]	[0.222]	[0.829]	[0.633]	[0.430]	[0.222]	[0.483]	[0.328]	[0.101]	[0.130]
Transportation	-0.058	-0.069	0.032	0.026	-0.000	-0.001	-0.007	-0.008	-0.038	-0.037
	(0.076)	(0.076)	(0.039)	(0.039)	(0.048)	(0.048)	(0.045)	(0.046)	(0.047)	(0.046)
Y_{t-1}	0.385***	0.401***	0.376***	0.384***	0.500***	0.512***	0.309***	0.316***	0.359***	0.355***
	(0.029)	(0.030)	(0.030)	(0.030)	(0.037)	(0.037)	(0.030)	(0.030)	(0.051)	(0.052)
Observations	3,753	3,753	3,753	3,753	3,753	3,753	3,753	3,753	2,279	2,279
R-squared	0.215	0.220	0.186	0.188	0.272	0.274	0.119	0.125	0.175	0.180
Controls [(a)]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls*Treated	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean Dep. Var. (Control)	0.0400	0.0400	0.0439	0.0439	-0.139	-0.139	0.00186	0.00186	-0.0631	-0.0631
Std Dev Dep. Var. (Control)	6.106	6.106	3.474	3.474	4.145	4.145	3.452	3.452	3.625	3.625
Share with treatment effect > 0	0.892	0.791	0.892	0.764	0.864	0.661	0.523	0.520	0.426	0.483
Panel B: Identification Index										
Treated*Identification	0.251***	0.273***	0.062	0.063	0.098	0.092	0.162***	0.160**	0.016	0.010
	(0.092)	(0.092)	(0.059)	(0.061)	(0.074)	(0.075)	(0.061)	(0.061)	(0.060)	(0.059)
	[0.025]	[0.015]	[0.353]	[0.388]	[0.353]	[0.388]	[0.022]	[0.019]	[0.785]	[0.868]
Treated	1.061***	-7.357**	0.358***	-2.026	0.285*	-2.172	0.013	-1.908	0.102	-3.601*
	(0.214)	(3.567)	(0.125)	(1.740)	(0.166)	(2.010)	(0.115)	(1.914)	(0.151)	(2.071)
	[0.000]	[0.126]	[0.012]	[0.426]	[0.092]	[0.426]	[0.915]	[0.320]	[0.784]	[0.163]
Treated*Y_{t-1}	-0.050	-0.072*	-0.006	-0.017	-0.055	-0.074	-0.022	-0.032	-0.131**	-0.125*
	(0.038)	(0.042)	(0.038)	(0.040)	(0.043)	(0.045)	(0.038)	(0.038)	(0.065)	(0.066)
	[0.477]	[0.249]	[0.866]	[0.658]	[0.477]	[0.249]	[0.574]	[0.397]	[0.102]	[0.122]
Identification	-0.021	-0.037	0.037	0.035	0.006	0.007	0.026	0.030	-0.028	-0.021
	(0.066)	(0.066)	(0.041)	(0.043)	(0.055)	(0.055)	(0.041)	(0.042)	(0.039)	(0.038)
Y_{t-1}	0.385***	0.401***	0.376***	0.384***	0.500***	0.512***	0.309***	0.316***	0.358***	0.356***
	(0.029)	(0.030)	(0.030)	(0.030)	(0.037)	(0.037)	(0.030)	(0.030)	(0.051)	(0.052)
Observations	3,753	3,753	3,753	3,753	3,753	3,753	3,753	3,753	2,279	2,279
R-squared	0.211	0.214	0.183	0.185	0.271	0.274	0.110	0.115	0.175	0.180
Controls[(a)]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls*Treated	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean Dep. Var. (Control)	0.0400	0.0400	0.0439	0.0439	-0.139	-0.139	0.00186	0.00186	-0.0631	-0.0631
Std Dev Dep. Var. (Control)	6.106	6.106	3.474	3.474	4.145	4.145	3.452	3.452	3.625	3.625
Share with treatment effect > 0	0.957	0.834	0.987	0.825	0.883	0.661	0.549	0.511	0.197	0.487

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. FWER adjusted p-values in square brackets, based on 10,000 replications. Increasing values of the dependent variable correspond to outcomes more consistent with the messages of Shuga.

(a) Controls include: female, age, currently attending school, years of education, English spoken at home, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table 6: Effect of announcing social norm (T2)

Dep. Var. (Y_{t-1}):	*Would reveal HIV status*		*Not ok date sugardaddy to finance education*		*Men should have one partner only*	
	(1)	(2)	(3)	(4)	(5)	(6)
T2	0.016	0.019	0.011	0.069	-0.037	-0.011
	(0.027)	(0.038)	(0.032)	(0.050)	(0.033)	(0.047)
Y_{t-1}	0.231***	0.235***	0.215***	0.246***	0.110***	0.129***
	(0.018)	(0.021)	(0.020)	(0.033)	(0.020)	(0.031)
Prior on community (r_{t-1})	0.125**	0.123**	-0.130***	-0.117**	0.001	-0.005
	(0.051)	(0.050)	(0.048)	(0.048)	(0.036)	(0.038)
T2 * Prior on community (r_{t-1})	-0.073	-0.071	-0.038	-0.065	0.073	0.085
	(0.067)	(0.067)	(0.064)	(0.065)	(0.054)	(0.055)
T2*Y_{t-1}		-0.006		-0.060		-0.038
		(0.037)		(0.042)		(0.041)
Observations	3,402	3,402	3,402	3,402	3,394	3,394
R-squared	0.091	0.091	0.080	0.081	0.044	0.045
Controls [(a)]	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var. (Control)	0.738	0.738	0.751	0.751	0.903	0.903

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (54 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. Sample includes treated individuals only (T1+T2).

(a) Controls include: female, age, currently attending school, years of education, English spoken at home, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table 7: Shuga and perceived norms

	Would reveal HIV+ status to partner	*Would prefer to keep HIV+ of family member a secret*	*Would buy fruit & veg from HIV+ shopkeeper*	*Think an HIV+ boy should be allowed to play football*	*Think testing for HIV means sleeping around*	*Think people with HIV should be blamed*	*Think HIV is a punishment for sleeping around*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Dep. var. ($r_{i,t}$) is respondent's expectation on the share of peers who [column heading]							
T1	0.018	0.003	0.037***	0.035**	-0.019	-0.017	-0.059***
	(0.014)	(0.025)	(0.012)	(0.015)	(0.020)	(0.019)	(0.022)
	[0.886]	[0.993]	[0.033]	[0.184]	[0.952]	[0.952]	[0.096]
T1*ri,t-1	-0.023	-0.008	-0.138***	-0.061*	0.020	0.016	0.065**
	(0.039)	(0.033)	(0.038)	(0.034)	(0.032)	(0.028)	(0.030)
	[0.998]	[0.998]	[0.006]	[0.537]	[0.998]	[0.998]	[0.335]
$r_{i,t-1}$ (a)	0.149***	0.139***	0.268***	0.204***	0.104***	0.163***	0.102***
	(0.029)	(0.024)	(0.027)	(0.023)	(0.022)	(0.020)	(0.023)
R-squared	0.035	0.041	0.107	0.127	0.039	0.051	0.059
Mean Dep. Var. (Control)	0.319	0.589	0.260	0.377	0.538	0.570	0.644
Panel B: Dep. var. ($Y_{i,t}$) is respondent's own opinion on [column heading]							
T1	0.030	0.089*	0.091***	0.088**	-0.066**	-0.009	-0.093**
	(0.040)	(0.047)	(0.025)	(0.035)	(0.032)	(0.038)	(0.036)
	[0.701]	[0.181]	[0.005]	[0.074]	[0.176]	[0.822]	[0.071]
T1*Prior on community ($r_{i,t-1}$)	0.064	-0.008	-0.022	-0.025	0.084	0.007	0.156**
	(0.069)	(0.060)	(0.079)	(0.051)	(0.051)	(0.060)	(0.063)
	[0.888]	[0.990]	[0.988]	[0.981]	[0.476]	[0.989]	[0.112]
T1*$Y_{i,t-1}$	-0.043	-0.075*	-0.111***	-0.040	-0.049	-0.032	-0.056
	(0.034)	(0.042)	(0.041)	(0.044)	(0.034)	(0.043)	(0.042)
	[0.570]	[0.387]	[0.006]	[0.594]	[0.570]	[0.594]	[0.570]
Prior on community ($r_{i,t-1}$)	0.057	0.040	0.081	0.096***	-0.029	0.117***	0.028
	(0.048)	(0.040)	(0.055)	(0.036)	(0.038)	(0.039)	(0.049)
$Y_{i,t-1}$ (b)	0.279***	0.223***	0.359***	0.284***	0.152***	0.193***	0.260***
	(0.028)	(0.028)	(0.031)	(0.030)	(0.024)	(0.032)	(0.028)
R-squared	0.108	0.086	0.131	0.124	0.058	0.073	0.100
Mean Dep. Var. (Control)	0.726	0.712	0.518	0.694	0.161	0.311	0.487
Observations	3,262	3,262	3,262	3,262	3,262	3,262	3,262
Controls(c)	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors in parentheses adjusted for clustering at the screening centre level (53 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. FWER adjusted p-values in square brackets based on 10,000 replications. Sample includes only T1 + Control.

(a) ri,t-1 is the respondent's baseline expectation of the share of community members who agree with the statement in the heading of each column.

(b) Yi,t-1 is a dummy taking value 1 if at baseline the respondent agreed with the statement in the heading of each column.

(c) Controls include: female, age, currently attending school, years of education, english spoken at home, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table 8: Friend invitations (T3)

Dep. Var. (Y_{t-1}):	*HIV knowledge*			*HIV attitudes*			*HIV testing*			*Attitudes towards risky sexual behavior*			*Risky sexual behavior (for sexually active)*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
T3	-0.105	-0.105	-0.149	-0.001	-0.001	-0.131	0.143	0.144	0.119	-0.068	-0.068	-0.161	-0.206*	-0.207*	-0.285
	(0.192)	(0.190)	(0.235)	(0.107)	(0.107)	(0.128)	(0.126)	(0.126)	(0.153)	(0.109)	(0.109)	(0.140)	(0.122)	(0.123)	(0.172)
	[0.897]	[0.902]	[0.693]	[0.991]	[0.992]	[0.693]	[0.697]	[0.696]	[0.693]	[0.897]	[0.902]	[0.693]	[0.386]	[0.407]	[0.412]
Y_{t-1}	0.319***	0.320***	0.320***	0.363***	0.367***	0.364***	0.434***	0.426***	0.433***	0.290***	0.288***	0.289***	0.211***	0.209***	0.211***
	(0.024)	(0.027)	(0.024)	(0.023)	(0.031)	(0.023)	(0.020)	(0.025)	(0.020)	(0.021)	(0.026)	(0.021)	(0.034)	(0.045)	(0.034)
T3 * Y_{t-1}		-0.001			-0.007			0.015			0.003			0.005	
		(0.034)			(0.029)			(0.031)			(0.035)			(0.054)	
		[0.901]			[0.999]			[0.994]			[0.999]			[0.999]	
Friend Invited is of Opposite Sex			0.898**			-0.122			-0.258			0.108			-0.413
			(0.412)			(0.214)			(0.265)			(0.217)			(0.255)
Brought a Friend			-0.219			0.326**			0.143			0.165			0.315
			(0.274)			(0.161)			(0.235)			(0.155)			(0.257)
Observations	3,402	3,402	3,402	3,402	3,402	3,402	3,402	3,402	3,402	3,402	3,402	3,402	2,117	2,117	2,117
R-squared	0.188	0.188	0.189	0.184	0.184	0.184	0.246	0.246	0.246	0.107	0.107	0.107	0.135	0.135	0.136
Controls [(a)]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep Var. (Control)	0.839	0.839	0.839	0.415	0.415	0.415	0.200	0.200	0.200	0.124	0.124	0.124	0.126	0.126	0.0827

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (54 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. Increasing values of the dependent variable correspond to outcomes more consistent with the messages of Shuga. Sample includes only treated individuals (T1+T2). FWER adjusted p-values in square brackets, based on 10,000 replications.

