Narrative Entertainment Shapes Policy Priorities: Evidence from Four Field Experiments in Tanzania

Salma Emmanuel, Donald P. Green, Dylan W. Groves,

Constantine Manda, Beatrice Montano, Bardia Rahmani *

October 8, 2024

Abstract

A growing body of work finds that entertainment-education interventions can influence attitudes, beliefs, and behaviors, but few studies consider their effects on audiences' political preferences. We present results from a series of experiments conducted in Tanzania that estimate the impact of four radio dramas on how audiences prioritize protecting the environment, countering gender-based violence, reducing early forced marriage, and improving access to HIV treatment. Interviewing listeners 2-4 weeks after they were exposed to the drama, we find that three of the four dramas significantly increased listeners' preference for hypothetical candidates promising to address the issue featured in the drama, and all four dramas elevated the perception that the issue represents a top priority for the community. Pooling across studies (N = 4,504), the effects of narrative messages on voting and prioritization persist more than 16 months after the audio screening and spill over to the spouses of audio screening attendees.

*The authors would like to thank the implementing partner, UZIKWASA, led by Novatus Urassa, with special thanks to Eric Mallya for his tireless efforts. We are grateful to the entire research team at Innovations for Poverty Action, particularly Martin Zuakulu, Rachel Jones, Kyu Khin Gar, Zachary Isdahl, Robert Mwandumbya, Neema Msechu, Gilbert Loshoock, Cosmas Sway, Zinabu Omary, Fatuma Yahaya, Frank Simon, Fadhili Mashaka, Stephano Kizigo, and Saulo Mwakisongo. We thank discussants and participants at the 2023 Meeting of the American Political Science Association. An earlier version of this paper benefitted from comments at the 2023 Meeting of the Association for Psychological Science and the 2023 Meeting of the Midwest Political Science Association. We thank Salma Mousa and Elizabeth Levy Paluck for their very helpful comments on earlier drafts. The studies described in this paper were reviewed and approved by Columbia Institutional Review Board (protocol AAAR5582) and Tanzania's Commission of Science and Technology (protocol 14528). These studies were made possible by financial support from Wellspring Philanthropic Fund. Corresponding author: dpg2110@columbia.edu

1 Introduction

Recent years have seen a surge of interest in narrative entertainment, such as soap operas, feature-length films, or short-format vignettes on radio or television. Governments and NGOs view narrative messages as a promising vehicle for changing attitudes and behaviors at scale (Frank and Falzone 2021), a view that draws increasing support from randomized trials on multiple continents. For example, Paluck and Green (2009) find substantial effects of a year-long radio drama on political norms in post-genocide Rwanda; Murrar and Brauer (2018) show that an educational television sitcom featuring Arab Muslim characters can reduce Canadians' prejudice towards Arabs and Muslims; Green, Wilke and Cooper (2020) find that video vignettes decrease violence against women in rural Uganda; and a host of studies find that edutainment can reduce HIV-related risk and stigma (Banerjee, Ferrara and Orozco-Olvera 2019; Creel et al. 2011; Lapinski and Nwulu 2008). A recent review of this literature by Rahmani et al. (2023) shows that not only is the pace of experimental research accelerating, but studies are increasingly situated in low- and middle-income countries, where weaker state capacity makes the prospect of persuasive messaging via mass media especially important.

Despite this extensive body of research, we know surprisingly little about whether narrative entertainment shapes an audience's political preferences. Although the Rahmani et al. (2023) meta-analysis finds narrative entertainment to be broadly influential, it also calls attention to an important gap in the literature: the 77 experiments they review focus almost exclusively on outcomes such as *acquiring knowledge* (e.g., learning how HIV is transmitted), *changing attitudes* about social groups (e.g., empathizing with ethnic minorities), *developing behavioral intentions* (e.g., willingness to report incidents of intimate partner violence to local authorities), and *altering behaviors* (e.g., bringing young children to health centers for immunizations). To date, few studies have assessed the hypothesis that exposure to narrative entertainment changes the importance that audiences place on the political and policy issues at the center of the story, whether through

collective action or some form of government intervention.[1] As we explain below, this hypothesis is suggested both by social psychological theories about narrative persuasion (Bandura 1976; Slater and Rouner 2002) as well as by "agenda setting" theories in political science that contend that news media shape what the public considers to be the most pressing issues of the day (Iyengar and Kinder 2010).

This paper attempts to fill this gap by pulling together evidence from recent experiments that present narrative dramas about social issues *and* measure the importance that audiences accord the issues depicted.[2] Do audiences become more likely to call the depicted problem a priority for their locality vis-a-vis other issues? Do audiences become more likely to express support for hypothetical candidates for office whose platform emphasizes the issue that was depicted in the narrative?

Our evidence comes from recent experimental evaluations of four dramas on the topics of protecting the environment, countering gender-based violence, reducing early forced marriage, and improving access to HIV treatment. As explained below, these studies employ similar lab-in-the-field designs and were conducted in the same region, northeastern Tanzania. Pooling the individual-level data from these studies enables us to inductively derive a series of scope conditions about the extent to which narrative entertainment produces large average effects on audiences' priorities. Interestingly, all four dramas produce substantial and statistically significant

[1]One exception is the placebo controlled experiment in Uganda reported in Green, Wilke and Cooper (2020) and Wilke, Green and Tan (2022). The latter reports that video vignettes about teacher absenteeism significantly increased the importance that participants accorded that issue eight months later but had no detectable effect on support for candidates promising to address this problem.

[2]Results from the midline but not the endline evaluation of the HIV drama were reported in (Green, Groves and Manda 2021). Green et al. (2022) presented the effects of the early forced marriage drama, but endline priorities are described only in the supplementary materials. The results from Experiment 2 have not been published.

effects on beliefs and attitudes measured weeks after exposure to the narratives; in that sense, all of the dramas "work" by the usual standards of this literature.[3]

Do these persuasive effects translate into changes in audiences' policy priorities? All four of the dramas made audience members significantly more likely to indicate weeks or months later that the issue depicted in the story is an important social or political concern for their community. Moreover, three of the four dramas produce substantial and statistically significant effects on vote preferences. The lone exception was the narrative that depicted public officials as being unresponsive to early forced marriage concerns when they were raised. In two of three instances in which the dramas produced large average treatment effects, they depicted people solving community problems through collective action, and in the third instance, the plot turns on the main character accessing government services. In many cases these effects remain detectable, albeit in attenuated form, over a year later despite the lack of continued messaging. Taken together, the results indicate that mass entertainment can lead to substantial and persistent changes in what issues people take to be important. Entertainment-education, like news media and political campaigns, may thus play a role in shaping citizens' policy priorities.

The rest of the paper is structured as follows. The next section reviews theories about why narrative entertainment may be thought to affect audience members' priorities. Special attention is paid to theories that imply heterogeneous treatment effects. We then describe the research designs (subjects, narratives, and outcome measurement) of the studies we analyze. After noting that all four narrative dramas influence attitudes directly connected to the main plot line, we turn specifically to the dramas' effects on listeners' priorities. The pooled results reveal sizable

[3]For example, the HIV drama significantly increased knowledge about how the disease is transmitted and willingness to disclose one's HIV+ status; the EFM drama reduced endorsement of the view that fathers should be the ones to pick their daughters' husbands; the environmental conservation drama increased awareness of climate change and support for conservation; and the gender-based violence drama increased awareness of safety concerns confronting women and girls as well as willingness to testify against offenders.

average effects that are statistically significant by even the most demanding standards. We then review evidence for long-term effects, spillovers to spouses, friends, and children, and treatment effect heterogeneity. We conclude by discussing how the hypotheses that emerge from these experiments might be tested directly in future experiments.

2 Theoretical Basis for Effects on Policy Priorities

Two prominent theories of narrative media help explain why narrative dramas might influence political preferences. One such theory is Social Cognitive Theory (SCT, Bandura 2004), which holds that people change their views to align with those displayed by appealing models. According to SCT, narratives persuade by presenting viewers with attractive characters who model appropriate behaviors. A corollary hypothesis, which may also imply heterogeneous treatment effects, is that viewers are especially likely to emulate the behavior of characters with whom they identify (Moyer-Gusé 2008).

The connection between SCT and policy priorities may occur either because the characters themselves model the pursuit of a policy objective or because the story line dramatizes the prevalence of a social problem. An example of the former is a trio of short films studied by Wilke, Green and Tan (2022) that depicted Ugandan villagers taking action to address chronic teacher absenteeism in their children's schools.

The second leading social psychological theory is the Extended Elaboration Likelihood Model (E-ELM, Slater and Rouner 2002), which highlights the persuasive opportunities that are unique to absorbing narratives. Ordinarily, people tend to counter-argue against messages they perceive to be persuasive, especially when such messages clash with their prior attitudes. Narrative messages circumvent counter-arguing and other forms of resistance by drawing audiences into an absorbing story and promoting identification with the characters (Green and Brock 2000; Slater and Rouner 2002). This theoretical perspective, which is widely used to explain narrative-driven attitude change (Moyer-Gusé 2008), may also be applied to the specific case of policy priorities. Those who might ordinarily resist the notion that a given issue is important and demands a policy response might soften their view in the wake of a transporting and memorable narrative.

The cross-disciplinary literature on agenda setting in political science and communication studies offers further reasons to expect narrative messages to influence audiences' political priorities. The leading scholarly works on agenda setting focus on two themes: how the framing of news stories affects the public's evaluation of issues and candidates (McCombs and Valenzuela 2020) and how the volume of news coverage that is devoted to an issue elevates the public's concern about it (Iyengar and Kinder 2010; Iyengar, Peters and Kinder 1982). News and narrative entertainment are obviously quite different, but both may achieve a high degree of source credibility with audiences under certain conditions. News stories that deploy the same message over days or weeks may reap persuasive benefits from repetition (Fazio et al. 2015) to audiences that might otherwise be inattentive, but narratives may also have the advantage of attracting interested audiences and conveying lengthy messages, sometimes spanning many episodes of a serialized drama. Moreover, as an agenda-setting device, narrative entertainment may have the advantage of conveying messages in a form that is more accessible and memorable than a typical news story (McCombs and Shaw 1972), perhaps leading audiences to impart the messages to others in their social network (Sinclair 2012).

Taken together, these theoretical perspectives offer multiple reasons to suspect that narrative entertainment may shape what political issues audiences believe to be important. However, three aspects of the empirical literature merit further exploration.

First, experimental evidence about the size and persistence of narrative drama's effects on political priorities is surprisingly sparse. Rahmani et al. (2023)'s meta-analysis identifies only 3 of 77 experimental studies of narrative communication that measure voting intention or prioritization outcomes. Of these, only one measures the effects more than a year after exposure. A vast literature casts doubt on the persistence of media messages' influence (Gerber et al. 2011); it remains to be seen whether the effects of narrative messaging on political prioritization and voting intentions are similarly short lived.

Second, no studies test for whether effects of narrative dramas on political priorities spread beyond the drama's immediate audience. Classic theories of traditional media effects hold that

individuals can be influenced when they directly engage with media messages or through discussion with opinion leaders or members of their social network (Lazarsfeld, Berelson and Gaudet 1944; Wanta and Wu 1992; Weaver, Zhu and Willnat 1992). However, the degree to which the agenda setting effects of narrative dramas are magnified by spread through social networks remains untested.

Finally, few studies of narrative dramas and political priorities test for effect heterogeneity by characteristics of the drama or audience (Banerjee, Ferrara and Orozco-Olvera 2019). Several questions about the optimal design of dramas remain unanswered. Is it essential for dramas to explicitly model political activity, as Social Cognitive Theory suggests? Should dramas feature protagonists from contexts with which audiences can easily identify, as recommended by certain versions of the Extended Elaboration Likelihood Model? Are narrative dramas most effective when they reinforce existing political views, or when they challenge conventional wisdom? Few research designs are situated to compare narrative dramas that are similar in most aspects but differ along these key dimensions.

In a related vein, current scholarship offers little guidance about the segments of rural populations that are most susceptible to the influence of narrative dramas. Are narrative dramas most persuasive among those who regularly use and trust media (Iyengar and Kinder 1987; Tsfati 2003), or among those who otherwise lack access to media messages? Among younger and more sociall progressive audience members, or among all audience members more or less equally? The optimal design and targeting of narrative dramas for influencing political priorities is of clear importance to governments and civil society, but empirical scholarship has yet to yield reliable, field-tested answers.

3 Study Context

The setting for the experiments is Tanzania's northeastern Tanga region. Tanga offers an appealing context for the study of narrative dramas on local political priorities in Sub-Saharan Africa. First, Tanga's rural population is broadly representative of remote rural populations that are frequent targets of radio campaigns in Sub-Saharan Africa. Roughly 75% of the respondents

in the sample have Standard 7 education or higher, 60% regularly listen to the radio, and 52% visit the city more than once a year (Table A1). The study sample has a higher proportion of Muslims (75%) and a lower proportion of cell-phone owning households (20%) than rural communities in other parts of Sub-Saharan Africa. However, as we demonstrate below, we observe no evidence that the effect of narrative dramas differs across these respondent attributes.