T3 is a dummy taking value 1 if the respondent was given extra tickets to bring up to two friends.

(a) Controls in each regression include: female, age, currently attending school, years of education, english spoken, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table 9: Spillovers

Dep. Var. (Y_t):	*HIV knowledge*		*HIV attitudes*		*HIV testing*		*Attitudes towards risky sexual behavior*		*Risky sexual behavior (for sexually active)*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Friend of Treated	0.640*	0.540	-0.082	-0.124	-0.512*	-0.545*	-0.138	-0.085	-0.337	-0.311
	(0.332)	(0.328)	(0.232)	(0.228)	(0.274)	(0.289)	(0.207)	(0.218)	(0.314)	(0.314)
	[0.106]	[0.199]	[0.812]	[0.587]	[0.106]	[0.178]	[0.625]	[0.704]	[0.455]	[0.536]
Friend of Treated*Opposite Sex		2.493**		1.317		0.840		-1.383		-1.089
		(1.234)		(0.961)		(1.143)		(1.057)		(1.735)
		[0.228]		[0.386]		[0.519]		[0.455]		[0.624]
Opposite Sex		-1.540*		-0.072		-0.456		0.562		0.211
		(0.786)		(0.592)		(1.011)		(0.735)		(0.926)
Y_{t-1}	0.434***	0.438***	0.395***	0.397***	0.301***	0.300***	0.273***	0.271***	0.300***	0.299***
	(0.060)	(0.060)	(0.060)	(0.061)	(0.071)	(0.072)	(0.060)	(0.061)	(0.100)	(0.101)
Friend of Treated * Y_{t-1}	-0.110	-0.113	-0.148*	-0.148*	0.013	0.012	0.032	0.034	-0.107	-0.106
	(0.072)	(0.072)	(0.075)	(0.075)	(0.082)	(0.083)	(0.072)	(0.073)	(0.118)	(0.119)
	[0.236]	[0.238]	[0.154]	[0.156]	[0.876]	[0.881]	[0.789]	[0.646]	[0.789]	[0.624]
Observations	924	924	924	924	924	924	924	924	580	580
R-squared	0.249	0.250	0.177	0.179	0.222	0.222	0.139	0.140	0.184	0.185
Controls [(a)]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var. (Control)	-0.185	-0.185	-0.316	-0.316	-0.164	-0.164	-0.107	-0.107	-0.213	-0.213

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. Increasing values of the dependent variable correspond to outcomes more consistent with the messages of Shuga. "Opposite sex" is a dummy equal to 1 if respondent has the opposite sex as his/her treated friend. Sample includes network friends that have non-missing Y_{t-1}. FWER adjusted p-values in square brackets, based on 10,000 replications.

(a) Controls in each regression include: female, age, currently attending school, years of education, english spoken, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table 10: Treatment effects and conformism

Dependent variable Y_t:	(1) *HIV knowledge*	(2) *HIV attitudes*	(3) *HIV testing*	(4) *Attitudes towards risky sexual behavior*	(5) *Risky sexual behavior (for sexually active)*
Panel A: Conformity					
Treated	0.775***	0.345***	0.335**	0.148	0.149
	(0.214)	(0.105)	(0.128)	(0.089)	(0.145)
	[0.001]	[0.004]	[0.012]	[0.188]	[0.312]
Treated*Conformity	0.192	-0.188*	0.028	-0.044	0.009
	(0.132)	(0.099)	(0.093)	(0.076)	(0.097)
	[0.276]	[0.167]	[0.768]	[0.815]	[0.926]
Conformity	-0.052	0.138*	-0.077	0.054	0.030
	(0.108)	(0.076)	(0.065)	(0.060)	(0.084)
Y_{t-1} *Treated	-0.067*	0.013	-0.034	0.004	-0.116**
	(0.036)	(0.036)	(0.032)	(0.033)	(0.058)
	[0.185]	[0.715]	[0.506]	[0.907]	[0.089]
Y_{t-1}	0.391***	0.357***	0.473***	0.288***	0.320***
	(0.029)	(0.029)	(0.027)	(0.027)	(0.048)
R-squared	0.198	0.180	0.262	0.099	0.152
P-value of joint significance	0.000	0.001	0.010	0.105	0.305
Panel B: Tradition					
Treated	0.784***	0.342***	0.336**	0.147	0.143
	(0.215)	(0.104)	(0.128)	(0.090)	(0.147)
	[0.001]	[0.004]	[0.011]	[0.203]	[0.336]
Treated*Tradition	-0.016	-0.152**	-0.007	-0.017	-0.131
	(0.141)	(0.065)	(0.087)	(0.087)	(0.115)
	[0.993]	[0.063]	[0.992]	[0.845]	[0.445]
Tradition	-0.031	0.100*	-0.036	0.057	0.174*
	(0.109)	(0.052)	(0.064)	(0.075)	(0.104)
Y_{t-1} * Treated	-0.065*	0.011	-0.033	0.001	-0.111*
	(0.035)	(0.037)	(0.032)	(0.034)	(0.057)
	[0.204]	[0.767]	[0.524]	[0.979]	[0.113]
Y_{t-1}	0.390***	0.359***	0.472***	0.290***	0.314***
	(0.029)	(0.030)	(0.027)	(0.027)	(0.047)
R-squared	0.198	0.180	0.261	0.099	0.154
P-value of joint significance	0.000	0.002	0.010	0.106	0.355
Panel C: Self Direction					
Treated	0.762***	0.333***	0.332**	0.146	0.153
	(0.213)	(0.104)	(0.128)	(0.091)	(0.141)
	[0.001]	[0.005]	[0.010]	[0.211]	[0.281]
Treated*Self-direction	0.086	0.035	0.142**	0.060	0.140*
	(0.112)	(0.065)	(0.058)	(0.074)	(0.079)
	[0.685]	[0.685]	[0.055]	[0.421]	[0.148]
Self-direction	-0.358***	-0.098*	-0.066*	-0.040	-0.161**
	(0.088)	(0.050)	(0.039)	(0.063)	(0.061)
Y_{t-1} * Treated	-0.060*	0.009	-0.032	0.004	-0.120**
	(0.035)	(0.038)	(0.032)	(0.035)	(0.056)
	[0.257]	[0.805]	[0.560]	[0.922]	[0.077]
Y_{t-1}	0.380***	0.357***	0.471***	0.290***	0.323***
	(0.028)	(0.031)	(0.027)	(0.028)	(0.046)
R-squared	0.202	0.180	0.262	0.099	0.153
P-value of joint significance	0.000	0.002	0.011	0.112	0.280
Controls [(a)]	Yes	Yes	Yes	Yes	Yes
Observations	4,986	4,986	4,986	4,986	3,070
Mean Dep. Var. (Control)	0.040	0.044	-0.139	0.00186	-0.063

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. Increasing values of the dependent variable correspond to outcomes more consistent with the messages of Shuga. FWER adjusted p-values in square brackets, based on 10,000 replications.

(a) Controls include: female, age, currently attending school, years of education, english spoken at home, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

ONLINE APPENDIX

The Entertaining Way to Behavioral Change: Fighting HIV with MTV

Abhijit Banerjee	Eliana La Ferrara	Victor Orozco
MIT	Bocconi University IGIER and LEAP	Oxford University and The World Bank

This version: April 2019

A. Sample selection

Locations

The study sites were 80 urban and peri-urban locations chosen in 7 towns across three states of South-West Nigeria. The distribution of locations across states and towns is as follows. Oyo state: Ibadan (50 locations), Ogbomosho (6 locations), Oyo (4 locations). Osun state: Ile-Ife (3 locations), Ilesha (4 locations), Osogbo (7 locations). Ondo state: Akure (6 locations). The selection of these towns balanced competing requirements such as: (i) excluding states earmarked by MTV as priority states for marketing Shuga (to avoid contamination of the control group); (ii) excluding areas where the integrity of the evaluation could have been compromised by security risks; (iii) choosing contiguous states to facilitate the logistical implementation. Locations were defined by drawing a 2-mile radius around each screening center where the intervention was implemented, and randomly selecting households within this radius. We ensured that there were buffer zones between communities to minimize the risk of contamination across study groups. These locations constitute our unit of randomization. Appendix Figure A1 shows the geographic distribution of treatment and control locations.

Main study participants

To identify study participants, we adopted a three-step recruitment strategy. First, enumerators visited a random selection of $200-225$ households in each location and collected basic demographic information about all young people aged 18 to 25 residing in those households.

The condition for a household to be in the study was that at least one of the members should be in the target age range of our intervention, i.e., 18-25 years old.

Second, the research team randomly selected one person age 18-25 in each household to be invited to the screening of a Nigerian movie, different from MTV Shuga. We denote this initial screening as Screening 0. The selection was stratified by gender, half males and half females. Out of 17,224 people invited to Screening 0, 6,613 attended, with a turnout rate of 38.4 percent. Appendix Table A1 compares those who chose to attend Screening 0 and those who did not, on a number of characteristics that we collected when we first visited the households, as well as the normalized pair-wise differences and the p-values for the test that the difference in means is zero.[1] The two samples are well balanced, with few variables showing significant differences, but of extremely small magnitude: *in all* cases the normalized difference is *below the threshold of* 0.25 *suggested by Imbens and Rubin (2015).* We thus conclude that those who attended Screening 0 were a representative subset of the households invited.

The third step was the selection of our baseline sample. In each location enumerators paid home visits to 64 individuals among those who attended Screening 0, randomly selected.[2] All were invited to participate in the study and administered the baseline survey if they agreed. *Appendix Table A2 shows balance in observable characteristics between the people who attended Screening 0 and were selected into the baseline sample and those that were not.* We also compared our sample to the 2008 Nigeria DHS, restricted to the South-Western region, and found that the two were quite similar in terms of religion (a third Muslim and two thirds Christian), years of education (around 11 years) and television-ownership rates (around 90 percent).

At the end of the survey, the main study participants received invitations to attend Screenings 1 and 2, organized in the two weekends following the interview. Those in the 'Friends treatment' also received invitations to bring up to two friends of their choice. Note that invitees were not told what they would watch, neither before Screening 0 nor before Screening 1. It is only when attending Screening 1 that they learnt they were watching Shuga or Gidi Up.