Tanzania's robust village-level governance institutions also make it a useful context for studying the drivers of political prioritization and vote choice in local politics. Tanzania has pursued a "decentralization by devolution" strategy since 1999, known as the Local Government Reform Program (LGRP). One feature of the decentralization program has been the empowerment of village/neighborhood chairpersons (mwenyekiti). Mwenykiti are tasked with coordinating village responses to local social and development problems and communicating villager needs to district officials. The villager political priorities that are the focus of this paper therefore influence long-term development outcomes in two ways: they directly influence village development mans and the selection of village and ward leaders, and they provide inputs into district-level policy planning. Village government elections are often closely contested and tend to focus on the local social and political problems that we focus on in this paper: improving service delivery, protecting local natural resources, and responding to perceived threats to social well-being.[4]

Finally, radio is the dominant mass medium in rural Tanga and narrative dramas are a familiar feature of local radio programming (Green et al. 2024). Radio stations in Tanzania develop episodic dramas in coordination with the government and non-governmental organizations to promote messages on topics ranging from public health to religious tolerance to voter turnout (Green et al. 2024). The format of the radio dramas screened for this study were therefore easily recognizable to villagers.

4 Research Design

We analyze data from two randomized experiments, each of which assigns exposure to one of

[4]In ruling party strongholds, primary elections still provide villagers with an opportunity to express preferences for different policy platforms.

two radio dramas. The first experiment presented certain audiences with an audio drama about overcoming HIV stigma, while others listened to a drama on the subject of early forced marriage. The second experiment presented certain audiences with an audio drama about violence against women, while others listened to an audio drama of similar length on the subject of environmental conservation. The maintained assumption in both experiments is that the the content of one drama has no effect on survey responses that subjects offer on the subject of the other drama or unrelated political priorities. We discuss evidence for this assumption in detail in Appendix D. What follows is a brief description of each experiment.

4.1 Experiment 1: HIV Stigma and Early Forced Marriage

4.1.1 Setting

This study was conducted in 30 villages in northeastern Tanzania (REDACTED). Two villages were selected from each of 15 wards then randomly assigned to either the HIV stigma treatment or early forced marriage treatment. The experiment is thus a pair-matched, placebo-controlled cluster randomized trial. Villages were eligible for selection only if they were located more than 8km from a major town, more than 4km away from any other major village, and outside the coverage area of Pangani FM, the radio station that broadcast the EFM drama that is the subject of this study.

4.1.2 Participants

In each village, approximately 40 randomly selected respondents were surveyed face-to-face at baseline and invited to attend a community screening of a radio drama a few days later. The content of the radio drama was not specified, and the interviewers at baseline were blinded to which radio program was to be presented. A screening was held in each village in the early evening to accommodate respondents' work obligations. The screenings were held in a classroom or other indoor community meeting place, and attendees were provided light snacks and refreshments. At all sites, two members of the field team briefly discussed the logistics of the screening and provided refreshments midway through the event but did not explain or discuss the content of the shows. Compliance with the assigned treatment was very high. Of the 1,205

respondents who completed a baseline survey and were invited to attend a screening with others surveyed from their village three or four days later, 1,035 (86%) attended.

4.1.3 Radio dramas

Wahapahapa (HIV). The radio drama *Wahapahapa* was a serialized Swahili radio drama developed by a Tanzanian media organization to increase awareness of HIV/AIDS treatments, reduce stigma towards HIV positive individuals, and increase listeners' willingness to disclose their HIV status. In order to expose participants to the main theme of the drama in the context of a single sitting, the original series was condensed to a 110-minute version.

The plot of *Wahapahapa* concerns Ray, a young musician living in an unnamed town who comes to terms with his own HIV positive status. The narrative also follows Ray as he discloses his status to his female romantic partner, employer, and friends, and receives their support in seeking out anti-retroviral drugs and medical support from clinics and physicians. The protagonist models the process of seeking out appropriate medical care, as opposed to traditional or spiritual remedies, and the audience sees the positive impact of anti-retroviral drugs on his well-being.

Tamapendo (EFM). While 15 villages were exposed to *Wahapahapa*, the other 15 villages received a 1 hour and 45-minute abridged audio screening of *Tamapendo*, a serialized Swahili radio program developed by the grassroots Tanzanian NGO UZIKWASA. *Tamapendo* focuses on reducing support for early forced marriage (EFM), and the abridged version makes no mention of HIV or access to medical care.

The abridged version of *Tamapendo* follows the story of a young girl, Fatuma, as she responds to the prospect of forced marriage. Fatuma's age is not stated explicitly, but she is understood to be between 14 and 17 years old. Fatuma is portrayed as an intelligent girl who wants to continue her education when she graduates from secondary school. However, her father arranges to marry her to a wealthy older man from outside the village without her consent. Fatuma's mother, best friend, and romantic interest each begin the story as passive bystanders to the forced marriage but grow to understand Fatuma's resistance to the arrangement and ultimately support Fatuma when she rejects the forced marriage in the final, climactic scene. As part of their character devel-

opment, conversations between Fatuma and her romantic interest model equitable gender roles in romantic relationships, such as mutual support for female education, equal household work, and an opposition to intimate partner violence. Their discussions contrast with the relationship between Fatuma's mother and father, which is depicted as hierarchical and abusive. Notably, the *Tamapendo* screening did not include any discussion about legal rights or state sanctions. In the only scene in which a character reaches out to village officials to ask for help escaping an arranged early marriage, the officials are depicted as skeptical and unresponsive.

4.1.4 Outcome Measures

The midline survey measured outcomes 13-16 days after the village screenings. 95.6% of baseline respondents (regardless of whether they attended the screening) completed this survey. The endline survey measured outcomes 16 months after the village screenings: 96.5% of baseline respondents completed this survey. As part of the endline survey, enumerators also interviewed the spouse, teenage children, and one close friend of the original respondent. The survey of teenage children and friends took place two months after the endline survey, approximately 18 months after the audio screening.

Community Priorities. The first measure of priorities gauges the importance accorded to HIV compared to other "goals for your village." Respondents were asked, "Here is a set of cards, which show different goals for your village. Please choose the three that are currently the most important to you, and the item that is least important" and then were handed a set of six cards, each with the name and image associated with a given goal: reduce the incidence of forced marriage, reduce the amount of crime, increase the availability of medicine for HIV/AIDS, increase the number of roads, increase the availability of electricity, increase the availability of water, reduce the number of people who do not have enough food to eat. The community priorities measure ranges from 0 (HIV unranked) to 3 (HIV ranked first). To avoid contamination by the *Tamapendo* treatment, the priority ranking associated with forced marriage was excluded. For example, if a respondent ranked forced marriage as their top priority and HIV as their second priority, HIV

was coded as the respondent's top priority.[5] This recoding scheme results in a scale that ranges from 0 (HIV rated least important or unranked) to 1 (HIV ranked first or ranked second if EFM is ranked first). Among those who attended the screening of the EFM drama, the average rank of the HIV priority is 0.18: 10% ranked it first (or second after EFM), 7% ranked it second, 10% third, and 73% unranked.

A similar coding procedure was used to assess the effects of *Tamapendo*. The forced marriage priority item excluded the HIV priority ranking; the resulting scale ranges from 0 (EFM rated least important or unranked) to 1 (EFM ranked first or ranked second if HIV is ranked first). EFM was ranked by those who attended the HIV drama on average 0.26: 15% ranked it first (or second after HIV), 13% second, 8% third, and 64% unranked.

In the endline survey, respondents were provided with 9 potential goals for the their village and asked to rank them in order of importance: increasing agriculture and fishing equipment; reducing crime in the village; reducing early forced marriage; education and schools; justice and resolving of problems e.g. land problems; availability of electricity; sanitation and waste management; improvement of infrastructure and roads; availability of health services like increasing access to HIV/AIDS medication.

To assess the effect of *Wahapahapa* on participants' midline priority ranking of HIV/AIDS medication, the goal of "reducing early forced marriage" was excluded from the priority list to avoid contamination. The resulting scale ranges from 0 (lowest priority rank) to 1 (highest priority rank) by dividing the rank by 8. The average score among placebo group compliers was 0.6, with 13% ranking it first and no placebo group compliers ranking it last.

The same ranking procedure was used to measure the effect of *Tamapendo* on long-term prioritization of early forced marriage. Among those who attended the HIV drama, the average priority score for early forced marriage at endline was 0.4: 7% ranked it first and 14% ranked it last.

[5]Experiment 2 sidesteps this measurement problem by putting the priorities questions for each drama into two distinct lists.

Electoral Priorities. The second measure of policy importance uses a series of votes in a hypothetical local election. Although Tanzania is generally considered a hybrid democracy with a hegemonic ruling party, public campaigns for village chairperson, or *mwenyikiti*, are common (Shayo 2021; Weghorst 2022). In addition to advertising their personal qualities, candidates run on platforms centered on local development and social issues over which they exercise control. Even in villages where opposition party support is muted, ruling party candidates publicly campaign during open party primaries (Babeiya 2022).

The midline and endline surveys used a random rotation of candidate names signaling different genders and religions to pit candidates against one another. Each candidate also proposed a different platform. One candidate seeks to "increase the availability of medicine for HIV/AIDS," while the other candidate's platform is to improve the quality of roads in the village or to reduce the amount of crime. The outcome variable for *Wahapahapa* scored 1 if the candidate running on the HIV platform attracts the respondent's vote; 0 otherwise. In control sites, 39% of compliers voted for the candidate campaigning to increase availability of HIV/AIDS treatments, and the correlation between preferring candidates with an HIV-platform and prioritizing HIV/AIDS services as a village policy goal is 0.26.

An analogous measure was used to assess the effects of exposure to *Tamapendo* on vote preferences. Here, the relevant platform was to "fight against child marriage"; an outcome was scored 1 if the respondent selects the candidate running on the EFM platform.[6] In control sites, 56% of compliers selected the candidate promising to fight against child marriage. The correlation between expressing support for candidates running on an EFM-platform ranking reducing forced marriage as a village policy goal is 0.32. Respondents were presented with a nearly identical campaign vignette during the 16 month endline survey.

[6]Note that the cases where the HIV and EFM platform were pitted against each other were dropped from the analysis. Both the 16 month endline measure and second experiment sidestep this problem by stipulating in the survey code that candidates with early forced marriage and HIV/AIDS platforms do not run against one another.

4.2 Experiment 2: Environmental Protection and Violence Against Women

Experiment 2 replicates and extends the results of Experiment 1. It examines two new issue areas, samples from a larger number of communities, and makes several changes to the ways in which political priorities are measured.

4.2.1 Setting

Experiment 2 was conducted in 34 villages in northeastern Tanzania, outside the coverage areas of Pangani FM, which was broadcasting the radio soap opera about gender-based violence. As in Experiment 1, villages were selected to be far enough from major roads that residents were unlikely to have heard the broadcasts while traveling. Villages within 10km of one another were paired and then randomly assigned to either the Environment or Gender-Based Violence treatment.

4.2.2 Participants

As in Experiment 1, in each village, approximately 40 randomly selected respondents were interviewed in person at baseline and invited to attend a community screening of a radio drama a few days later. The content of this radio drama was not specified, and the interviewers were blinded to which radio program was to be presented. Compliance with the assigned treatment was again high: 90.7% of respondents were present at the start of the screening and 91.3% were present at the end.

4.2.3 Radio dramas

Mikoko Yetu (Environment). This 1-hour Swahili-language radio drama on the topic of environmental conservation was developed in collaboration with local producers. The drama follows Bakari, a ranger from a coastal fishing village in Tanga, as he seeks to persuade his community to oppose a corrupt bargain between the village leader and foreign developers that would destroy the mangroves upon which the fishermen's livelihood depends.

The main plotline centers on Bakari's efforts to run for office on a platform to protect the mangroves; meanwhile the incumbent village leader campaigns on his plan to sell the mangroves to make way for a large-scale, foreign-owned shrimp farm. Bakari receives assistance from villagers

receptive to his message; however, he encounters resistance from some villagers who believe the corrupt village leader's claims of an economic windfall from the deal. The story culminates in Bakari winning the election and the community turning its back on the former village leader after discovering that he took bribes and directed a campaign of violent intimidation. Along the way, discussions between the characters make reference to other environmental issues such as the causes and consequences of climate change and the threat of illegal logging.

Boda Bora (GBV). This Swahili-language radio drama was a serialized, multi-week program set in the Tanga Region of Tanzania and written by the local NGO UZIKWASA. The drama tells the story of a grassroots campaign to prevent and report instances of sexual violence against women and girls. The research team worked with the NGO to condense the radio drama by focusing on key themes, so that it could be presented to participants in a single sitting of about 1 hour, and to add message-relevant narration. The abridged version of *Boda Bora* follows a young motorcycle taxi driver, Juma, as he seeks to mobilize his peers to stop engaging in and facilitating sexual assault and child prostitution. Juma organizes a collective effort to deter potential perpetrators and report sexual violence to authorities.

4.2.4 Outcome Measures

Four weeks after attending the screening, respondents were recontacted and invited to participate in a follow-up survey that measured outcomes. Enumerators collected outcome measures again 16 months after the screenings. Attrition was minimal for both surveys: 96.4% of baseline respondents completed the midline survey, and 94% of baseline respondents completed the endline survey. At endline, enumerators also interviewed the female spouses and teenage children of a randomly selected subset of respondents. Due to budget constraints, enumerators in Experiment 2 did not interview male spouses or the friends of respondents.

As in Experiment 1, outcomes focus on priorities and vote preferences. However, following the measurement approach of Experiment 1's endline survey, both outcome measures were refined by breaking the questions into two sub-parts, one involving environmental protection but not gender-based violence and the second involving gender-based violence but not environmen-

tal protection.

Community Priorities. Respondents were asked two questions about their community priorities. In the first question, respondents were given three laminated cards, each with "different goals for your village": protecting the environment, improving health and clinics, and increasing agricultural equipment. Respondents were asked to rank the cards in order from their top priority for the village to their lowest priority. Responses were scored 1 if environmental protection was chosen as a first priority, 0.5 if chosen as a second priority, and 0 if chosen as a last priority. The average rank of the environmental priority among those who attended the GBV audio screening is 0.37, where 13% ranked it first, 47% second, and 39% ranked it last.

Later in the survey, respondents were given four new laminated cards listing different "social problems in villages in Tanzania": people hurting the environment, men abandoning their family, public officials taking bribes, and people not practicing religion. Respondents were scored 1 if they listed hurting the environment as their most important problem, 0.75 if they ranked it second, 0.25 if third, and 0 if it was their lowest priority. The average score for hurting the environment among compliers in the placebo group was 0.48, with 18% of respondents listing it as the top priority, 29% as their second priority, 30% as their third, and 22% as their last.

In order to assess the effects of exposure to *Boda Bora*, we asked respondents a second round of questions about the village goals and social problems they prioritize most. Among village goals, respondents selected between reducing sexual violence, improving access to water, and improving cell phone reception. On a scale of 0 (ranked last) to 1 (ranked first), the average rank of anti-GBV goal among the placebo group of compliers is 0.47: 22% ranked it first, 52% ranked it second, and 26% ranked it last. Among social problems, respondents ranked the following four: sexual violence against young girls, alcoholism, not paying back loans, and kids not going to school and people not working. The average priority score for sexual violence among placebo attendees was 0.48, with 18% ranking as their village's largest social problem, 29% ranking it

second, 29% third, and 24% last.[7]

The correlation between the two environmental priority measures among compliers in the control group for the *Miko Yetu* drama was 0.15; the correlation between the two gender based violence measures among compliers in the control group for the *Boda Bora* drama was was 0.09. In keeping with our pre-analysis plan, we created an index comprising the mean of the two priority scores.

At the 16 month endline survey, respondents were once against to asked to rank village goals and social problems. The index scores remained steady in the year between the midline and endline survey: the average priority rank among screening attendees in the control group for environmental priorities was 0.43 at midline and 0.43 at endline, while the priority score among screening attendees in the control group for gender based violence was 0.47 at midline and 0.48 at endline.

Electoral Priorities. Vote preferences were elicited in the same way as in the first experiment, by asking respondents to imagine a village about one day's walk away that is having an election for village chairperson, where there are two candidates giving speeches. The name of the candidates, signifying religion and gender, were randomly assigned. To assess the effects of *Mikoko Yetu*, the hypothetical election featured a candidate running on a platform to protect the environment with the slogan "Cut Down One Tree, Plant Many Trees," facing off against a candidate promising to either improve roads (with slogan "Make our roads better") or education ("Better schools for our children") for the community. The platform of the non-environment candidate was randomly assigned. Responses are scored 1 for voting for the environmental protection platform and 0 otherwise. The effects of *Boda Bora* are assessed according to respondents' support for a candidate (again, with a randomly attached name) who promises to fight against

[7]This 0.48 figure differs from the 0.47 reported in the table; the latter averages responses concerning both social problems and village goals.

sexual violence in the village with slogan "Protect our girls from sugar daddies[8] and rapists," whereas the opposing candidate promises to either improve roads or education with the same slogans as before.

At the 16 month endline, respondents were presented with the same candidate choices. Average support for candidates campaigning to protect the environment in the *Mikoko Yetu* control group was 38% at both midline and endline. In the *Boda Bora* experiment, 49% of control respondents selected the candidate promising to fight sexual violence at midline and 55% supported the candidate at endline.

4.3 Estimation

We focus our analyses on subjects who attended the audio screenings. We estimated the effect of the narrative dramas using ordinary least squares regression (OLS). If Y_i is the outcome for subject i, and T_i is a binary indicator of the subject's treatment status, the regression model

$$Y_i = \beta T_i + \gamma_1 block_{1i} + \gamma_2 block_{2i} \ldots + \gamma_k block_{ki} + u_i$$

expresses the outcome as a function of the narrative drama treatment, an indicator for each of the 17 blocks, and an unobserved error term u_i. We focus our analysis on the complier average causal effect (CACE), which is given by β. In this context, compliers are those who would attend an audioscreening if invited; by restricting our regression analysis to those who actually attended the audioscreenings, we isolate the average causal effect among compliers. We calculate exact p-values using randomization inference and cluster standard errors at the village level.

In keeping with the pre-analysis plan, we also use the LASSO procedure to select prognostic covariates from a set of demographic variables collected at baseline as well as a binary indicator for each ward block. Because we do not observe meaningful differences between estimates with and without adjustment for LASSO-selected covariates, we present results with LASSO-

[8]In the Tanzanian context, "shugadadi" (sugar daddy) is a common term that refers to a wealthy man who dangles expensive gifts to lure girls into having sex.

adjustment in the appendix.

5 Results

5.1 Community Priorities

Table 1 assembles the results from the four narrative dramas 2-4 weeks after the screenings. The results for *Wahapahapa* leave no doubt that the drama increased endorsement of the goal of "increasing the availability of medicine for HIV/AIDS" at midline. The average among compliers in the control group was 0.18; the treatment boosted this mean by approximately 0.12 (SE=0.02).

By comparison, *Tamapendo* had a somewhat more muted effect on respondents' political priorities: it elevated the ranking of "reduce the incidence of forced marriage" by almost 0.07 (SE=0.025) over a mean of 0.26 in the control group. This apparent effect is substantively meaningful, given that the village-level standard deviation is 0.10, reaching the conventional levels of statistical significance using a one-tailed test.

The statistical results for *Mikoko Yetu* are clearer. Although the average priority accorded to "protecting the environment" was 0.05 (SE=0.01) higher in the treatment group than in the control group (mean=0.43), the larger N in Experiment 2 means that the randomization inference p-value is less than 0.05.

The statistical evidence is also quite clear for *Boda Bora*, which increased the mean support for "reducing sexual violence" by almost 0.08 over a 0.47 base in the control group. With or without covariate adjustment, the p-values are less than 0.01.

Overall, although the point estimates vary somewhat across the four evaluations, the joint significance of the four estimates is unequivocal. Column (3) in Table 1 shows the results when data are pooled. The point estimate is 0.078 (SE=0.009), regardless of whether covariates are included in the regression model. This estimated effect is just over one-half of a village-level standard deviation.

5.2 Vote Preferences

Three of the four dramas produced large shifts in vote choice. Respondents who attended the

Wahapahapa screening were 18 percentage points more likely to vote for a candidate promising to increase access to HIV/AIDS care as opposed to a candidate who proposes to improve roads or crack down on stealing in the village ($p < 0.01$). To put this estimate in perspective, the average vote share received by the HIV-focused candidate in the control group was 39%. Respondents exposed to *Mikoko Yetu* became 6.8 percentage points more likely to vote for hypothetical candidates promising to protect the environment over candidates running on alternative platforms ($p < 0.05$), again in comparison to a base rate of 38% in the control group. And exposure to *Boda Bora* substantially increased respondents' propensity to vote for candidates whose platform highlighted the need to take action against sexual predators. A candidate running on this platform received 9.8 percentage points higher vote share than the control group, half of whom voted in support ($p < 0.01$).

Tamapendo had unexpectedly negative effects on vote support for a candidate who pledges to "fight against child marriage." One possible explanation is that the outcome question was posed to a relatively small number of respondents (N=660), and the resulting estimate is a sampling fluke. This interpretation is supported by the endline results (see below), which suggest a positive long-term effect. Another possibility is *Tamapendo* painted local officials as unresponsive to pleas for help when the issue arose in the village, perhaps sending the signal to beware of politicians' hypocrisy on this issue.

Table 1: Effects of Narrative Dramas on Voting Preferences and Political Priorities
2-4 weeks after exposure

	Pooled			Experiment 1						Experiment 2					
				HIV			EFM			Enviro			GBV		
	Index	Vote	Rank	Index	Vote	Rank	Index	Vote	Rank	Index	Vote	Rank	Index	Vote	Rank
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Treatment	0.078***	0.074***	0.079***	0.140***	0.175***	0.120***	0.025	-0.057	0.066**	0.049**	0.068**	0.049***	0.090***	0.098***	0.079***
Standard Error	0.009	0.015	0.008	0.022	0.031	0.020	0.023	0.030	0.024	0.015	0.033	0.011	0.013	0.025	0.009
RI p-value	<0.001	<0.001	<0.001	<0.001	0.001	0.001	0.217	0.929	0.049	0.016	0.040	0.002	<0.001	0.004	<0.001
Hypothesis	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Control Mean	0.38	0.45	0.35	0.25	0.39	0.18	0.36	0.56	0.26	0.41	0.38	0.43	0.48	0.49	0.47
Control SD	0.12	0.14	0.14	0.09	0.15	0.08	0.11	0.15	0.10	0.06	0.09	0.05	0.06	0.08	0.06
DV Range	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]
Blocked FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Covariates			72	16	14	17	26	8	22	19	0	14	12	6	16
Adj-R^2	0.10	0.04	0.16	0.09	0.09	0.06	0.08	0.06	0.08	0.05	0.00	0.08	0.05	0.03	0.05
Observations	4,504	3,822	4,462	1,001	660	1,001	1,001	660	1,001	1,251	1,251	1,230	1,251	1,251	1,230

Note: * p < 0.1, ** p < 0.05, and *** p < 0.01. Standard errors clustered at the village level. RI p-value come from one-tailed tests in the hypothesized direction. "Vote" columns report results for the *Vote Preferences* outcome as described in the text. In summary, they report results for responses to the question: *"Imagine a village about one day's walk from here is having an election for village chairperson. There are two candidates giving speeches. Let me tell you about each one and you can tell me which of the two you think should be elected. The first candidate is named [randomize: Mr. Salim, Mr. John, Mrs. Mwanahidi, Mrs. Nema] and [he / she] promises to [Target topic]. Their slogan is [Target topic slogan]. The second candidate is named [randomize: Mr. Salim, Mr. John, Mrs. Mwanahidi, Mrs. Nema] and [he/she] promises to [randomize: 2 other issues]. Their slogan is [2 other slogans]. Which of these two candidates do you think should be elected?"* The responses are scored as 1 for voting for the targeted topic platform, 0 otherwise. "Rank" columns report results for the *Community Priorities* outcome as described in the text. In summary, the report results for responses to the question: *"Here is a set of cards, which show different goals for your village. Now, please rank them in order from the most important to the least."*. The responses are scored between 1 when the Target topic is ranked first, and 0 when it is ranked last.

5.3 Long-term Effects

Participants in Experiment 1 were re-interviewed approximately 16 months after the audio screenings of *Wahapahapa* (EFM) and *Tamapendo* (HIV), with very high response rates (93%). Respondents were again asked to prioritize community goals and to choose among hypothetical candidates running on different platforms. As shown in Table 2, the effects of *Wahapahapa* (HIV) on priorities dissipated by the endline, but the voting effects remains large and statistically robust. The point estimate is 0.085 (SE=0.020), and the randomization inference p-value is 0.006. The effects of *Tamapendo*, which seemed equivocal for voting shortly after the audio screening, are weakly positive more than a year later. The fact that the long-term estimate for voting is positive suggests that the negative estimate from the midline was a statistical anomaly.

Long-term results from Experiment 2 also suggest that effects persisted. In this study, follow-up interviews occurred 16-17 months after the audio screenings, again with a response rate of 94%. The *Boda Bora* drama continued to influence both village priorities and vote preferences; those exposed to this drama ranked "reducing sexual violence" more highly and were more likely to vote for a candidate who pledges to take action on this issue. By the same token, the *Mikoko Yetu* drama continued to elevate both the priority ranking of environmental issues and the vote share of candidates who promised to prioritize environmental conservation.

Collectively, the four dramas produced an average shift in vote choice of 4.1 percentage points (SE = 0.013, RI p-value = 0.017) and an average shift in issue priority ranking of 0.023 (SE = 0.008, RI p-value = 0.040). The effects on the combined index at endline are jointly significant (p = 0.005) and roughly half the size of effects observed at midline.