Network members

We complemented our data collection on the main study participants with a sample of con-

[1] The normalized difference is a scale-free measure of the difference in distributions, recommended by Imbens and Wooldridge (2009):

$$\Delta = \frac{\overline{X}_A - \overline{X}_B}{\sqrt{S_A^2 + S_B^2}}$$

where $\overline{X}_A$ and $\overline{X}_B$ are the means of covariate X in groups A and B, respectively, and S_A^2 and S_B^2 are the corresponding sample variances of X.

[2] In some locations we have 65 respondents instead of 64, because our teams consisted of eight enumerators working in parallel. Enumerators were given a target number of respondents to interview per day and they did not know how many their colleagues had interviewed until the end of the day.

tacts from their social networks that we use for detecting potential spillovers. In all locations, before taking the baseline survey, study participants were asked to list two friends to whom they regularly talked and who lived in the community ('network members').

In each location, we administered the baseline and the follow-up surveys to a random sample of 15 network members among those indicated by main respondents who were *not* in the 'Friends' treatment. No more than one friend per main respondent was interviewed. By construction, contacts of those main respondents who had been given extra tickets for friends (and who could have therefore attended the screenings) are not part of the spillover sample.

B. Timeline of activities

In each location, all baseline survey and screening activities were concluded in four weeks. Implementation was rolled out so that activities in a given location were completed before moving to the next, in order to minimize attrition due to subjects forgetting about the screenings, travelling or relocating.

The overall timeline of the project across all locations was as follows: listing, Screening 0 and the baseline survey occurred between mid-September and the first week of December 2014; Screenings 1 and 2 between October and end of December; the endline survey and health camps between end of May and August 2015. In each location, baseline and endline were 8 weeks apart.

The typical sequence of activities in a given location was as follows:
MONTH 1

- Week 1: listing
- Weekend 1: Screening 0
- Weeks 2-3: baseline survey
- Weekend 3: Screening 1
- Weekend 4: Screening 2

MONTH 8

- Week 1: Follow-up survey
- Weekend 1: Health camp 1 - Chlamydia testing and distribution of referral letters for HIV testing
- Weekend 3: Health camp 2 - test results for Chlamydia test and post-test counselling

C. Construction of indexes

C.1 Outcome indexes

We construct our outcome indexes related to HIV and risky sexual behavior using two methods. The first method follows Kling et al. (2007): we take equally weighted averages of the z-scores of the variables that enter each index, where the sign of each variable is oriented so that answers consistent with Shuga's message translate into higher values of the index.[3] Appendix Tables A6 and A7 report list of variables contained in each index, with a sign (+) or (−) to denote whether the variable enters the index with a positive or negative sign. Variables are oriented so that the impact of treatment on each component of the index should be positive.

The second method uses principal components analysis. Starting from the same lists of variables as above, we extract the first principal component for each family of outcomes. The individual variables and their loading factors are shown in Appendix Tables A6 and A7.

C.2 Indexes of conformism

In our survey we included a series of questions aimed at measuring how strongly individuals identified with three of the values categorized by Schwartz (2012): conformity, tradition and self-direction. For each category, respondents were read four questions describing people with certain characteristics and were asked how similar each person was to them, with answers on a 5-point scale ranging from "not like me at all" to "very much like me". We aggregate the four questions in an index using principal component analysis (see Appendix Table A23 for the list of variables and loading factors) and we construct the following three indexes.

Conformity: captures how inclined an individual is to restrain his/her own choices if these were to upset others or violate social norms. People with a high value of this index believe that people should do what they are told, be obedient and polite, and they generally have a taste for smooth social interaction, even at the cost of self-restraint.

Tradition: captures individuals' acceptance and commitment to the values that their culture or religion promote. Respondents who identify with this profile believe that people should be humble and be satisfied with what they have. 'Tradition' and 'Conformity' are similar in the sense that they capture individuals' willingness to subordinate to what is expected from them, but they differ in the group to which one subordinates him/herself: in the case of conformity it is mainly people (e.g., parents or peers), while in the case of tradition it is religious and cultural customs.

[3]To deal with missing values we follow Kling et al. (2007): if a respondent has a non-missing value for at least one of the variables in an index, we impute any missing values for the other variables using the random assignment group mean. This implies that differences between treatment and control means of an index coincide with the average of treatment and control means of the variables in that index (when divided by their standard deviations).

Self-direction: captures how inclined an individual is to think and act independently. Respondents with a high value of this index like to be curious, creative, free to make their own choices and to rely on themselves.

D. Health camps

We set up 80 'health camps' (one per location) in schools. Respondents were invited to the health camp when they took the follow-up survey. Each health camp was set up for the weekend immediately following the follow-up survey: the data collected at health camps is therefore about 8 months after the baseline survey and about 7.5 months after the last screening of our intervention.

Participants were informed about testing by counsellors and were offered the opportunity to test for Chlamydia through urine sample collection. During the same session they also participated in a game that consisted in choosing between N50 (approximately equivalent to 0.25 USD at the time) and a certain number of condom packs. The number was randomly determined and could vary from 1 to 3, with each pack worth approximately N50 on the market. At the end of the session participants received contact details of HIV counselling and testing centres in their town and were given a voucher that would entitle them to free HIV testing at one of these centers. After the specimens were analyzed and the results for Chlamydia were available, participants were invited for a second visit to the health camp, where they were informed of the outcome and –if they tested positive for Chlamydia– they were prescribed treatment. We did not test anyone for HIV hence do not know who is HIV positive or negative, but only if they took the test.

From our sample, 3828 individuals attended the health camp, and all got tested for Chlamydia and participated in the condom game; 74 of them tested positive for Chlamydia, and 213 redeemed the voucher to get tested for HIV.

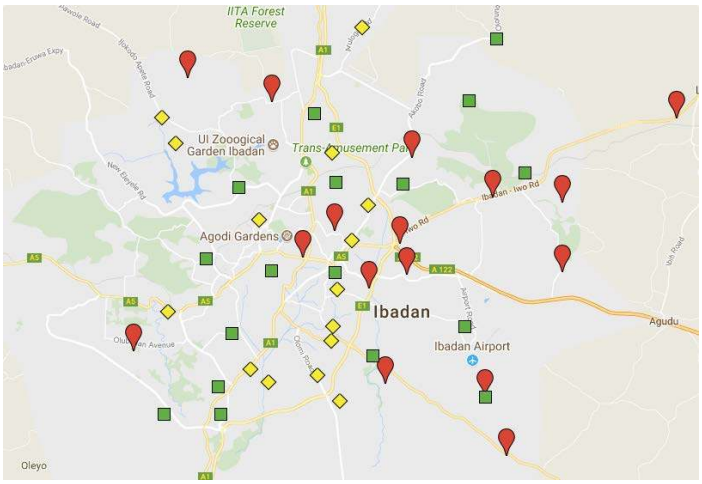
In our analysis we will use the following outcomes 'objectively' collected at health camps: (i) whether participants attended the health camp and took the Chlamydia test; (ii) whether they tested positive for Chlamydia; (iii) whether they redeemed the voucher to get tested for HIV; and (iv) whether they chose condoms over money when given the choice.

Attendance to health camp was relatively high: 77 percent of the study participants attended the health camp, and on average this share was the same in treatment and control locations. Appendix Table A4 shows how baseline characteristics and baseline values of our outcomes correlate with the decision to participate in the health camp. People currently attending school and living outside the family were less likely to attend, possibly due to conflict with school schedules. While treatment status is uncorrelated with the decision to attend (column 1), we also test whether observable characteristics may have played a different role within the treatment and the control samples. We find that higher values of the *HIV testing*

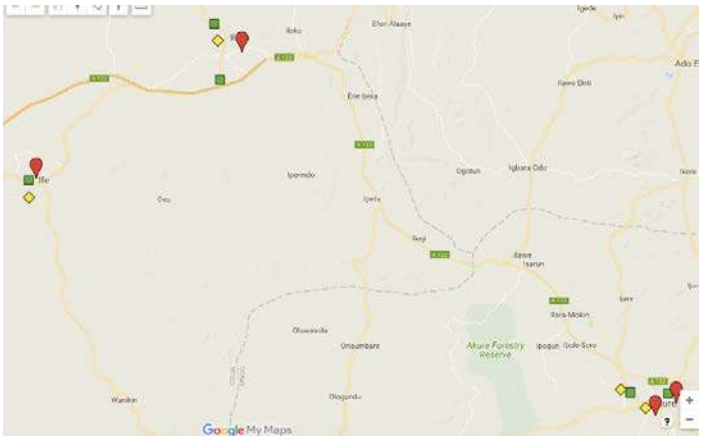
index at baseline predict attendance among treated individuals, while lower values of the *Risky Sexual Behavior index* predict it among control ones. In Appendix Tables A6 and A9 we show that these differences do not explain our estimated impacts on outcomes collected at health camps, as our estimates are robust to including the baseline values of these indexes (and their interaction with treatment) among the regressors.

Appendix Figures and Tables

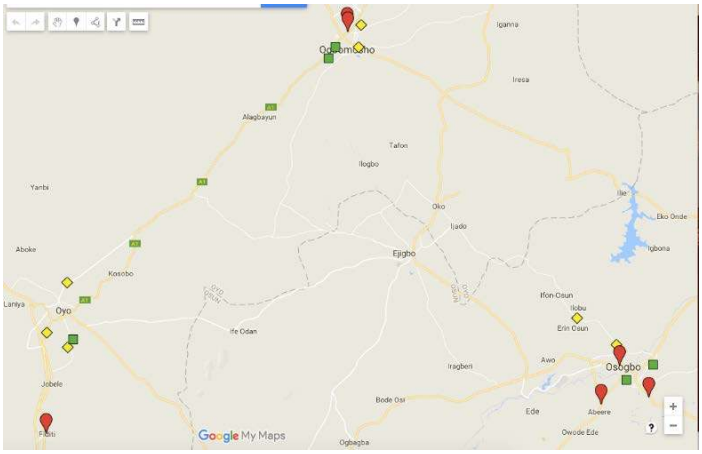
Figure A1: Location of treatment and control centres



(a) Ibadan



(b) Ife, Ilesha, Akure



(c) Oyo, Ogbomosho, Osogbo

Table A1: Summary statistics on invitees to Screening 0 (S0), by participation status

Variable Name	*Mean Did not participate in S0*	*Mean Participated in S0*	*Diff=0 (p-value)*	*Normalized Difference*[(a)]
Female	0.47	0.47	0.65	0.01
Age	20.74	20.60	0.00	-0.04
Highest Educ. Level Attained is Primary	0.01	0.01	0.06	-0.03
Highest Educ. Level Attained is Junior Sec.	0.01	0.01	0.12	-0.02
Highest Educ. Level Attained is Senior Sec.	0.87	0.88	0.12	0.02
Highest Educ. Level Attained > Senior Sec.	0.11	0.11	0.55	-0.01
Speaks mostly English at home	0.07	0.06	0.28	-0.01
First Preferred Language is Yoruba	0.81	0.80	0.11	-0.02
First Preferred Language is English	0.17	0.18	0.03	0.02
Second Most Preferred Language is Yoruba	0.15	0.17	0.00	0.03
Second Most Preferred Language is English	0.78	0.78	0.69	0.00
Number of household members aged 18-25	1.03	1.04	0.00	0.04
Muslim religion	0.39	0.37	0.00	-0.03
No. Obs	10.102	6348		

Notes: Sample includes individuals invited to attend Screening 0, i.e. a movie unrelated to Shuga.

(a) Normalized difference is the difference in the sample means of treatment and control groups divided by the square root of the sum of sample variances.

Table A2: Summary statistics on Screening 0 attendees, by inclusion in baseline sample

	Mean Not selected for baseline (i)	Mean for Baseline Sample (ii)	Normalized Difference[(a)] (ii)-(i)	P-value (i) = (ii)
Variable Name				
Female	0.48	0.47	0.00	0.84
Age	20.46	20.63	0.05	0.02
Highest Educ. Level Attained is Primary	0.01	0.01	-0.04	0.11
Highest Educ. Level Attained is Junior Sec.	0.01	0.00	-0.05	0.03
Highest Educ. Level Attained is Senior Sec.	0.87	0.88	0.02	0.51
Highest Educ. Level Attained > Senior Sec.	0.10	0.11	0.01	0.76
Speaks mostly English at home	0.06	0.06	-0.01	0.81
First Preferred Language is Yoruba	0.84	0.79	-0.09	0.00
First Preferred Language is English	0.15	0.19	0.07	0.00
Second Most Preferred Language is Yoruba	0.15	0.18	0.06	0.02
Second Most Preferred Language is English	0.81	0.77	-0.07	0.00
Number of household members aged 18-25	1.04	1.04	0.02	0.47
Muslim religion	0.36	0.37	0.02	0.29
No. Obs	1.182	5.166		

Notes: Sample includes individuals that attended Screening 0, i.e. a movie unrelated to Shuga.

(a) Normalized difference is the difference in the sample means of treatment and control groups divided by the square root of the sum of sample variances.

Table A3: Variable definitions

Variable name	*Definition*
Panel A: Outcomes	
Indexes	
HIV knowledge	Index of knowledge about transmission, testing and treatment of HIV
HIV respondent's attitudes	Index of attitudes toward HIV positive people
HIV testing	Index of testing behavior
Attitudes towards risky sexual behavior	Index of attitudes towards risky sexual behavior
Risky sexual behavior (for sexually active)	Index of risky sexual behavior (defined for sexually active individuals)
Individual variables	
HIV transmitted during pregnancy	Dummy=1 if knows that HIV can be transmitted during pregnancy
Has heard of ARVs	Dummy=1 if, when specifically asked, respondent says he/she has heard of ARV drugs
Second test necessary	Dummy =1 if knows that a second test is necessary
Window period 3 months	Dummy=1 if knows that a 3 month period is necessary before retest
Can get HIV through intercourse	Dummy=1 if knows that HIV can be contracted via sexual intercourse
Would buy from an HIV+ shopkeeper	Dummy=1 if would buy food from an HIV positive shopkeeper
An HIV+ boy should play footbal	Dummy=1 if agrees that an HIV positive boy should be allowed to play football
People HIV+ should not be blamed	Dummy=1 if agrees that HIV positive people should not be blamed
HIV is not punishment for sleeping around	Dummy=1 if says that HIV is not a punishment for sleeping around
Would reveal HIV status	Dummy=1 if would reveal own HIV status to partner
Tested last 6 months (self-reported)	Dummy=1 if has been tested less than 6 months ago
Tested at health camp (observed)	Dummy=1 if has attended the health camp and has been tested for STDs
Chose condoms over N50	Dummy=1 if chose condoms over money in experimental game at health camp
Tested positive for Chlamydia	Dummy=1 if tested positive for Chlamydia
Men should have one partner only	Dummy=1 if agrees that men should only date one partner at a time
Women should have one partner only	Dummy=1 if agrees that women should only date one partner at a time
Not ok date sugardaddy to finance educ	Dummy=1 if does not consider appropriate dating a sugardaddy, even if he offers to pay for education
Not ok date sugardaddy for money	Dummy=1 if does not consider appropriate dating a sugardaddy in exchange for money
Not ok date sugardaddy to go out	Dummy=1 if does not consider appropriate dating a sugardaddy even if he brings the girl out
If a woman brings a condom does not mean she's not serious	Dummy=1 if disagrees with the statement that if a woman brings a condom, her man will think she's not serious
Has not had multiple concurrent sexual partners	Dummy=1 if has not had multiple concurrent sexual partners
No. of current sexual partners if sexually active	Number of current sexual partners
Used condom the last time he/she had sex	Dummy=1 if used a condom during last sexual intercourse
Panel B: Control variables	
Indexes	
Transportation	Index of immersion in the narrative while watching the show
Identification	Index of identification with the characters
Conformity	Index of propensity to subordinate to norms (e.g., instilled parents or peers)
Tradition	Index of acceptance and commitment to the values that religion or culture promote
Self-direction	Index of inclination to think and act independently
Socioeconomic controls	
Female	Dummy=1 if female
Age	Age of respondent
Currently attending school	Dummy=1 if currently attending school
Years of education	Years of education
Speaks English	Dummy=1 if speaks english as primary or secondary language at home
Single	Dummy=1 if does not have a partner
Does not live with the family	Dummy=1 if does not live with family
Household size	Number of components of respondent's family
Wealth index	Principal component index from dwelling characteristics and durable goods ownership
Owns his/her house	Dummy=1 if dwelling where respondent lives is owned
Father's education > secondary	Dummy=1 if father obtained a level of education higher than secondary
Mother's education > secondary	Dummy=1 if mother obtained a level of education higher than secondary
Muslim	Dummy=1 if Muslim
Yoruba	Dummy=1 if native language is Yoruba

Table A4: Summary statistics at follow-up

	No. Obs.	*Mean*	*Std. Dev.*	*Min*	*Max*
	(1)	(2)	(3)	(4)	(5)
Panel A: Outcome variables					
Indexes					
HIV knowledge	4986	0.585	6.127	-22.006	25.984
HIV attitudes	4986	0.297	3.476	-10.114	7.089
HIV testing	4986	0.093	4.441	-5.052	11.732
Attitudes towards risky sexual behavior	4986	0.085	3.417	-15.113	4.495
Risky sexual behavior (for sexually active)	3618	0.067	3.525	-31.687	5.1
Individual variables					
HIV transmitted during pregnancy	4986	0.669	0.471	0	1
Mentions drugs to live longer with HIV	4986	0.739	0.438	0	1
Recognizes ARV when mentioned by enumerator	4986	0.332	0.471	0	1
Has heard of ARVs	4986	0.292	0.447	0	1
Second test necessary	4986	0.372	0.484	0	1
Window period 3 months	4986	0.16	0.367	0	1
Can get HIV through intercourse	4986	0.976	0.153	0	1
Would buy from an HIV+ shopkeeper	4986	0.521	0.5	0	1
An HIV+ boy should play footbal	4986	0.696	0.46	0	1
People HIV+ should not be blamed	4986	0.693	0.461	0	1
HIV is not punishment for sleeping around	4986	0.52	0.5	0	1
Would reveal HIV status	4986	0.727	0.446	0	1
Tested last 6 months (self-reported)	4982	0.103	0.304	0	1
Tested at health camp (observed)	3828	0.056	0.229	0	1
Men should have one partner only	4976	0.894	0.308	0	1
Women should have one partner only	4986	0.926	0.262	0	1
Not ok date sugardaddy to finance educ	4986	0.744	0.437	0	1
Not ok date sugardaddy for money	4986	0.696	0.46	0	1
Not ok date sugardaddy to go out	4986	0.905	0.293	0	1
If a woman brings a condom it does not mean she's not serious	4986	0.625	0.484	0	1
Has not had multiple concurrent sexual partners	3339	0.803	0.398	0	1
No. of current sexual partners if sexually active	4978	0.673	0.867	0	21
Used condom the last time he/she had sex	3084	0.489	0.5	0	1
Chose condoms over N50	3827	0.393	0.488	0	1
Tested positive for Chlamydia	4986	0.302	0.459	0	1
Panel B: Control variables					
Indexes					
Transportation	3753	0	1.945	-9.074	3.422
Identification	3753	0	2.051	-8.651	3.621
Conformity	5166	0.036	1.31	-7.009	1.515
Tradition	5166	0.026	1.276	-7.241	1.911
Self-direction	5166	0.004	1.449	-1.846	6.152
Socioeconomic controls [(a)]					
Female	5166	0.473	0.499	0	1
Age	5166	20.615	2.362	18	26
Currently attending school	5166	0.348	0.476	0	1
Years of education	5166	11.597	1.085	0	12
Speaks English	5166	0.961	0.194	0	1
Single	5166	0.227	0.419	0	1
Does not live with the family	5166	0.241	0.428	0	1
Household size	5166	4.328	2.362	1	19
Wealth index	5150	1.75	0.587	0	3
Owns his/her house	5165	0.385	0.487	0	1
Father's education > secondary	3928	0.334	0.472	0	1
Mother's education > secondary	4393	0.226	0.418	0	1
Muslim	5166	0.373	0.484	0	1
Yoruba	5166	0.92	0.271	0	1

Notes. (a) Summary statistics for these variables are calculated at baseline, because this is how they are included in all regressions.

Table A5: Attrition between baseline and follow-up

Dep. Var. =1 if interviewed at follow-up	(1)	(2)	(3)	(4)	(5)
Treated	-0.004	-0.004	-0.004	-0.004	-0.000
	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)
HIV knowledge	-0.000				
	(0.000)				
HIV attitudes		-0.001			
		(0.001)			
HIV testing			0.001		
			(0.001)		
Attitudes towards risky sexual behavior				-0.000	
				(0.001)	
Risky sexual behavior (for sexually active)					-0.001
					(0.001)
Female	-0.022***	-0.022***	-0.023***	-0.022***	-0.029***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)
Age	0.001	0.001	0.001	0.001	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Currently attending school	-0.013**	-0.013**	-0.013**	-0.013**	-0.016*
	(0.006)	(0.006)	(0.006)	(0.006)	(0.008)
Years of education	0.001	0.001	0.000	0.000	0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
English Spoken	0.002	0.002	0.002	0.002	0.009
	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)
Single	-0.003	-0.003	-0.003	-0.003	0.006
	(0.005)	(0.005)	(0.005)	(0.005)	(0.009)
Does not live with the family	-0.014	-0.014	-0.014	-0.014	-0.021**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.011)
Household size	0.000	0.000	0.000	0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Wealth index	0.008	0.007	0.007	0.007	0.010
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
Home owner	-0.004	-0.004	-0.004	-0.004	-0.015*
	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
Father's education > secondary	0.010**	0.009**	0.009**	0.009**	0.009
	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)
Mother's education > secondary	0.005	0.005	0.004	0.004	0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.007)
Muslim	0.012*	0.012**	0.012**	0.012**	0.016**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.008)
Yoruba	0.018	0.018	0.018	0.018	0.006
	(0.013)	(0.013)	(0.013)	(0.013)	(0.014)
Constant	0.948***	0.949***	0.963***	0.956***	0.914***
	(0.041)	(0.042)	(0.043)	(0.043)	(0.049)
	5,166	5,166	5,166	5,166	3,246
R-squared	0.094	0.094	0.094	0.093	0.080
Controls	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var. (Control)	0.966	0.966	0.966	0.966	0.966

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. The dependent variable is an indicator for whether the individual has been interviewed at follow-up.