Table 2: Effects of Narrative Dramas on Voting Preferences and Political Priorities
16-17 months after exposure

	Pooled			Experiment 1						Experiment 2					
				HIV			EFM			Enviro			GBV		
	Index	Vote	Rank	Index	Vote	Rank	Index	Vote	Rank	Index	Vote	Rank	Index	Vote	Rank
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Treatment	0.032***	0.041**	0.023**	0.049**	0.085***	0.013	0.024	0.014	0.028	0.027	0.032	0.021	0.032	0.035	0.029
Standard Error	0.009	0.013	0.008	0.014	0.020	0.014	0.019	0.023	0.015	0.019	0.028	0.017	0.019	0.025	0.017
RI p-value	0.005	0.017	0.040	0.018	0.006	0.269	0.186	0.332	0.106	0.171	0.216	0.196	0.133	0.173	0.129
Hypothesis	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Control Mean	0.49	0.50	0.47	0.55	0.50	0.59	0.49	0.56	0.41	0.41	0.38	0.43	0.51	0.55	0.48
Control SD	0.09	0.13	0.09	0.07	0.12	0.05	0.07	0.10	0.06	0.07	0.09	0.06	0.09	0.14	0.06
DV Range	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]
Blocked FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj-R^2	0.05	0.03	0.06	0.01	0.02	0.00	0.01	0.00	0.01	0.02	0.01	0.02	0.02	0.02	0.01
Observations	4,450	4,450	4,386	1,003	1,003	971	1,003	1,003	971	1,222	1,222	1,222	1,222	1,222	1,222

Note: * p < 0.1, ** p < 0.05, and *** p < 0.01. Standard errors clustered at the village level. RI p-value come from one-tailed tests in the hypothesized direction. "Vote" columns report results for the *Vote Preferences* outcome as described in the text. In summary, they report results for responses to the question: *"Imagine a village about one day's walk from here is having an election for village chairperson. There are two candidates giving speeches. Let me tell you about each one and you can tell me which of the two you think should be elected. The first candidate is named [randomize: Mr. Salim, Mr. John, Mrs. Mwanahidi, Mrs. Nema] and [he / she] promises to [Target topic]. Their slogan is [Target topic slogan]. The second candidate is named [randomize: Mr. Salim, Mr. John, Mrs. Mwanahidi, Mrs. Nema] and [he/she] promises to [randomize: 2 other issues]. Their slogan is [2 other slogans]. Which of these two candidates do you think should be elected?"* The responses are scored as 1 for voting for the targeted topic platform, 0 otherwise. "Rank" columns report results for the *Community Priorities* outcome as described in the text. In summary, the report results for responses to the question: *"Here is a set of cards, which show different goals for your village. Now, please rank them in order from the most important to the least."*. The responses are scored between 1 when the Target topic is ranked first, and 0 when it is ranked last.

5.4 Spillover Effects

Did the influence of narrative dramas extend to the spouses, children, and teenage children of screening attendees? In Experiment 1, enumerators interviewed the spouses of participants 16 months after the audio screenings of *Wahapahapa* and *Tamapendo* and interviewed the teenage children and friends of participants 18 months after the screenings. In Experiment 2, enumerators interviewed the female spouses and teenage children of participants 16 months after the screenings of *Boda Bora* and *Mikoko Yetu*.

As Table 3 demonstrates, the radio drama increased spouses' willingness to vote for candidates espousing platforms consistent with the drama's message as well as spouses' priority rankings of the issues that were the dramas' focus. Averaged across all dramas, the coefficients for voting and priority ranking are 0.024 and 0.027, respectively. The average effect on the combined index is 0.025 (SE = 0.01, RI p-value 0.048), about two-thirds the size of the effect on the drama attendees. This suggests that the effects of narrative dramas may extend beyond the audience members who consume them, which potentially amplifies narrative entertainment's cumulative influence. On the other hand, we observed no discernible influence of the dramas on the political priorities of the teenage children and close friends of attendees (Table A6 and Table A7).

Table 3: Effects of Narrative Dramas on Voting Preferences and Political Priorities of Spouses
16-17 months after exposure

	Pooled			Experiment 1						Experiment 2					
				HIV			EFM			Enviro			GBV		
	Index	Vote	Rank	Index	Vote	Rank	Index	Vote	Rank	Index	Vote	Rank	Index	Vote	Rank
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Treatment	0.025**	0.024	0.027*	0.023	-0.005	0.053**	0.012	0.045	-0.017	0.009	-0.002	0.020	0.071*	0.068*	0.073
Standard Error	0.010	0.016	0.013	0.019	0.026	0.019	0.017	0.030	0.022	0.018	0.037	0.020	0.031	0.037	0.044
RI p-value	0.048	0.130	0.087	0.222	0.563	0.049	0.288	0.154	0.689	0.367	0.523	0.249	0.064	0.097	0.126
Hypothesis	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Control Mean	0.57	0.49	0.64	0.57	0.54	0.59	0.45	0.51	0.39	0.36	0.34	0.38	1.06	0.52	1.60
Control SD	0.29	0.15	0.52	0.05	0.10	0.07	0.07	0.10	0.09	0.07	0.12	0.08	0.13	0.16	0.19
DV Range	[0-2]	[0-1]	[0-2]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-2]	[0-1]	[0-2]
Blocked FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj-R^2	0.35	0.02	0.62	0.01	0.01	0.01	0.00	-0.00	0.01	-0.02	-0.03	-0.01	0.02	0.01	0.01
Observations	2,071	2,067	2,054	703	703	686	686	682	686	341	341	341	341	341	341

Note: * p < 0.1, ** p < 0.05, and *** p < 0.01. Standard errors clustered at the village level. RI p-values come from one-tailed tests in the hypothesized direction. "Vote" columns report results for the *Vote Preferences* outcome as described in the text. In summary, they report results for responses to the question: *"Imagine a village about one day's walk from here is having an election for village chairperson. There are two candidates giving speeches. Let me tell you about each one and you can tell me which of the two you think should be elected. The first candidate is named [randomize: Mr. Salim, Mr. John, Mrs. Mwanahidi, Mrs. Nema] and [he / she] promises to [Target topic]. Their slogan is [Target topic slogan]. The second candidate is named [randomize: Mr. Salim, Mr. John, Mrs. Mwanahidi, Mrs. Nema] and [he/she] promises to [randomize: 2 other issues]. Their slogan is [2 other slogans]. Which of these two candidates do you think should be elected?"* The responses are scored as 1 for voting for the targeted topic platform, 0 otherwise. "Rank" columns report results for the *Community Priorities* outcome as described in the text. In summary, the report results for responses to the question: *"Here is a set of cards, which show different goals for your village. Now, please rank them in order from the most important to the least."*. The responses are scored between 1 when the Target topic is ranked first, and 0 when it is ranked last.

5.5 Assessing Treatment Effect Heterogeneity

We now turn our attention to whether and in what ways the primary treatment effects vary by the characteristics of the narrative drama and experimental subjects. The four dramas vary in important respects (Table 4). Two dramas directly model changes in political beliefs and behaviors (GBV and Environment); one (HIV) does mention politics and another (EFM) suggests political engagement is counterproductive. Two dramas (HIV, GBV) affirm widely held community priorities while two (EFM, Environment) challenge conventional prioritization. Three dramas (EFM, GBV, Environment) were set in recognizably Tanga villages while one (HIV) was set in a generic Tanzanian village. Two dramas (EFM, GBV) were written and performed by Tangan writers and actors, and two (HIV, Environment) were not.

Our overarching finding is that these differences are not essential determinants of narrative persuasion effects. The drama that generated the largest and most persistent impacts, HIV, did not directly model changes in political beliefs and was neither created nor set in Tanga. Nonetheless, the success of all four dramas despite their differences suggests that prevailing assumptions about the development of narrative dramas require more evidence, at least if the goal is to influence political priorities.

Table 4: Overview of Narrative Dramas

	Model of Political Activity	**Message Runs with Prevailing Priorities**	**Local Setting**	**Local Content Creation**	**Protagonist Identity**
EFM	Negative	No	Yes	Yes	Female, Youth
HIV	None	Yes	No	No	Male, Adult
GBV	Positive	Yes	Yes	Yes	Female, Youth
Environment	Positive	No	Yes	No	Male, Adult

Note: A summary of all dramas appears in the Appendix.

Are there more meaningful variation in treatment effects when we partition the *subjects* according to their covariate profiles? Machine learning tools automate and systematize the search for treatment effect heterogeneity across many covariates. Here, we use the Generalized Random

Forest (GRF) algorithm developed by Athey, Tibshirani and Wager (2019).

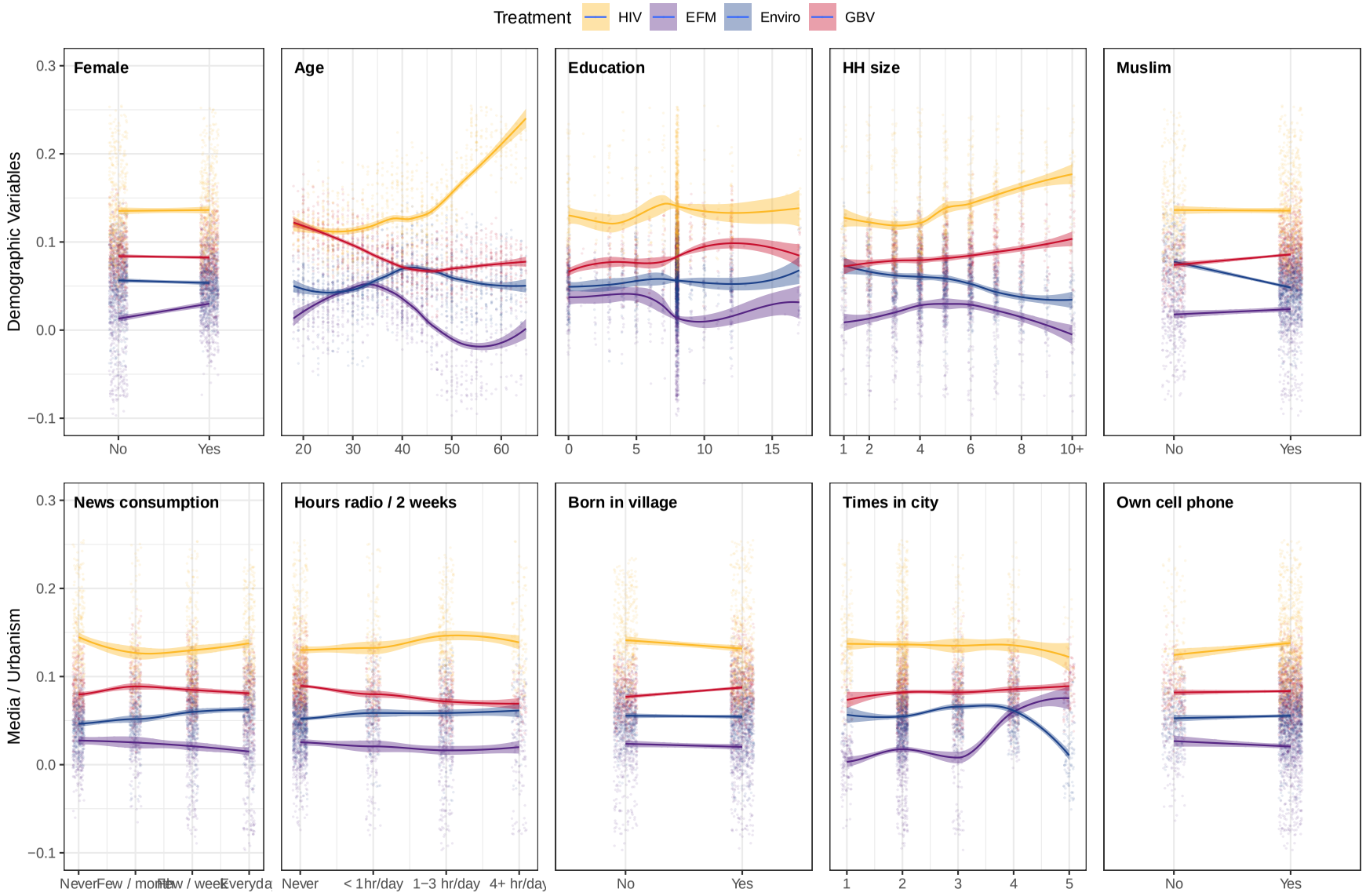
The starting point for our assessment of treatment effect heterogeneity is an omnibus test based on the cumulative distribution function of the estimated CATEs for each covariate profile (Chernozhukov et al. 2018). Applying this omnibus test to each of the four dramas, we find limited evidence of treatment effect heterogeneity across the four dramas. The smallest of the four p-values, for the HIV-drama, is 0.059 (see Figure A2). For the other dramas, CATEs vary across covariate profiles no more than would be expected by chance if all CATEs were truly identical. Turning to the long-term endline (Figure A3), evidence for treatment effect heterogeneity is nil: the smallest of the four p-values is 0.53.

The same conclusion emerges when we look specifically at the degree of treatment effect heterogeneity associated with each covariate. We use LOESS to flexibly characterize potentially nonlinear patterns of treatment-by-covariate interactions. Each of the tiled plots in Figure 1 describes how responsiveness to each of the four treatments changes as the values of each covariate change.