Table A6: HIV indexes

	Sign with which variable enters index	***Loading factor***
HIV knowledge		
# of correct sources of contagion listed	+	0.338
Can get HIV through intercourse	+	0.164
Has mentioned drugs to live longer with HIV (not ARVs)	+	0.155
Has mentioned ARV	+	0.206
Recognizes ARV when mentioned by enumerator	+	0.288
# of correct ways to avoid contracting HIV listed	+	0.312
Avoid HIV knowing your/your partner's status	+	0.048
Window period 3 months	+	0.343
Knows that an early negative test is no guarantee of no HIV	+	0.364
Second test necessary	+	0.396
HIV transmitted during pregnancy	+	0.235
Says exist drugs to reduce transmission risk to baby	+	0.250
Says HIV can be transmitted from mother to baby during delivery	+	0.237
Says HIV can be transmitted from mother to baby by breastfeeding	+	0.167
HIV attitudes		
Would not prefer to keep HIV of family member a secret	+	-0.066
Would reveal HIV status	+	0.228
Would buy from an HIV+ shopkeeper	+	0.389
An HIV+ boy should play footbal	+	0.406
If a young person get tested for HIV, he has been sleeping around	+	0.441
People with HIV should be blamed	+	0.490
HIV/AIDS is a punishment for sleeping around.	+	0.439
HIV testing		
Tested for HIV at least once	+	0.501
Tested last 12 months (self reported)	+	0.467
Tested last 6 months (self-reported)	+	0.371
Asked him(her)self for the test	+	0.345
Tested and picked up results	+	0.493
Knows a place to get HIV test	+	0.174

Notes: Shaded cells refer to ordinal variables. The variables are coded so that higher values correspond to higher levels of disagreement.

Table A7: Indexes of risky sexual behavior

Attitudes towards risky sex	***Sign with which variable enters index***	***Loading factor***
"Men who are not married should not only have sex with one partner"	+	0.227
"Women who are not married should not only have sex with one partner"	+	0.231
"OK for a young girl to date an older married man if he offers to pay for her educaton"	+	0.555
"OK for a young girl to date an older married man if her family needs financial support"	+	0.557
"OK for a young girl to date an older married man if he offers to take her out"	+	0.513
If a woman brings a condom it does not mean she's no serious	+	0.113
Risky sexual behavior		
Not multiple concurrent sexual partners	+	0.506
In the last 6 months had only one partner in the same month	+	0.464
Number of current sexual partners	-	-0.475
Used condom last time he/she had sex	+	-0.070
Has a main partner	+	0.198
Has not an "other"partner	+	0.510

Notes: Shaded cells refer to ordinal variable. The variables are coded so that higher values correspond to higher levels of disagreement.

Table A8: Correlates of Health Camp attendance

	Full sample			*Treated*			*Control*		
Dep. Var=1 if attended health camp	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	0.000	-0.001	0.005						
	(0.015)	(0.015)	(0.019)						
HIV knowledge		-0.001			-0.000			-0.003*	
		(0.001)			(0.001)			(0.002)	
HIV attitudes		0.003			0.002			0.003	
		(0.002)			(0.002)			(0.004)	
HIV testing		0.001			0.002			-0.002	
		(0.001)			(0.002)			(0.003)	
Attitudes towards risky sexual behavior			0.003			0.003			0.004
			(0.002)			(0.002)			(0.004)
Risky sexual behavior (sexually active)			0.001			0.003			-0.003
			(0.002)			(0.003)			(0.002)
Female	0.000	0.001	-0.029	0.016	0.015	-0.019	-0.028	-0.026	-0.040
	(0.015)	(0.015)	(0.018)	(0.019)	(0.019)	(0.022)	(0.028)	(0.028)	(0.033)
Age	0.003	0.002	0.005	0.003	0.002	0.005	0.003	0.003	0.004
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)	(0.006)
Currently attending school	-0.057***	-0.058***	-0.063***	-0.058***	-0.060***	-0.057***	-0.058**	-0.057**	-0.081**
	(0.014)	(0.014)	(0.016)	(0.017)	(0.017)	(0.018)	(0.023)	(0.024)	(0.036)
Years of education	-0.013**	-0.013**	-0.005	-0.009	-0.010	-0.002	-0.020*	-0.019*	-0.012
	(0.006)	(0.006)	(0.007)	(0.008)	(0.008)	(0.009)	(0.010)	(0.010)	(0.011)
Speaks English	-0.012	-0.009	-0.033	-0.023	-0.024	-0.046	-0.004	0.003	-0.016
	(0.032)	(0.032)	(0.031)	(0.044)	(0.044)	(0.043)	(0.040)	(0.040)	(0.038)
Single	-0.024	-0.024	-0.059***	-0.026	-0.027	-0.058**	-0.020	-0.020	-0.061
	(0.015)	(0.015)	(0.021)	(0.020)	(0.020)	(0.026)	(0.024)	(0.024)	(0.038)
Does not live with the family	-0.153***	-0.151***	-0.144***	-0.155***	-0.154***	-0.157***	-0.150***	-0.150***	-0.115***
	(0.021)	(0.022)	(0.022)	(0.027)	(0.027)	(0.029)	(0.036)	(0.036)	(0.033)
Household size	0.013***	0.013***	0.014***	0.011***	0.011***	0.011**	0.015***	0.014***	0.018**
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.007)
Wealth index	-0.006	-0.006	-0.007	0.004	0.002	0.006	-0.030	-0.026	-0.040*
	(0.014)	(0.014)	(0.015)	(0.018)	(0.018)	(0.019)	(0.021)	(0.021)	(0.023)
Owns his/her house	0.011	0.012	0.013	0.012	0.013	0.010	0.011	0.013	0.018
	(0.015)	(0.015)	(0.018)	(0.019)	(0.019)	(0.022)	(0.024)	(0.024)	(0.030)
Father obtained education higher then sec.	-0.032*	-0.032*	-0.024	-0.043**	-0.044**	-0.035	-0.012	-0.010	0.001
	(0.017)	(0.017)	(0.019)	(0.021)	(0.021)	(0.024)	(0.030)	(0.031)	(0.031)
Mother obtained education higher then sec.	0.024	0.024	0.028	0.021	0.019	0.014	0.036	0.038	0.058
	(0.018)	(0.018)	(0.022)	(0.020)	(0.020)	(0.023)	(0.035)	(0.035)	(0.049)
Muslim	0.024*	0.023*	0.027*	0.021	0.021	0.025	0.022	0.019	0.022
	(0.013)	(0.013)	(0.015)	(0.015)	(0.015)	(0.016)	(0.031)	(0.030)	(0.036)
Native language Yoruba	0.069***	0.068***	0.075***	0.042	0.042	0.039	0.125***	0.124***	0.139**
	(0.026)	(0.026)	(0.027)	(0.034)	(0.034)	(0.034)	(0.040)	(0.040)	(0.051)
Sample	All	All	Sexually active	All	All	Sexually active	All	All	Sexually active
Observations	4,986	4,986	3,618	3,402	3,402	2,487	1,584	1,584	1,131
R-squared	0.067	0.068	0.071	0.064	0.065	0.068	0.085	0.088	0.100
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var	0.773		0.767	0.765		0.763	0.773		0.767

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level . 80 clusters in the Full Sample, 54 clusters in the Treated Sample, 26 clusters in the Control Sample. ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. All regressors are measured at baseline.

Table A9: Exogeneity of treatment assignment

Dep. Var. =1 if Treated.					
Panel A: Controls	*Coeff.*	*Std. Err.*			
Female	0.003	(0.013)			
Age	-0.002	(0.004)			
Currently attending school	0.019	(0.022)			
Years of education	0.000	(0.008)			
English Spoken	0.034	(0.038)			
Single	-0.005	(0.022)			
Does not live with the family	-0.038	(0.032)			
Household size	-0.005	(0.004)			
Wealth index	-0.013	(0.027)			
Home owner	-0.089**	(0.039)			
Father obtained education higher then sec.	-0.048***	(0.017)			
Mother obtained education higher then sec.	-0.019	(0.021)			
Muslim	0.008	(0.032)			
Yoruba Native	0.012	(0.052)			
Constant	0.784***	(0.172)			
Observations	5,166				
R-squared	0.021				
P-val F-test of joint significance	0.147				
Panel B: Outcomes	(1)	(2)	(3)	(4)	(5)
HIV knowledge	0.002				
	(0.001)				
HIV attitudes		-0.000			
		(0.003)			
HIV testing			0.001		
			(0.002)		
Attitudes towards risky sexual behavior				-0.000	
				(0.002)	
Risky sexual behavior (for sexually active)					0.002
					(0.002)
Constant	0.820***	0.782***	0.796***	0.784***	0.745***
	(0.168)	(0.171)	(0.168)	(0.171)	(0.176)
Observations	5,166	5,166	5,166	5,166	3,246
R-squared	0.022	0.021	0.021	0.021	0.025
Controls [(a)]	Yes	Yes	Yes	Yes	Yes
P-val F-test of joint significance	0.156	0.159	0.188	0.142	0.304

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. The dependent variable takes value one if the individual has been assigned to treatment.

(a) Controls in each regression of panel B include: female, age, currently attending school, years of education, english spoken, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table A10: Impact on Indexes calculated with principal component method

Dep. Var. (Y_t):	*HIV knowledge*			*HIV attitudes*			*HIV testing*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Full Sample	*Females*	*Males*	*Full Sample*	*Females*	*Males*	*Full Sample*	*Females*	*Males*
Treated	0.184***	0.302***	0.089	0.177***	0.156***	0.207***	0.146**	0.161*	0.131
	(0.055)	(0.079)	(0.083)	(0.042)	(0.059)	(0.055)	(0.057)	(0.085)	(0.083)
	[0.002]	[0.001]	[0.288]	[0.000]	[0.021]	[0.001]	[0.011]	[0.069]	[0.225]
Treated*Y_{t-1}	-0.066*	-0.016	-0.114**	-0.011	-0.006	-0.014	-0.028	-0.039	-0.035
	(0.039)	(0.049)	(0.045)	(0.040)	(0.051)	(0.053)	(0.033)	(0.042)	(0.050)
	[0.257]	[0.932]	[0.043]	[0.787]	[0.932]	[0.798]	[0.630]	[0.725]	[0.731]
Y_{t-1}	0.433***	0.420***	0.442***	0.436***	0.487***	0.391***	0.543***	0.591***	0.492***
	(0.031)	(0.041)	(0.033)	(0.032)	(0.040)	(0.046)	(0.026)	(0.034)	(0.038)
Observations	4,986	2,323	2,663	4,986	2,323	2,663	4,971	2,320	2,651
R-squared	0.208	0.240	0.204	0.188	0.247	0.157	0.279	0.378	0.195
Controls[(a)]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep var (Control)	0.189	0.177	0.199	0.119	0.188	0.0594	0.0895	0.236	-0.0391
P-value test joint sig	0.000	0.000	0.138	0.000	0.012	0.001	0.008	0.064	0.059

Dep. Var. (Y_t):	*Attitudes towards risky sexual behavior*			*Risky sexual behavior (for sexually active)*		
	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample	*Females*	*Males*	*Full Sample*	*Females*	*Males*
Treated	0.042	-0.025	0.112**	0.060	0.151	0.029
	(0.036)	(0.056)	(0.056)	(0.075)	(0.166)	(0.140)
	[0.432]	[0.661]	[0.096]	[0.432]	[0.622]	[0.836]
Treated*Y_{t-1}	0.017	0.007	0.022	-0.166**	-0.263*	-0.150**
	(0.031)	(0.046)	(0.042)	(0.064)	(0.154)	(0.075)
	[0.571]	[0.887]	[0.607]	[0.026]	[0.236]	[0.098]
Y_{t-1}	0.278***	0.295***	0.260***	0.396***	0.386***	0.405***
	(0.024)	(0.037)	(0.036)	(0.054)	(0.145)	(0.064)
Observations	4,973	2,320	2,653	1,682	760	922
R-squared	0.113	0.147	0.101	0.265	0.146	0.125
Controls [(a)]	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var. (Control)	0.0858	0.144	0.0346	-0.0488	0.837	-0.753
P-value test joint sig	0.254	0.647	0.0418	0.370	0.509	0.339

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. P-values in square brackets corrected for multiple hypothesis testing using FWER. Increasing values of the dependent variable correspond to outcomes more consistent with the messages of Shuga.