The tiled plots suggest that treatment effect heterogeneity is limited. Treatment effects do not seem to change appreciably across the range of each covariate: for the most part, the trajectories depicted in the graphs look flat. For example, those who say they "never" listen to the news have roughly the same CATE as those who report listening to news "every day." And in those rare instances where we see the modeled CATE change over the range of the covariate, as in the case of age, we do not see a similar pattern across all four dramas. Older respondents seem to be the strongly influenced by the HIV drama; younger respondents seem to be the most strongly influenced by GBV; the other two dramas seem to generate little by way of age-by-treatment interaction. Although one cannot rule out the possibility that interactions might turn up for covariates other than the ten featured here, further exploratory analysis that considers issue-specific covariates, such as pre-treatment baseline attitudes, provides no clear cases of treatment effect heterogeneity. Moreover, Figure A4 demonstrates that any suggestive indications of variation in CATEs over pre-treatment indicators dissipated 16-17 months after the treatment. It appears,

therefore, that the average treatment effect aptly summarizes the treatment effects experienced by a wide array of subgroups.

Figure 1: **Treatment-by-Covariate Interactions**, Political Priorities index



6 Discussion

The entertainment-education literature has repeatedly generated encouraging effects across a range of social issues. Most randomized trials to date have focused on changing beliefs, reducing stigma, increasing empathy, encouraging prosocial behaviors, and discouraging high-risk behaviors. Although some studies have addressed political issues such as corruption, relatively little attention has been devoted to politically relevant outcomes such as elevating the importance that voters attach to policy problems.

Pulling together evaluations of four audio dramas, this paper demonstrates that narrative dramas indeed shape audiences' policy priorities. Not only do we find effects that are statistically significant approximately two to four weeks after exposure, these effects are substantively large by almost any standard. Whether we consider the issues that respondents rank as priorities for their village or the votes they would cast for hypothetical candidates running on different issue platforms, exposure to a radio drama substantially increased the primacy of this issue. The average effect across all four dramas is slightly more than one-half of a village-level standard deviation.

Remarkably, these effects seem to persist for more than a year after audiences were exposed to the dramas. Although effects decay between midline and endline, it remains notable that a single exposure to narrative entertainment continues to shape audiences' priorities more than a year later. Moreover, endline survey data suggest that the effects of exposure to narrative dramas diffuse to at least some members of compliers' social network. The priorities expressed by spouses seem to show the indirect influence of exposure to the narrative dramas, but we find no apparent effects among compliers' children or friends. One wonders whether interpersonal diffusion would be more widespread were the audience exposed to an ongoing serialized drama as opposed to a condensed version presented in a single sitting.

Perhaps surprisingly, when we use machine learning methods to model treatment effect heterogeneity, we find that the narratives' agenda-setting effects are similar for respondents with very different covariate profiles. Variables such as education, religion, media consumption do

not seem to moderate the effects of the narrative messages. Put another way, the four dramas' effects are apparent across many segments of society, perhaps an indication of the extent to which narratives "transported" the audience and enticed listeners to identify with characters different from themselves. Whether unique features of narrative entertainment lead to more consistent effects than those exerted by news, campaigns, and other non-narrative forms of communication is an intriguing direction for future research.

Although evidence suggesting the lack of treatment effect heterogeneity may be taken to suggest the generality of agenda-setting effects, the broader research program of experimentally testing variations in narrative content remains in its infancy. We may infer that the *Tamapendo* (EFM) drama did nothing to inspire support for politicians who denounce early forced marriage because it portrayed local officials as aloof and unconcerned about this issue, but this inference is post hoc. Much better would be a head-to-head test of competing versions of *Tamapendo* (EFM), some of which portray sympathetic local officials who articulate and enforce norms against early forced marriage. In a similar vein, one might envision an experimental research program that examines whether modeling political participation as part of the story line induces audiences to go beyond prioritizing the featured issue, taking further behavioral steps such as expressing their views to public officials or working collectively to bring about change.

References

Athey, Susan, Julie Tibshirani and Stefan Wager. 2019. "Generalized Random Forests." *Annals of Statistics* 47(2):1148–1178.

Babeiya, Edwin. 2022. "Tactics and Pathologies of Party Primaries: A Test for Democratic Consolidation in Tanzania." *The African Review* pp. 1–24.

Bandura, Albert. 1976. *Social Learning Theory*. 1st edition ed. Englewood Cliffs, New Jersey: Prentice-Hall.

Bandura, Albert. 2004. "Health Promotion by Social Cognitive Means." *Health Education & Behavior* 31(2):143–164.

Banerjee, Abhijit, Eliana La Ferrara and Victor H Orozco-Olvera. 2019. "The Entertaining Way to Behavioral Change: Fighting HIV with MTV." *NBER Working Paper* .

Chernozhukov, Victor, Mert Demirer, Esther Duflo and Iván Fernández-Val. 2018. "Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments, with an Application to Immunization in India." *NBER Working Paper* .

Chipman, Hugh A., Edward I. George and Robert E. McCulloch. 2010. "BART: Bayesian Additive Regression Trees." *The Annals of Applied Statistics* 4(1).

Creel, A. H., R. N. Rimal, G. Mkandawire, K. Böse and J. W. Brown. 2011. "Effects of a Mass Media Intervention on HIV-related Stigma: 'Radio Diaries' Program in Malawi." *Health Education Research* 26(3):456–465.

Fazio, Lisa K., Nadia M. Brashier, B. Keith Payne and Elizabeth J. Marsh. 2015. "Knowledge Does Not Protect against Illusory Truth." *Journal of Experimental Psychology: General* 144(5):993–1002.

Frank, Lauren B. and Paul Falzone, eds. 2021. *Entertainment-Education Behind the Scenes: Case Studies for Theory and Practice*. Palgrave Macmillan.

Gerber, Alan, James Gimpel, Donald Green and Daron Shaw. 2011. "How large and long-lasting are the persuasive effects of televised campaign ads? Results from a randomized field experiment." *American Political Science Review* 105(1):135–150.

Green, Donald P., Anna M. Wilke and Jasper Cooper. 2020. "Countering Violence Against Women by Encouraging Disclosure: A Mass Media Experiment in Rural Uganda." *Comparative Political Studies* 53(14):2283–2320.

Green, Donald P., Dylan W. Groves and Constantine Manda. 2021. "A Radio Drama's Effects on HIV Attitudes and Policy Priorities: A Field Experiment in Tanzania." *Health Education & Behavior* 48(6):842–851.

Green, Donald P., Dylan W. Groves, Constantine Manda, Beatrice Montano and Bardia Rahmani. 2022. "A Radio Drama's Effects on Attitudes Toward Early and Forced Marriage: Results From a Field Experiment in Rural Tanzania." *Comparative Political Studies* p. 00104140221139385.

Green, Donald P., Dylan W. Groves, Constantine Manda, Beatrice Montano and Bardia Rahmani. 2024. "The Effects of Independent Local Radio on Tanzanian Public Opinion: Evidence from a Planned Natural Experiment." *The Journal of Politics* 86(1):231–240.

Green, Donald P. and Holger L. Kern. 2012. "Modeling Heterogeneous Treatment Effects in Survey Experiments with Bayesian Additive Regression Trees." *Public Opinion Quarterly* 76(3):491–511.

Green, Melanie C. and Timothy C. Brock. 2000. "The Role of Transportation in the Persuasiveness of Public Narratives." *Journal of Personality and Social Psychology* 79(5):701–721.

Humphreys, Macartan, Raul Sanchez de la Sierra and Peter van der Windt. 2013. "Fishing, Commitment, and Communication: A Proposal for Comprehensive Nonbinding Research Registration." *Political Analysis* 21(1):1–20.

Iyengar, Shanto and Donald R. Kinder. 1987. *News That Matters: Television and American Opinion*. News That Matters: Television and American Opinion University of Chicago Press.

Iyengar, Shanto and Donald R. Kinder. 2010. *News That Matters: Television and American Opinion, Updated Edition.* Chicago, IL: University of Chicago Press.

Iyengar, Shanto, Mark D. Peters and Donald R. Kinder. 1982. "Experimental Demonstrations of the "Not-So-Minimal" Consequences of Television News Programs." *American Political Science Review* 76(4):848–858.

Lapinski, Maria Knight and Paul Nwulu. 2008. "Can a Short Film Impact HIV-related Risk and Stigma Perceptions? Results from an Experiment in Abuja, Nigeria." *Health Communication* 23:403–412.

Lazarsfeld, P. F., B. Berelson and H. Gaudet. 1944. *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign, Legacy Edition.* The People's Choice Oxford, England: Duell, Sloan & Pearce.

McCombs, Maxwell E. and Donald L. Shaw. 1972. "The Agenda-Setting Function of Mass Media." *The Public Opinion Quarterly* 36(2):176–187.

McCombs, Maxwell and Sebastian Valenzuela. 2020. *Setting the Agenda: Mass Media and Public Opinion.* John Wiley & Sons.

Moyer-Gusé, Emily. 2008. "Toward a Theory of Entertainment Persuasion: Explaining the Persuasive Effects of Entertainment-Education Messages." *Communication Theory* 18(3):407–425.

Murrar, Sohad and Markus Brauer. 2018. "Entertainment-Education Effectively Reduces Prejudice." *Group Processes & Intergroup Relations* 21(7):1053–1077.

Paluck, Elizabeth Levy and Donald P. Green. 2009. "Deference, Dissent, and Dispute Resolution: An Experimental Intervention Using Mass Media to Change Norms and Behavior in Rwanda." *American Political Science Review* 103(4):622–644.

Rahmani, Bardia, Dylan Groves, Beatrice Montano and Donald Green. 2023. "Narrative Entertainment Changes Minds: A Meta-Analysis of Recent Experiments." *ISERP Working Papers* .

Shayo, Deodatus Patrick. 2021. "Citizen Participation in Local Government Elections in the Age of Crowdsourcing: Explorations and Considerations in Tanzania." *Program on Governance and Local Development Working Paper* 47.

Sinclair, Betsy. 2012. *The Social Citizen: Peer Networks and Political Behavior*. Chicago Studies in American Politics Chicago, IL: University of Chicago Press.

Slater, Michael D. and Donna Rouner. 2002. "Entertainment—Education and Elaboration Likelihood: Understanding the Processing of Narrative Persuasion." *Communication Theory* 12(2):173–191.

Tsfati, Yariv. 2003. "Does Audience Skepticism of the Media Matter in Agenda Setting?" *Journal of Broadcasting & Electronic Media* 47(2):157–.

Wanta, Wayne and Yi-Chen Wu. 1992. "Interpersonal Communication and the Agenda-Setting Process." *Journalism Quarterly* 69(4):847–855.

Weaver, David H., Jian-Hua Zhu and Lars Willnat. 1992. "The Bridging Function of Interpersonal Communication in Agenda-Setting." *Journalism Quarterly* 69(4):856–867.

Weghorst, Keith. 2022. *Activist Origins of Political Ambition: Opposition Candidacy in Africa's Electoral Authoritarian Regimes*. New York, NY: Cambridge University Press.

Wilke, Anna M., Donald P. Green and Benjamin Tan. 2022. "Encouraging Community Action Against Teacher Absenteeism: A Mass Media Experiment in Rural Uganda." *The Journal of Development Studies* 58(5):915–930.

Supplemental Materials for
Narrative Entertainment Shapes Policy Priorities

October 8, 2024

Contents

*The authors would like to thank the implementing partner, UZIKWASA, led by Novatus Urassa, with special thanks to Eric Mallya for his tireless efforts. We are grateful to the entire research team at Innovations for Poverty Action, particularly Martin Zuakulu, Rachel Jones, Kyu Khin Gar, Zachary Isdahl, Robert Mwandumbya, Neema Msechu, Gilbert Loshoock, Cosmas Sway, Zinabu Omary, Fatuma Yahaya, Frank Simon, Fadhili Mashaka, Stephano Kizigo, and Saulo Mwakisongo. We thank discussants and participants at the 2023 Meeting of the American Political Science Association. An earlier version of this paper benefitted from comments at the 2023 Meeting of the Association for Psychological Science and the 2023 Meeting of the Midwest Political Science Association. We thank Salma Mousa and Elizabeth Levy Paluck for their very helpful comments on earlier drafts. The studies described in this paper were reviewed and approved by Columbia Institutional Review Board (protocol AAAR5582) and Tanzania's Commission of Science and Technology (protocol 14528). These studies were made possible by financial support from Wellspring Philanthropic Fund. Corresponding author: dpg2110@columbia.edu

A Covariate Balance

Table A1: **Balance**

	Experiment 1			Experiment 2		
	EFM	HIV	RI p-value	GBV	Environment	RI p-value
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.50 (0.49)	0.52 (0.49)	0.786	0.50 (0.50)	0.52 (0.50)	0.999
Age	38.04 (11.95)	38.50 (11.88)	0.657	40.55 (13.13)	39.96 (12.78)	0.236
Education	7.16 (3.31)	7.35 (3.18)	0.780	7.24 (3.27)	7.36 (3.15)	0.662
HH Size	4.82 (2.41)	5.00 (2.21)	0.745	5.05 (2.09)	5.05 (2.17)	0.501
Muslim	0.64 (0.48)	0.65 (0.47)	0.821	0.75 (0.43)	0.80 (0.40)	0.588
Daily news	0.28 (0.45)	0.29 (0.45)	0.563	0.20 (0.40)	0.24 (0.42)	0.807
Any radio	0.65 (0.48)	0.67 (0.47)	0.772	0.54 (0.50)	0.53 (0.50)	0.417
Radio frequency	1.05 (1.12)	1.02 (1.08)	0.286	0.73 (0.97)	0.76 (0.98)	0.660
Any news	0.27 (0.45)	0.28 (0.45)	0.579	0.30 (0.46)	0.28 (0.45)	0.254
News per week	1.57 (1.17)	1.58 (1.18)	0.558	1.39 (1.12)	1.50 (1.13)	0.839
Born in village	0.55 (0.50)	0.58 (0.49)	0.627	0.58 (0.49)	0.59 (0.49)	0.614
Ever visit city	0.15 (0.36)	0.18 (0.38)	0.804	0.01 (0.11)	0.04 (0.20)	0.954
Own cell phone	0.85 (0.36)	0.80 (0.40)	0.113	0.78 (0.42)	0.81 (0.40)	0.857
Visit city frequency	2.33 (0.99)	2.31 (1.01)	0.414	2.95 (1.05)	2.81 (1.03)	0.161
Gender equality index	0.72 (0.22)	0.74 (0.21)	0.944	2.28 (1.03)	2.40 (1.03)	0.919
Know others in village	2.70 (0.76)	2.65 (0.74)	0.287	2.52 (0.83)	2.61 (0.84)	0.944

Note: * p < 0.1, ** p < 0.05, and *** p < 0.01. Standard deviations appear in parentheses. Randomization Inference p-values are based on 5,000 re-randomizations.

B Compliance and Attrition

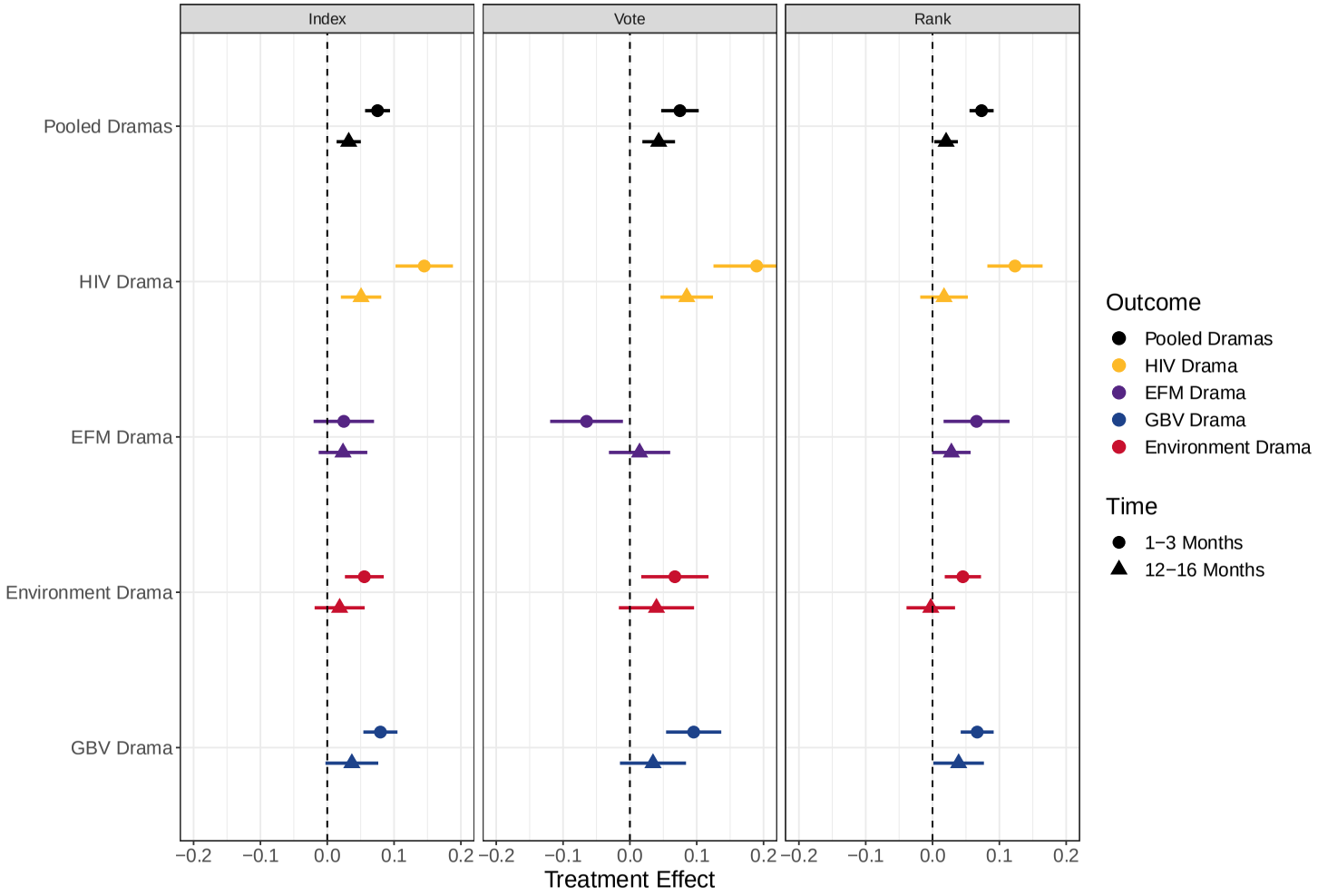
Table A2: **Compliance and Attrition**
Midline and Endline

	Experiment 1			Experiment 2		
	EFM	HIV	RI p-value	GBV	Environment	RI p-value
	(1)	(2)	(3)	(4)	(5)	(6)
Compliance	0.856 (0.351)	0.862 (0.345)	0.606	0.948 (0.221)	0.913 (0.282)	0.026
Midline Attrition	0.045 (0.207)	0.043 (0.204)	0.339	0.010 (0.101)	0.021 (0.142)	0.881
Endline Attrition	0.036 (0.187)	0.033 (0.180)	0.415	0.044 (0.206)	0.034 (0.181)	0.116

Note: * p < 0.1, ** p < 0.05, and *** p < 0.01. Standard deviations appear in parentheses. Compliance is coded as 1 if the respondent attended any part of the screening and 0 otherwise. Midline attrition is coded as 1 if the respondent completed the baseline survey but did not complete the midline survey and 0 if the respondent completed the baseline and midline survey. Endline attrition is coded as 1 if the respondent completed the baseline survey but did not complete the endline survey and 0 if the respondent completed the baseline and endline survey. Randomization Inference p-values are based on 5,000 re-randomizations. Although one p-value falls below the 0.05 significance level, it would no longer be significant if we were to apply a Bonferroni correction for multiple comparisons.

C Visualization of Primary Results

Figure A1: Effects of Narrative Dramas on Voting Preferences and Political Priorities



D Did the Placebo Treatment Affect Outcome Measures Other Dramas?

The maintained assumption of the placebo-controlled design is that the effect of the treatment drama did not influence political priorities unrelated to the treatment. To test this assumption, we can construct a rank order of priorities with the topic of the treatment dramas (e.g. EFM, HIV, Environment, and GBV) removed. Do respondents in treatment and placebo villages prioritize unrelated political issues differently at midline or endline? We find no evidence that they do: the treatment effect on only one of 29 unrelated priorities exhibited a randomization inference p-value below 0.05, less than would be exhibited by random chance.

Table A3: Effects of Narrative Dramas on Political Priorities of Issues Unrelated to the Drama

	Experiment 1, 2-4 Weeks After Exposure		
Variable	EFM Mean (SD)	HIV Mean (SD)	RI p-value
Food	1.06 (1.24)	1.08 (1.25)	0.794
Roads	1.47 (1.26)	1.29 (1.23)	0.258
Electricity	1.64 (1.20)	1.79 (1.22)	0.292
	Experiment 1, 16-17 Months After Exposure		
Variable	EFM Mean (SD)	HIV Mean (SD)	RI p-value
Education	6.89 (1.78)	6.92 (1.80)	0.812
Electricity	5.83 (1.88)	6.01 (2.04)	0.370
Sanitation	6.64 (1.78)	6.42 (1.87)	0.324
Roads	6.93 (1.84)	6.64 (1.83)	0.118
Crime	4.27 (2.12)	3.93 (1.88)	0.086
Justice	4.70 (1.72)	4.92 (1.83)	0.338
	Experiment 2, 2-4 Weeks After Exposure		
Variable	GBV Mean (SD)	Enviro Mean (SD)	RI p-value
Cell Service	2.08 (0.27)	2.06 (0.25)	0.704
Water	2.92 (0.27)	2.94 (0.25)	0.704
Alcoholism	3.03 (0.78)	2.99 (0.80)	0.382
Loans	2.52 (0.69)	2.57 (0.71)	0.404
Misbehaving Youth	3.45 (0.70)	3.44 (0.70)	0.822
Bribery	2.77 (0.77)	2.88 (0.82)	0.044
Health	2.85 (0.36)	2.80 (0.40)	0.084
Agriculture	2.15 (0.36)	2.20 (0.40)	0.084
Bad Fathers	3.56 (0.50)	3.56 (0.50)	0.788
Low Religiosity	3.44 (0.50)	3.44 (0.50)	0.788
	Experiment 2, 16-17 Months After Exposure		
Variable	GBV Mean (SD)	Enviro Mean (SD)	RI p-value
Cell Service	2.14 (0.34)	2.12 (0.33)	0.628
Water	2.86 (0.34)	2.88 (0.33)	0.628
Alcoholism	3.11 (0.76)	3.16 (0.75)	0.574
Loans	2.49 (0.68)	2.48 (0.69)	0.882
Misbehaving Youth	3.40 (0.73)	3.37 (0.73)	0.604
Bribery	2.81 (0.76)	2.84 (0.77)	0.626
Health	2.85 (0.35)	2.83 (0.38)	0.270
Agriculture	2.15 (0.35)	2.17 (0.38)	0.270
Bad Fathers	3.61 (0.49)	3.64 (0.48)	0.342
Low Religiosity	3.39 (0.49)	3.36 (0.48)	0.342

E Treatment Effect Heterogeneity

E.1 Explanation of Machine Learning

This ML routine proceeds in three steps. First, the algorithm splits the data into a training sample, used to partition the data into sub-groups, and an estimation sample, used to estimate treatment effects across each sub-group.[1] Second, the algorithm uses the training sample to build regression trees using recursive partitioning. Trees start as a single node containing the entire training sample. The "root" node is then split into child nodes. To generate child nodes, the algorithm randomly selects a subset of variables as candidates for splitting. It then examines all possible values of each selected variable and considers splitting the sample into two nodes at this value. The algorithm identifies the variable and value that maximizes the "goodness of split" and divides the sample between observations that fall above and below the splitting value.[2] If the algorithm finds no valid splits at a given node, the node is considered a leaf of the final tree. When all branches of a tree reach their terminal node, the tree is considered defined.

After generating a regression tree using the training sample, the algorithm fits the estimation sample to the tree by placing each observation in the terminal node that matches its covariate profile. The algorithm then calculates a predicted average treatment effect at each terminal node and assigns it to every observation in the node. A respondent's predicted conditional average treatment effect (CATE) is calculated as the average predicted treatment effect across 25,000 trees.[3]

E.2 Assessing Treatment Effect Heterogeneity

To this point, we have focused on the average treatment effect of each drama at midline and endline. In this section we turn our attention to the question of whether and in what ways these effects vary when we partition the subjects according to their covariate profiles.

When the sources of treatment effect heterogeneity are not known *ex ante*, restricting analyses to pre-registered hypotheses may hinder the discovery of unanticipated results; on the other hand, ad hoc exploration of treatment effect heterogeneity runs the risk of "fishing" and false discovery (Humphreys, Sanchez de la Sierra and van der Windt 2013).

Machine learning tools, originally developed to predict individualized responses to medical interventions, automate and systematize the search for treatment effect heterogeneity across many covariates. Here, we use the Generalized Random Forest (GRF) algorithm developed by Athey, Tibshirani and Wager (2019).[4]

Our search for treatment-by-covariate interactions focused on ten variables that pertain to all four dramas.[5] Five reflect basic demographic characteristics: gender, age, education, household

[1] Partitioning the data into two different samples to be used as training and estimation samples allows to satisfy a condition called *honesty* to reduce bias in the search (Athey, Tibshirani and Wager 2019).

[2] If a potential split creates a node with fewer than 10 treatment or 10 placebo observations, it is rejected. We use 10 as the minimum number of observations criteria, but note that we obtain similar estimates when the minimum number of observations per leaf is set to 5, 10, or 15.

[3] For a practical reference see the GRF package documentation, and for a more theoretical treatment of the GRF algorithm see Athey, Tibshirani and Wager (2019).