(a) Controls in each regression include: female, age, currently attending school, years of education, english spoken, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table A11: Impact on outcome Indexes, ANCOVA

Dep. Var. (Y_t):	*HIV knowledge*	*HIV attitudes*	*HIV testing*	*Attitudes towards risky sexual behavior*	*Risky sexual behavior (for sexually active)*
	(1)	(2)	(3)	(4)	(5)
Treated	0.781***	0.339***	0.336**	0.148	0.150
	(0.214)	(0.103)	(0.128)	(0.090)	(0.146)
Y_{t-1}	0.346***	0.366***	0.449***	0.292***	0.241***
	(0.020)	(0.019)	(0.016)	(0.017)	(0.028)
Observations	4,986	4,986	4,986	4,986	3,070
R-squared	0.197	0.179	0.261	0.099	0.149
Controls[(a)]	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var. (Control)	0.0400	0.0439	-0.139	0.00186	-0.0631
Std Dev Dep. Var (Control)	6.106	3.474	4.145	3.452	3.625

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. FWER adjusted p-values in square brackets, based on 10,000 replications. Increasing values of the dependent variable correspond to outcomes more consistent with the messages of Shuga.

(a) Controls in each regression include: female, age, currently attending school, years of education, english spoken, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table A12: Impact on outcome indexes, by gender

Dep. Var. (Y_t):	*HIV knowledge*		*HIV attitudes*		*HIV testing*		*Attitudes towards risky sexual behavior*		*Risky sexual behavior (for sexually active)*	
	Females	*Males*	*Females*	*Males*	*Females*	*Males*	*Females*	*Males*	*Females*	*Males*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Cross-section										
Treated	1.191***	0.674**	0.201	0.480***	0.440*	0.283	-0.133	0.426***	-0.170	0.335
	(0.361)	(0.316)	(0.183)	(0.172)	(0.248)	(0.178)	(0.167)	(0.161)	(0.121)	(0.220)
	[0.005]	[0.076]	[0.265]	[0.021]	[0.152]	[0.120]	[0.422]	[0.022]	[0.296]	[0.136]
Observations	2323	2663	2323	2663	2323	2663	2323	2663	1526	2092
R-squared	0.097	0.098	0.078	0.044	0.144	0.057	0.046	0.022	0.037	0.018
Panel B: Conditional specification										
Treated	1.174***	0.462	0.275*	0.408***	0.421**	0.260	-0.074	0.363**	-0.118	0.293
	(0.313)	(0.285)	(0.140)	(0.137)	(0.199)	(0.179)	(0.138)	(0.152)	(0.199)	(0.244)
	[0.001]	[0.204]	[0.074]	[0.012]	[0.074]	[0.204]	[0.811]	[0.037]	[0.811]	[0.230]
Treated*Y_{t-1}	-0.023	-0.106***	0.020	-0.002	-0.053	-0.033	-0.024	0.012	-0.060	-0.104
	(0.049)	(0.039)	(0.039)	(0.052)	(0.040)	(0.046)	(0.048)	(0.048)	(0.096)	(0.067)
	[0.854]	[0.025]	[0.854]	[0.964]	[0.469]	[0.723]	[0.792]	[0.798]	[0.792]	[0.233]
Y_{t-1}	0.379***	0.399***	0.389***	0.334***	0.523***	0.425***	0.311***	0.279***	0.228***	0.326***
	(0.042)	(0.028)	(0.029)	(0.046)	(0.033)	(0.036)	(0.039)	(0.041)	(0.085)	(0.057)
R-squared	0.225	0.194	0.238	0.146	0.361	0.181	0.126	0.094	0.080	0.079
Controls[(a)]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2323	2663	2323	2663	2323	2663	2323	2663	1287	1783
Mean Dep. Var. (Control)	0.001	0.0741	0.00715	0.0760	0.182	-0.420	0.208	-0.178	1.353	-1.078
Std dev Dep. Var. (Control)	5.944	6.248	3.222	3.683	4.186	4.091	3.304	3.568	2.090	4.119

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectivelyì. FWER adjusted p-values in square brackets, based on 10,000 replication. Increasing values of the dependent variable correspond to outcomes more consistent with the messages of Shuga.

(a) Controls in each regression include: female, age, currently attending school, years of education, english spoken, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table A13: Impact on selected HIV outcomes

Coefficient on:	CROSS-SECTION		CONDITIONAL SPECIFICATION						
	Treated		*Treated*		*Treated*$*Y_{t-1}$				
	Coeff (Std.Err)	p-value FWER	Coeff (Std.Err)	p-value FWER	Coeff (Std.Err)	p-value FWER	p-value (3)+(5)	No. Obs	Mean in control group
Dep. Var. (Y_t):	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Tested for HIV (observed) (a)	0.031**							3828	0.033
	(0.013)	[0.014]							
Tested last 6 months (self-reported)	0.025***		0.026***		-0.038		0.005	4971	0.086
	(0.009)	[0.014]	(0.009)	[0.004]	(0.063)	[0.550]			
HIV transmitted during pregnancy	0.069***		0.127***		-0.098***		0.000	4986	0.621
	(0.019)	[0.002]	(0.033)	[0.002]	(0.032)	[0.025]			
Has mentioned ARV drugs spontaneously or when mentioned by enumerator	0.037**		0.032**		-0.019		0.486	4986	0.047
	(0.017)	[0.041]	(0.015)	[0.045]	(0.036)	[0.623]			
Second test necessary	0.047***		0.055		-0.039		0.003	4986	0.343
	(0.015)	[0.013]	(0.018)	[0.014]	(0.037)	[0.726]			
Window period 3 months	0.049***		0.045***		-0.009		0.000	4986	0.129
	(0.012)	[0.001]	(0.012)	[0.003]	(0.062)	[0.955]			
Can get HIV through intercourse	0.01*		-0.037		0.050		0.285	4986	0.969
	(0.006)	[0.136]	(0.034)	[0.500]	(0.034)	[0.552]			
Would buy from an HIV+ shopkeeper	0.047***		0.082***		-0.094***		0.000	4986	0.487
	(0.015)	[0.013]	(0.020)	[0.000]	(0.032)	[0.024]			
An HIV+ boy should play football	0.051***		0.080***		-0.045		0.006	4986	0.662
	(0.016)	[0.012]	(0.028)	[0.022]	(0.036)	[0.581]			
People HIV+ should not be blamed	0.022		0.045		-0.031		0.240	4986	0.676
	(0.018)	[0.422]	(0.037)	[0.417]	(0.041)	[0.701]			
HIV is not punishment for sleeping around	0.045**		0.027		0.021		0.063	4986	0.486
	(0.023)	[0.137]	(0.028)	[0.417]	(0.034)	[0.701]			
Would reveal HIV status to partner	0.015		0.049*		-0.044		0.098	4986	0.713
	(0.013)	[0.422]	(0.029)	[0.263]	(0.034)	[0.581]			

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. FWER adjusted p-values in square brackets, adjusted within family of variables (e.g., knowledge, attitudes, testing) and based on 10,000 replications. "P-Value (3)+(5)" tests the null hypothesis Treated + Treated*Y_{t-1}m =0, where Y_{t-1}m denote the sample mean of Y_{t-1} at baseline. Estimates in cols. 1-2 come from the cross sectional model; estimates in cols. 3-5 from the model that includes the lagged dependent variable and its interaction with treatment. Col. 9 reports the mean of the dependent variable at follow-up. All regressions include the following controls: female, age, currently attending school, years of education, English spoken at home, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as native language.

(a) For this outcome we cannot estimate the model including the lag of the dependent variable because teasting through health-camps was only offered at follow-up

Table A14: Treatment Effect on the probability of HIV

]

Dep. Var:	*Prob. of being currently HIV positive*			*Prob. that partner is HIV positive*			*Prob. that a person your age in the community has HIV*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	0.000	0.001	0.000	-0.006	-0.005	-0.002	0.027**	0.024**	0.034**
	(0.002)	(0.002)	(0.003)	(0.005)	(0.005)	(0.005)	(0.011)	(0.011)	(0.013)
	[0.852]	[0.721]	[0.947]	[0.410]	[0.528]	[0.928]	[0.055]	[0.070]	[0.035]
Y_{t-1}		0.088***	0.073**		0.110***	0.142***		0.116***	0.137***
		(0.017)	(0.028)		(0.019)	(0.027)		(0.016)	(0.030)
Treated*Y_{t-1}			0.025			-0.047			-0.031
			(0.034)			(0.037)			(0.037)
			[0.654]			[0.531]			[0.654]
Observations	4.986	4.986	4.986	4.066	3.452	3.452	4.986	4.986	4.986
R-squared	0.006	0.015	0.016	0.018	0.031	0.031	0.021	0.034	0.034
Controls$^{(a)}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var. (Control)	0.019	0.019	0.019	0.058	0.055	0.055	0.337	0.337	0.337

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. FWER adjusted p-values in square brackets, based on 10,000 replications.

(a) Controls in each regression include: female, age, currently attending school, years of education, english spoken at home, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table A15: Heterogeneous effects by stability of partnerr

Dep. Var.:	*HIV knowledge*	*HIV attitudes*	*HIV testing*	*Attitudes towards risky sexual behavior*	*Risky sexual behavior (for sexually active)*
	(1)	(2)	(3)	(4)	(5)
Treated	0.690	0.702	0.999**	0.615	0.081
	(0.708)	(0.434)	(0.395)	(0.414)	(0.561)
	[0.339]	[0.208]	[0.044]	[0.276]	[0.884]
Treated * Stable Partnert-1	-0.045	-0.496	-0.927**	-0.623	0.027
	(0.757)	(0.456)	(0.401)	(0.426)	(0.614)
	[0.956]	[0.485]	[0.071]	[0.281]	[0.947]
Stable Partnert-1	-0.489	0.332	0.879***	0.565	-0.745
	(0.610)	(0.395)	(0.243)	(0.341)	(0.549)
Treated*Y_{t-1}	-0.043	-0.016	-0.011	0.064	-0.110*
	(0.047)	(0.045)	(0.037)	(0.046)	(0.062)
	[0.750]	[0.926]	[0.926]	[0.165]	[0.162]
Y_{t-1}	0.371***	0.365***	0.436***	0.260***	0.339***
	(0.039)	(0.040)	(0.031)	(0.036)	(0.050)
Observations	2,827	2,827	2,827	2,827	2,777
R-squared	0.195	0.179	0.250	0.109	0.160
Controls[(a)]	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var. (Control)	0.040	0.0439	-0.139	0.00186	-0.0631
Std Dev Dep. Var. (Control)	6.106	3.474	4.145	3.452	3.625

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. Increasing values of the dependent variable correspond to outcomes more consistent with the messages of Shuga. FWER adjusted p-values in square brackets, based on 10,000 replications

(a) Controls in each regression include: female, age, currently attending school, years of education, english spoken, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks yoruba as a native language.