[4] GRF is one of several machine learning approaches. Similar results are obtained using Bayesian additive regression trees (Chipman, George and McCulloch 2010; Green and Kern 2012).

[5] A total of sixteen covariates was fed to the GRF package; we focus only on the ten of substantive interest. None of the other ten, which are listed in Appendix I, interacts with the treatment in any apparent way and are given low importance scores by the GRF algorithm.

size, and whether the respondent is Muslim. Another five variables might be characterized as relating to exposure to outside ideas: news consumption, number of hours listening to radio over the previous two weeks, whether the respondent was born in the village, and cell phone ownership. We model outcomes measured a few weeks after treatment so as to be able to leverage data from both experiments. In order to maximize power, we model an additive index of the two outcome measures — community importance and voting preferences — by taking a simple average of the two scores.[6]

The starting point for our assessment of treatment effect heterogeneity is an omnibus test based on the cumulative distribution function of the estimated CATEs for each covariate profile (**?**). Under the null hypothesis of no treatment effect heterogeneity, the CDF is mostly flat, as variation among CATEs is solely due to chance. On the other hand, an upward-sloping CDF function suggests a high degree of treatment effect heterogeneity, as some covariate profiles show much larger effects than others.

Applying this omnibus test to each of the four dramas, we find limited evidence from treatment effect heterogeneity across the four dramas. The smallest of the four p-values, for the HIV-drama, is 0.047 (see Figure A2). For the other dramas, CATEs vary across covariate profiles basically no more than would be expected by chance if all CATEs were truly identical. Turning to the long-term endline, evidence for treatment effect heterogeneity is nil. The smallest of the four p-values is 0.27.

Although the omnibus test is instructive, it lacks power (**?**, p. 4). A more telling assessment looks specifically at the degree of treatment effect heterogeneity associated with specific covariates. In order to describe how the estimated CATEs vary with each covariate, we use LOESS to flexibly characterize potentially nonlinear patterns of treatment-by-covariate interactions. Each of the tiled plots in Figure 1 describes how treatment responsiveness changes as the values of each covariate change. To facilitate comparison for each experiment and each type of outcome measure, the plots describe the CATE trajectory for each of the four dramas. Flat LOESS lines imply homogeneous treatment effects; conversely, fitted lines that move up or down on the vertical dimension imply treatment-by-covariate interactions. Especially noteworthy are instances in which the fitted lines all move up or down in unison, as this suggests a robust interaction that transcends one particular topic.

On the whole, the evidence suggests that treatment effect heterogeneity is limited. Treatment effects do not seem to change appreciably across the range of each covariate; for the most part, the trajectories depicted in the graphs look flat. For example, those who say they "never" listen to the news have roughly the same CATE as those who report listening to news "every day." And in those rare instances where we see the modeled CATE change over the range of the covariate, as in the case of age, we do not see a similar pattern across all four dramas. Older respondents seem to be the strongly influenced by the HIV drama; younger respondents seem to be the most strongly influenced by GBV; the other two dramas seem to generate little by way of age-by-treatment interaction. Although one cannot rule out the possibility that interactions might turn up for covariates other than the ten featured here, further exploratory analysis that considers issue-specific covariates, such as pre-treatment baseline attitudes, provides no clear

[6]The two outcomes are mildly correlated for each of the four drama's outcomes. When we focus solely on respondents who attended placebo audioscreenings (so as to assess the correlation in the absence of treatment), the correlation between community priorities and voting preferences is 0.06 for environmental protection, 0.17 for gender-based violence, 0.29 for HIV, and 0.32 for early/forced marriage.

cases in of covariate-predicted treatment responsiveness. Moreover, Figure A4 demonstratives that any suggestive indications of variation in CATEs over pre-treatment indicators dissipated or reversed 15 months after the treatment. It appears, therefore, that the average treatment effect aptly summarizes the treatment effects experienced by a wide array of subgroups.

Figure A2: **Distribution of Individual CATEs**

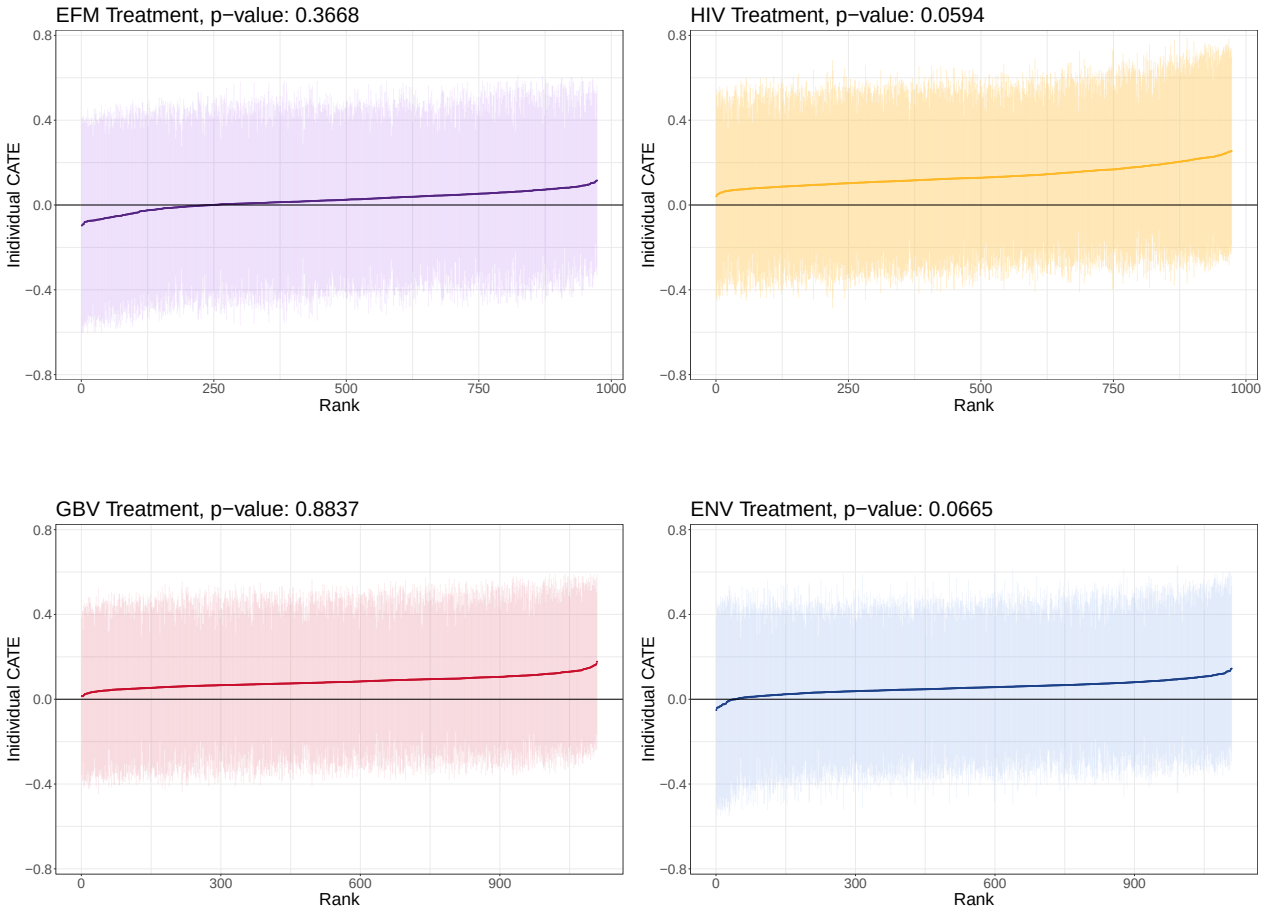


Figure A3: **Distribution of Individual CATEs** [*Endline*]

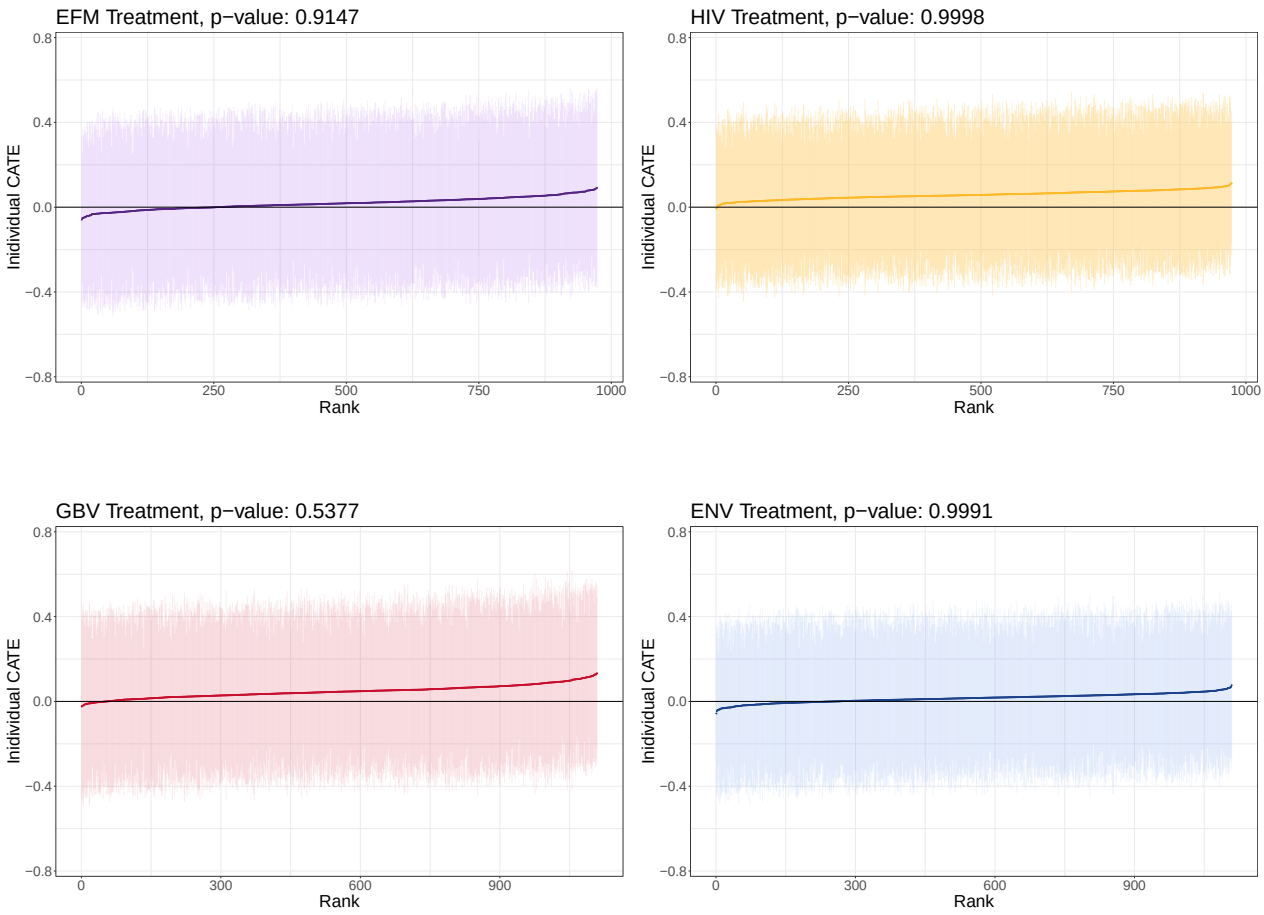
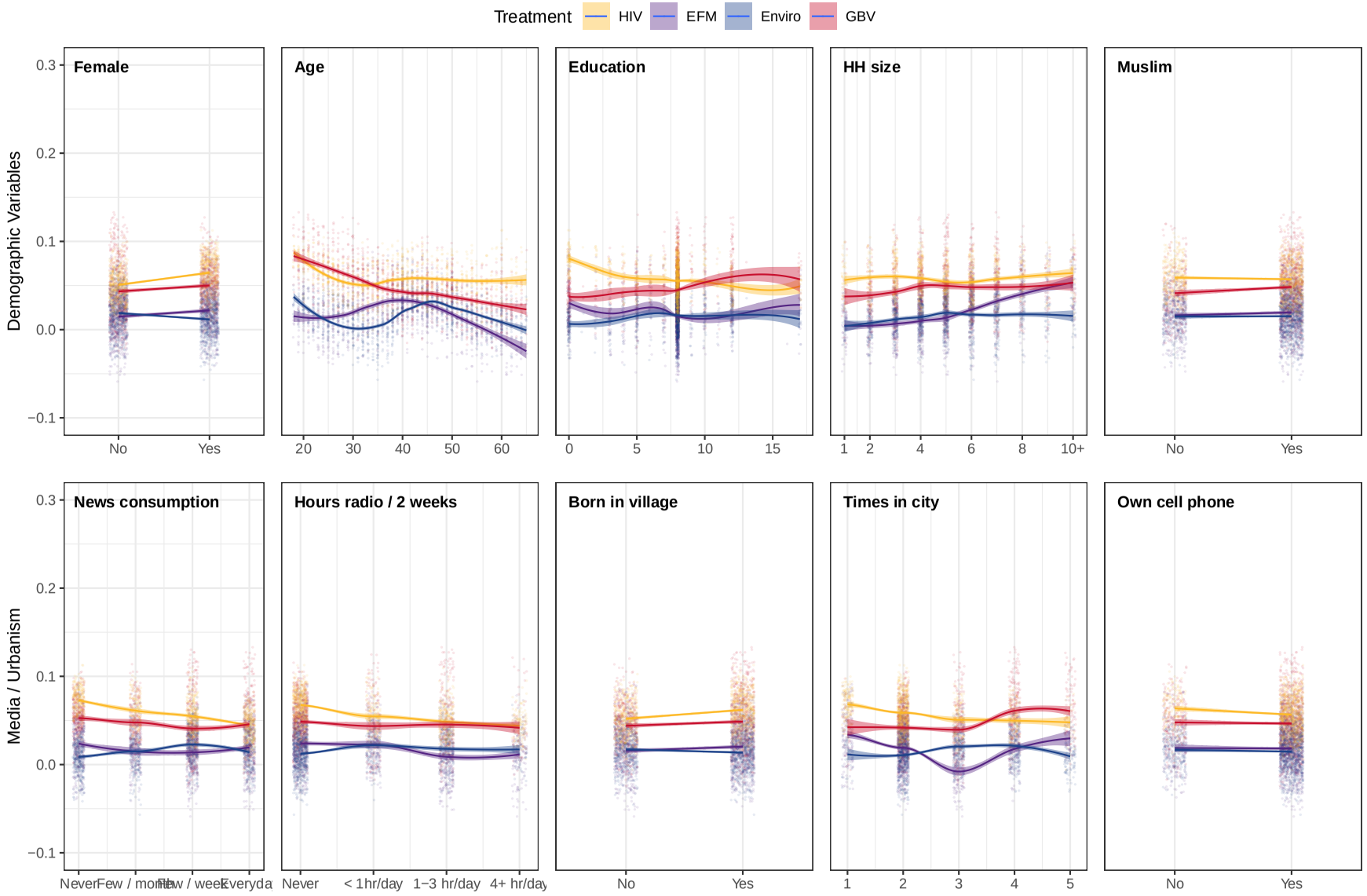