The dummy "Stable partner" is measured at baseline and takes value one if the respondent reports being married or living with someone or having a main partner.

Table A16: Impact on selected risky sexual behavior outcomes

Coefficient on:	CROSS-SECTION		CONDITIONAL SPECIFICATION						
	Treated		*Treated*		*Treated*Y_{t-1}*				
	Coeff (Std.Err)	P-Value FWER	Coeff (Std.Err)	P-Value FWER	Coeff (Std.Err)	P-Value FWER	P-Value (3)+(5)	No. Obs	Mean in control group
Dep. Var. (Y_t):	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Men should have one partner only	0.036***		0.074**		-0.048		0.048	4973	0.867
	(0.012)	[0.013]	(0.037)	[0.100]	(0.036)	[0.347]			
Women should have one partner only	0.027***		0.103***		-0.088**		0.004	4986	0.906
	(0.010)	[0.021]	(0.035)	[0.014]	(0.036)	[0.051]			
Has not had multiple concurrent sexual partners	0.027*		0.113***		-0.103**		0.007	2623	0.780
	(0.015)	[0.232]	(0.041)	[0.018]	(0.041)	[0.028]			
Number of current sexual partners if sexually active	-0.051		0.215**		-0.222***		0.014	1850	1.323
	(0.034)	[0.232]	(0.086)	[0.049]	(0.07)	[0.020]			
If a woman brings condom does not mean she's not serious	0.014		0.012		-0.004		0.640	4986	0.613
	(0.015)	[0.347]	(0.026)	[0.636]	(0.030)	[0.901]			
Used condom the last time he/she had sex	-0.003		-0.021		0.038		0.513	2308	0.494
	(0.019)	[0.844]	(0.032)	[0.506]	(0.053)	[0.574]			
Have you ever used a condom?	0.018		0 .069*		-0.088**		0.054	3070	0.737
	(0.014)	[0.495]	(0.035)	[0.153]	(0.039)	[0.086]			
Are you currently using condoms?	-0.014		-0.027		0.01		0.508	2990	0.775
	(0.015)	[0.588]	(0.04)	[0.759]	(0.047)	[0.972]			
The first time you will have sex, do you plan to use a condom?	0.012		0.003		0.006		0.875	1311	0.168
	(0.02)	[0.588]	(0.019)	[0.874]	(0.077)	[0.972]			

Notes: Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. FWER adjusted p-values in square brackets based on 10,000 replications. "P-Value (5)+(6)" tests the null hypothesis Treated+ Treated*Y_{t-1}m =0, where by Y_{t-1}m we denote the sample mean of Y_{t-1} at baseline.

Controls in each regression include: female, age, currently attending school, years of education, english spoken at home, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language

Table A17: Involvement in the narrative, loading factors

	Loading factor
Transportation index	
You could easily imagine what was going on in the show.	0.3055
You were not distracted by activities in the room around you.	0.1387
You could imagine yourself being part of the story.	0.2800
You were really following the story.	0.3851
You wanted to learn how the story ended.	0.3673
It affected you emotionally.	0.3113
You were thinking of ways the story could have ended in a different way.	0.3146
While watching the show, you did not found your mind wandering.	0.0933
You found the story relevant to your everyday life.	0.2632
You had a clear picture of the characters in the story.	0.3504
You did not found it easy to forget about it.	0.1667
You feel the story has changed your life.	0.3183
Identification index	
While viewing the show, you felt as if you were part of the action	0.3174
While viewing the show, you forgot yourself and you were fully absorbed	0.2869
You were able to understand the events in the show like the characters understood them	0.3607
You have a good understanding of the characters	0.3688
You understand the reasons why the characters did what they did.	0.3510
While viewing the show you could feel the emotions the characters displayed.	0.3385
During the show, you felt you could read the characters? minds.	0.2808
At key moments in the show, you felt you had experienced the same thing that the characters were going through	0.2117
While viewing the show, you wanted the characters to succeed in achieving their goals.	0.3156
When the characters succeeded you felt joy, but when they failed, you felt sad.	0.2996

Table A18: Correlates of Transportation and Identification

Dep. Var.:	*Transportation Index*		*Identification Index*	
	Coeff.	*Std. Err.*	*Coeff.*	*Std. Err.*
Female	-0.010	(0.088)	-0.262***	(0.098)
Age	0.016	(0.015)	0.018	(0.015)
Currently attending school	0.065	(0.077)	-0.048	(0.072)
Years of education	-0.060	(0.040)	-0.072*	(0.040)
English spoken at home	0.237	(0.162)	0.374**	(0.151)
Single	-0.036	(0.083)	0.060	(0.085)
Does not live with the family	-0.055	(0.105)	0.048	(0.120)
Household size	-0.006	(0.017)	-0.016	(0.020)
Wealth index	0.116*	(0.063)	0.157**	(0.074)
Home owner	-0.087	(0.104)	-0.126	(0.111)
Father obtained education higher then sec.	-0.029	(0.087)	0.051	(0.087)
Mother obtained education higher then sec.	-0.084	(0.100)	-0.123	(0.102)
Muslim	0.015	(0.086)	-0.100	(0.084)
Yoruba	0.043	(0.128)	0.126	(0.146)
Constant	0.009	(0.520)	0.026	(0.651)
Observations	3,753		3,753	
R-squared	0.019		0.030	
P-value F-test of joint significance	0.391		0.001	

Notes: Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. Increasing values of the dependent variable correspond to outcomes more consistent with the messages of Shuga.

Table A19: Transportation and Identification, cross sectional specification

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable:	*HIV knowledge*		*HIV attitudes*		*HIV testing*		*Attitudes towards risky sexual behavior*		*Risky sexual behavior (for sexually active)*	
Panel A: Transportation Index										
Treated*Transportation	0.491***	0.492***	0.204***	0.207***	0.140	0.140	0.351***	0.357***	0.090	0.082
	(0.109)	(0.110)	(0.073)	(0.072)	(0.085)	(0.086)	(0.073)	(0.074)	(0.063)	(0.065)
	[0.000]	[0.000]	[0.015]	[0.013]	[0.112]	[0.098]	[0.000]	[0.000]	[0.168]	[0.208]
Treated	1.145***	-7.852**	0.358**	-1.990	0.332*	-1.364	-0.012	-1.983	0.045	-2.862
	(0.262)	(3.824)	(0.164)	(1.975)	(0.186)	(2.349)	(0.120)	(1.982)	(0.128)	(1.988)
	[0.000]	[0.121]	[0.054]	[0.549]	[0.073]	[0.562]	[0.915]	[0.319]	[0.915]	[0.269]
Transportation	-0.047	-0.051	0.001	-0.003	-0.011	-0.015	-0.021	-0.021	-0.057	-0.050
	(0.077)	(0.077)	(0.046)	(0.045)	(0.056)	(0.055)	(0.052)	(0.053)	(0.043)	(0.045)
Observations	3,753	3,753	3,753	3,753	3,753	3,753	3,753	3,753	2,667	2,667
R-squared	0.103	0.107	0.056	0.058	0.092	0.097	0.041	0.047	0.102	0.107
Controls[(a)]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls*Treated	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean Dep. Var. (Control)	0.0400	0.0400	0.0439	0.0439	-0.139	-0.139	0.00186	0.00186	-0.0631	-0.0631
Std Dev Dep. Var. (Control)	6.106	6.106	3.474	3.474	4.145	4.145	3.452	3.452	3.625	3.625
P-value test joint sig	3.77e-05	0.0434	0.0325	0.317	0.0781	0.563	0.919	0.320	0.727	0.154
Panel B: Identification Index										
Treated*Identification	0.310***	0.322***	0.124*	0.118*	0.140*	0.143*	0.181**	0.174**	0.028	0.014
	(0.097)	(0.096)	(0.064)	(0.064)	(0.076)	(0.077)	(0.069)	(0.069)	(0.055)	(0.058)
	[0.006]	[0.006]	[0.117]	[0.139]	[0.117]	[0.139]	[0.024]	[0.033]	[0.612]	[0.823]
Treated	1.172***	-7.976**	0.377**	-2.060	0.333*	-1.398	0.011	-2.108	0.043	-2.882
	(0.256)	(3.830)	(0.161)	(1.994)	(0.190)	(2.353)	(0.121)	(2.007)	(0.130)	(1.977)
	[0.000]	[0.122]	[0.042]	[0.516]	[0.083]	[0.541]	[0.935]	[0.313]	[0.935]	[0.267]
Identification	-0.006	-0.014	0.012	0.014	-0.009	-0.015	0.018	0.026	-0.035	-0.022
	(0.072)	(0.072)	(0.047)	(0.047)	(0.056)	(0.056)	(0.049)	(0.050)	(0.040)	(0.043)
Observations	3,753	3,753	3,753	3,753	3,753	3,753	3,753	3,753	2,667	2,667
R-squared	0.098	0.102	0.052	0.054	0.092	0.097	0.031	0.036	0.102	0.107
Controls[(a)]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls*Treated	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Mean Dep. Var. (Control)	0.0400	0.0400	0.0439	0.0439	-0.139	-0.139	0.00186	0.00186	-0.0631	-0.0631
Std Dev Dep. Var. (Control)	6.106	6.106	3.474	3.474	4.145	4.145	3.452	3.452	3.625	3.625
P-value test joint sig	1.70e-05	0.0405	0.0218	0.305	0.0838	0.554	0.928	0.297	0.741	0.149

Notes: Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. P-values in square brackets corrected for multiple hypothesis testing using FWER. Increasing values of the dependent variable correspond to outcomes more consistent with the messages of Shuga.

(a) Controls in each regression include: female, age, currently attending school, years of education, english spoken, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table A20: Impact of T2 on outcome indexes

	(1) *HIV knowledge*	(2) *HIV attitudes*	(3) *HIV testing*	(4) *Attitudes towards risky sexual behavior*	(5) *Risky sexual behavior (for sexually active)*
Panel A: Cross-section					
T2	-0.089	-0.065	0.019	-0.075	0.057
	(0.209)	(0.137)	(0.174)	(0.122)	(0.132)
Observations	3,402	3,402	3,402	3,402	2,487
R-squared	0.091	0.057	0.093	0.029	0.084
Panel B: Conditional specification					
T2	-0.074	0.037	0.101	-0.049	-0.053
	(0.195)	(0.128)	(0.156)	(0.121)	(0.148)
T2*Y_{t-1}	0.036	0.057	-0.028	-0.027	0.005
	(0.044)	(0.044)	(0.037)	(0.040)	(0.067)
Y_{t-1}	0.301***	0.334***	0.447***	0.304***	0.209***
	(0.032)	(0.031)	(0.027)	(0.029)	(0.050)
Observations	3.402	3.402	3.402	3.402	2.117
R-squared	0.188	0.184	0.246	0.107	0.135
Controls[(a)]	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var. (Control)	1.080	0.534	0.278	0.176	0.0740
Std Dev Dep. Var. (Control)	6.043	3.373	4.609	3.433	3.562

Notes: Standard errors in parentheses adjusted for clustering at the screening centre level (54 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. Increasing values of the dependent variable correspond to outcomes more consistent with the messages of Shuga.

(a) Controls in each regression include: female, age, currently attending school, years of education, english spoken at home, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table A21: Spillovers, cross sectional specification

Dep. Var. (Y_t):	*HIV knowledge*		*HIV attitudes*		*HIV testing*		*Attitudes towards risky sexual behavior*		*Risky sexual behavior (for sexually active)*	
	(1)	(2)	(3)	(4)	('5)	(6)	(7)	(8)	(9)	(10)
Friend of Treated	0.642*	0.587	-0.063	-0.080	-0.454	-0.520	-0.300	-0.253	-0.408	-0.366
	(0.376)	(0.375)	(0.256)	(0.258)	(0.296)	(0.313)	(0.210)	(0.220)	(0.319)	(0.324)
	[0.216]	[0.263]	[0.772]	[0.749]	[0.195]	[0.263]	[0.290]	[0.439]	[0.290]	[0.439]
Friend of Treated*Opposite Sex		1.388		0.636		1.719		-1.214		-1.181
		(1.221)		(1.028)		(1.266)		(1.069)		(1.448)
		[0.541]		[0.559]		[0.541]		[0.532]		[0.532]
Opposite Sex		-0.799		0.187		-0.812		0.630		0.724
		(0.738)		(0.732)		(0.984)		(0.800)		(0.756)
Observations	924	924	924	924	924	924	924	924	680	680
R-squared	0.119	0.119	0.093	0.094	0.139	0.140	0.057	0.058	0.125	0.126
Controls [(a)]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var. (Control)	-0.185	-0.185	-0.316	-0.316	-0.164	-0.164	-0.107	-0.107	-0.213	-0.213
Std Dev Dep. Var. (Control)	6.063	6.063	3.601	3.601	4.410	4.410	3.512	3.512	3.862	3.862

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. Increasing values of the dependent variable correspond to outcomes more consistent with the messages of Shuga. Sample includes network friends that have non-missing Y_{t-1}. FWER adjusted p-values in square brackets, based on 10,000 replications.

(a) Controls in each regression include: female, age, currently attending school, years of education, english spoken, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table A22: Balance between main respondents and network members used for spillover analysis

	Mean main respondents	*Mean network members*	*Diff=0 (p-value)*	*Normalized Diff.*[(a)]	*No. Obs.*
	(1)	(2)	(3)	(4)	(5)
Indexes					
HIV knowledge	0.141	0.300	0.456	0.018	6042
HIV attitudes	0.007	-0.158	0.167	-0.033	6042
HIV testing	-0.029	0.189	0.141	0.035	6042
Attitudes towards risky sexual behavior	-0.022	-0.367	0.003	-0.071	6042
Risky sexual behavior (sex act)	0.023	0.107	0.573	0.017	3808

Notes: The table reports sample means at baseline.

(a) Normalized difference is the difference in the sample means of main respondents and their network members divided by the square root of the sum of the sample variances.

Table A23: Indexes of conformism, loading factors

	Loading factor
Conformity index	
He/she believes that people should do what they're told. He/she thinks people should follow rules at all times, even when no-one is watching.	0.4832
It is important to him/her to always behave properly. He/she wants to avoid doing anything people would say is wrong.	0.5541
He/she believes he/she should always show respect to his/her parents and to older people. It is important to him/her to be obedient.	0.4393
It is important to him/her to be polite to other people all the time. He/she tries never to disturb or irritate others.	0.5162
Tradition index	
He/she thinks it's important not to ask for more than what you have. He/she believes that people should be satisfied with what they have.	0.5239
Religious belief is important to him/her. He/she tries hard to do what his/her religion requires.	0.5127
He/she thinks it is best to do things in traditional ways. It is important to him/her to keep up the customs he has learned.	0.4236
It is important to him/her to be humble and modest. He/she tries not to draw attention to himself/herself.	0.5321
Self-Direction index	
Thinking up new ideas and being creative is important to him/her. He/she likes to do things in his/her own original way.	0.4765
It is important to him/her to make his own decisions about what he/she does. He/she likes to be free to plan and to choose what to do himself/herself.	0.5445
He/she thinks it's important to be interested in things. He/she likes to be curious and to try to understand all sorts of things.	0.4678
It is important to him/her to be independent. He likes to rely on himself.	0.5076

Table A24: Treatment and conformism, cross section

Dependent Variable:	(1) *HIV knowledge*	(2) *HIV attitudes*	(3) *HIV testing*	(4) *Attitudes towards risky sexual behavior*	(5) *Risky sexual behavior (for sexually active)*
Panel A: Conformity					
Treated	0.890***	0.346**	0.355**	0.147	0.126
	(0.244)	(0.138)	(0.148)	(0.090)	(0.131)
	[0.001]	[0.025]	[0.025]	[0.208]	[0.341]
Treated*Conformity	0.186	-0.147	-0.017	-0.057	-0.069
	(0.149)	(0.110)	(0.099)	(0.081)	(0.087)
	[0.216]	[0.463]	[0.876]	[0.669]	[0.669]
Conformity	-0.001	0.188**	-0.045	0.177***	0.121*
	(0.128)	(0.089)	(0.070)	(0.061)	(0.071)
R-squared	0.086	0.052	0.092	0.024	0.091
P-Value of joint signif.	0.000	0.014	0.018	0.113	0.348
Panel B: Tradition					
Treated	0.900***	0.344**	0.354**	0.144	0.118
	(0.244)	(0.136)	(0.147)	(0.090)	(0.131)
	[0.000]	[0.024]	[0.024]	[0.216]	[0.370]
Treated*Tradition	-0.031	-0.127*	0.017	0.007	-0.139
	(0.151)	(0.068)	(0.088)	(0.090)	(0.106)
	[0.975]	[0.174]	[0.982]	[0.926]	[0.321]
Tradition Index	-0.012	0.145***	-0.055	0.137*	0.215**
	(0.114)	(0.054)	(0.060)	(0.078)	(0.094)
R-squared	0.085	0.051	0.092	0.024	0.092
P-Value of joint signif.	0.000	0.014	0.018	0.114	0.386
Panel C: Self Direction					
Treated	0.865***	0.334**	0.349**	0.144	0.125
	(0.239)	(0.136)	(0.149)	(0.093)	(0.127)
	[0.000]	[0.028]	[0.028]	[0.225]	[0.326]
Treated*Self-direction	0.134	0.005	0.125*	0.038	0.137*
	(0.121)	(0.071)	(0.068)	(0.081)	(0.078)
	[0.288]	[0.937]	[0.206]	[0.640]	[0.164]
Self-direction	-0.538***	-0.149***	-0.093**	-0.077	-0.154**
	(0.100)	(0.052)	(0.046)	(0.068)	(0.059)
R-squared	0.095	0.054	0.092	0.021	0.091
P-Value of joint signif.	0.000	0.016	0.021	0.123	0.324
Controls [(a)]	Yes	Yes	Yes	Yes	Yes
Observations	4,986	4,986	4,986	4,986	3,618
Mean Dep. Var. (Control)	0.04	0.044	-0.139	0.00186	-0.063
Std Dev Dep. Var. (Control)	0.000	0.0133	0.0186	0.113	0.383

Notes. Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. Increasing values of the dependent variable correspond to outcomes more consistent with the messages of Shuga. FWER adjusted p-values in square brackets, based on 10,000 replications.

(a) Controls in each regression include: female, age, currently attending school, years of education, english spoken at home, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.

Table A25: Heterogeneity by education and English

Dep. Var. (Y^t):	*HIV knowledge*		*HIV attitudes*		*HIV testing*		*Attitudes towards risky sexual behavior*		*Risky sexual behavior (for sexually active)*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated	-3.648*	-0.084	-1.808	-0.239	0.779	-0.762	-1.780	-0.315	-1.748	0.377
	(2.173)	(0.890)	(1.147)	(0.466)	(1.129)	(0.688)	(1.088)	(0.568)	(1.059)	(0.527)
	[0.285]	[0.924]	[0.285]	[0.852]	[0.492]	[0.622]	[0.225]	[0.733]	[0.225]	[0.733]
Treated*Years of Education	0.382**		0.185*		-0.038		0.166*		0.163*	
	(0.185)		(0.096)		(0.098)		(0.094)		(0.090)	
	[0.135]		[0.135]		[0.697]		[0.173]		[0.173]	
Years of Education	0.007		0.050		0.062		-0.019		-0.100	
	(0.161)		(0.087)		(0.055)		(0.084)		(0.068)	
Treated*English spoken at home		0.903		0.604		1.144*		0.482		-0.242
		(0.926)		(0.479)		(0.686)		(0.591)		(0.529)
		[0.387]		[0.387]		[0.284]		[0.674]		[0.674]
English spoken at home		0.580		-0.528		-0.682		-0.376		-0.103
		(0.749)		(0.394)		(0.620)		(0.501)		(0.454)
Treated*Y lagged	-0.077**	-0.068*	-0.002	0.007	-0.032	-0.035	-0.005	-0.001	-0.118**	-0.117**
	(0.036)	(0.036)	(0.038)	(0.037)	(0.032)	(0.032)	(0.034)	(0.033)	(0.056)	(0.056)
	[0.103]	[0.177]	[0.934]	[0.861]	[0.536]	[0.483]	[0.887]	[0.985]	[0.078]	[0.087]
Y lagged	0.399***	0.393***	0.367***	0.361***	0.471***	0.473***	0.296***	0.293***	0.322***	0.321***
	(0.028)	(0.029)	(0.031)	(0.030)	(0.027)	(0.027)	(0.028)	(0.027)	(0.046)	(0.046)
Observations	4.986	4,986	4.986	4,986	4.986	4,986	4.986	4,986	3.07	3,070
R-squared	0.199	0.198	0.180	0.179	0.261	0.262	0.099	0.099	0.153	0.153
Controls[(a)]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var .(Control)	0.0400	0.585	0.0439	0.297	-0.139	0.0925	0.0019	0.0849	-0.0631	0.0156
Std Dev Dep. Var. (Control)	6.106	6.127	3.474	3.476	4.145	4.441	3.452	3.417	3.625	3.546
P-value test joint sig	0.0972	0.925	0.119	0.609	0.492	0.272	0.106	0.581	0.103	0.476

Notes: Standard errors in parentheses adjusted for clustering at the screening centre level (80 clusters). ***, ** and * denote significance at the 1, 5 and 10 percent level, respectively. Increasing values of the dependent variable correspond to outcomes more consistent with the messages of Shuga. FWER adjusted p-values in square brackets, based on 10,000 replications

(a) Controls in each regression include: female, age, currently attending school, years of education, English spoken at home, single, does not live with the family, household size, wealth index, home owner, father obtained higher than secondary education, mother obtained higher than secondary education, muslim, speaks Yoruba as a native language.