Figure A4: **Treatment-by-Covariate Interactions [*Endline*]**, Political Priorities index



F Outcome Measures

F.1 Voting

F.1.1 Midline (2-4 weeks)

Early and Forced Marriage

- **Question**: Imagine a village about one day's walk from here is having an election for village chairperson. There are two candidates giving speeches. Let me tell you about each one and you can tell me which of the two you think should be elected. The first candidate is named [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] and [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] promises to [*fight against child marriage, make hiv-aids treatment more available, improve roads, crack down on stealing in the village*]. The second candidate is named [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] and [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] promises to [*fight against child marriage, make hiv-aids treatment more available, improve roads, crack down on stealing in the village*]. Which of these two candidates do you think should be elected?
- **Coding**: 1 if respondent votes for candidate with "fight against child marrage" platform, 0 if candidate votes against "fight against child marriage" platform.
- Respondents who were asked to vote between the candidate with the "fight against child marriage" platform and the "making hiv-aids treatment more available" platform were dropped.

HIV/AIDS

- **Question**: Imagine a village about one day's walk from here is having an election for village chairperson. There are two candidates giving speeches. Let me tell you about each one and you can tell me which of the two you think should be elected. The first candidate is named [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] and [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] promises to [*fight against child marriage, make hiv-aids treatment more available, improve roads, crack down on stealing in the village*]. The second candidate is named [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] and [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] promises to [*fight against child marriage, make hiv-aids treatment more available, improve roads, crack down on stealing in the village*]. Which of these two candidates do you think should be elected?
- **Coding**: 1 if respondent votes for candidate with "make hiv-aids treatment more available" platform, 0 if candidate votes against "make hiv-aids treatment more available" platform.
- Items in brackets were assigned with equal probability. Names were matched to the religion the names are commonly associated with (e.g. Mr. John and Mrs. Rose were always given as Christian)

- Respondents who were asked to vote between the candidate with the "fight against child marriage" platform and the "making hiv-aids treatment more available" platform were dropped.

Gender Based Violence

- **Question**: Imagine a village about one day's walk from here is having an election for village chairperson. There are two candidates giving speeches. Let me tell you about each one and you can tell me which of the two you think should be elected." The first candidate is named [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] is a [*Muslim/Christian*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] promises to fight against sexual violence in the village. Their slogan is Protect our Girls from Sugar Daddies and Rapists. The second candidate is named [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] is a [*Muslim/Christian*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] promises to [*improve roads in the village. Their slogan is Make Our Roads Better/improve education in our village. Their Slogan is Better Schools for our Children*]. Which of these two candidates do you think should be elected?"

- **Coding**: 1 if respondent votes for candidate with "fight against sexual violence" platform, 0 if candidate votes against "fight against sexual violence" platform.

- Items in brackets were assigned with equal probability. Names were matched to the religion the names are commonly associated with (e.g. Mr. John and Mrs. Rose were always given as Christian)

- Whether the candidate with the "fight against sexual violence" platform appeared before or after the opposing candidate was randomized.

Environment

- **Question**: Imagine a village about one day's walk from here is having an election for village chairperson. There are two candidates giving speeches. Let me tell you about each one and you can tell me which of the two you think should be elected. "The first candidate is named [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] is a [*Muslim/Christian*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] promises to protect the environment. Their slogan is Cut Down One Tree, Plant Many Trees. The second candidate is named [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] is a [*Muslim/Christian*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] promises to [*improve roads in the village. Their slogan is Make Our Roads Better/improve education in our village. Their Slogan is Better Schools for our Children*]. Which of these two candidates do you think should be elected?"

- **Coding**: 1 if respondent votes for candidate with "protect the environment" platform, 0 if candidate votes against "protect the environment" platform.

- Items in brackets were assigned with equal probability. Names were matched to the religion the names are commonly associated with (e.g. Mr. John and Mrs. Rose were always given as Christian)

- Whether the candidate with the "protect the environment" platform appeared before or after the opposing candidate was randomized.

F.1.2 Endline / Partner / Friend Survey (18 months)

Early and Forced Marriage

- **Question**: Imagine a village about one day's walk from here is having an election for village chairperson. There are two candidates giving speeches. Let me tell you about each one and you can tell me which of the two you think should be elected.

  "The first candidate is named [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] is a [*Muslim/Christian*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] promises to reduce the number of early marriages in the village.

  The second candidate is named [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] is a [*Muslim/Christian*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] promises to [reduce crime in the village / improve the roads in the village]

  Which of these two candidates do you think should be elected?"

- **Coding**: 1 if respondent votes for candidate with 'reduce the number of early marriages in the village" platform, 0 if candidate votes against "reduce the number of early marriages in the village" platform.

- Items in brackets were assigned with equal probability. Names were matched to the religion the names are commonly associated with (e.g. Mr. John and Mrs. Rose were always given as Christian)

- Whether the candidate with the "reduce the number of early marriages in the village" platform appeared before or after the opposing candidate was randomized.

HIV / AIDS

- **Question**: Imagine a village about one day's walk from here is having an election for village chairperson. There are two candidates giving speeches. Let me tell you about each one and you can tell me which of the two you think should be elected.

  "The first candidate is named [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] is a [*Muslim/Christian*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] promises villagers to make HIV/AIDS treatment more available in the village.

  The second candidate is named [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] is a [*Muslim/Christian*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] promises to [reduce crime in the village / improve the roads in the village]

  Which of these two candidates do you think should be elected?"

- **Coding**: 1 if respondent votes for candidate with ‘reduce the number of early marriages in the village” platform, 0 if candidate votes against “reduce the number of early marriages in the village” platform.
- Items in brackets were assigned with equal probability. Names were matched to the religion the names are commonly associated with (e.g. Mr. John and Mrs. Rose were always given as Christian)
- Whether the candidate with the “reduce the number of early marriages in the village” platform appeared before or after the opposing candidate was randomized.

Gender Based Violence

- **Question**: Imagine a village about one day’s walk from here is having an election for village chairperson. There are two candidates giving speeches. Let me tell you about each one and you can tell me which of the two you think should be elected." The first candidate is named [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] is a [*Muslim/Christian*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] promises to fight against sexual violence in the village. Their slogan is Protect our Girls from Sugar Daddies and Rapists. The second candidate is named [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] is a [*Muslim/Christian*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] promises to [*improve roads in the village. Their slogan is Make Our Roads Better/improve education in our village. Their Slogan is Better Schools for our Children*]. Which of these two candidates do you think should be elected?"
- **Coding**: 1 if respondent votes for candidate with “fight against sexual violence” platform, 0 if candidate votes against “fight against sexual violence” platform.
- Items in brackets were assigned with equal probability. Names were matched to the religion the names are commonly associated with (e.g. Mr. John and Mrs. Rose were always given as Christian)
- Whether the candidate with the “fight against sexual violence” platform appeared before or after the opposing candidate was randomized.

Environment

- **Question**: Imagine a village about one day’s walk from here is having an election for village chairperson. There are two candidates giving speeches. Let me tell you about each one and you can tell me which of the two you think should be elected. "The first candidate is named [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] is a [*Muslim/Christian*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] promises to protect the environment. Their slogan is Cut Down One Tree, Plant Many Trees. The second candidate is named [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] is a [*Muslim/Christian*]. [*Mr. Salim/Mr. John/Mrs. Rose/Mrs. Mwanaidi*] promises to [*improve roads in the village. Their slogan is Make Our Roads Better/improve education in our village. Their Slogan is Better Schools for our Children*]. Which of these two candidates do you think should be elected?"

- **Coding**: 1 if respondent votes for candidate with "protect the environment" platform, 0 if candidate votes against "protect the environment" platform.
- Items in brackets were assigned with equal probability. Names were matched to the religion the names are commonly associated with (e.g. Mr. John and Mrs. Rose were always given as Christian)
- Whether the candidate with the "protect the environment" platform appeared before or after the opposing candidate was randomized.

F.2 Political Priorities

F.2.1 Midline (2-4 weeks)

Early and Forced Marriage

- **Question**: Here is another set of cards, which show different goals for your village. Please choose the three that are currently the most important to you, and the item that is least important: Reduce the number of people who do not have enough food to eat; Reduce the incidence of forced marriage; Increase the number of roads; Increase the availability of water; Increase the availability of electricity; Reduce the amount of crime; Increase the availability of medicine for HIV/AIDS.
- **Coding**: 1 if EFM was ranked first, 0.66 if ranked second, 0.33 if ranked third, and 0 if unranked. If HIV/AIDS was ranked ahead of EFM, the score moved up one category (e.g. from 0.33 to 0.66).

HIV/AIDS

- **Question**: Here is another set of cards, which show different goals for your village. Please choose the three that are currently the most important to you, and the item that is least important: Reduce the number of people who do not have enough food to eat; Reduce the incidence of forced marriage; Increase the number of roads; Increase the availability of water; Increase the availability of electricity; Reduce the amount of crime; Increase the availability of medicine for HIV/AIDS.
- **Coding**: 1 if HIV/AIDs was ranked first, 0.66 if ranked second, 0.33 if ranked third, and 0 if unranked. If EFM was ranked ahead of HIV/AIDS, the score moved up one category (e.g. from 0.33 to 0.66).

GBV

- **Question 1**: Here is set of cards, which show different goals for your village. Please place the cards in order from least important to most important. Options: Reducing sexual violence; Access to water; Improved cell phone reception.
- **Coding**: 1 if sexual violence ranked first, 0.5 if ranked second, 0 if ranked third

- **Question 2**: Here is set of cards, which show different social problems in villages in Tanzania. Now, please put them in order, from biggest problem to smallest problem. Options: Sexual violence against young girls; Alcoholism; Not paying back loans; Kids not going to school and people not working.
- **Coding**: 1 if sexual violence against young girls ranked first, 0.66 if ranked second, 0.33 if ranked third, 0 if ranked last.
- **Index Coding**: Sum of two priority measures divided by 2.

Environment

- **Question 1**: Here is set of cards, which show different goals for your village: Caring for the environment; Health + clinics; Increasing agricultural equipment. Please place the cards in order from least important to most important.
- **Coding**: 1 if caring for the environment ranked first, 0.5 if ranked second, 0 if ranked third
- **Question 2**: Here is set of cards, which show different social problems in villages in Tanzania. Now, please put them in order, from biggest problem to smallest problem. Options: Hurting the environment; Men abandoning their family; Bribing of public officials; Not practicing religion much.
- **Coding**: 1 if hurting the environment ranked first, 0.66 if ranked second, 0.33 if ranked third, 0 if ranked last.
- **Index Coding**: Sum of two priority measures divided by 2.

F.2.2 Endline (18 months)

Early and Forced Marriage

- **Question**: Here is another set of cards, which show different goals for your village. Please put them in order, from most important to you to least important to you: Increasing agriculture and finshing equipment; Reducing crime in the village; Reducing early force marriage; Education / schools; Justice and resolving of problems eg land problems; Availabilty of electricity; Sanitation/waste management; Improvement of infrastructure and roads; Availability of health services like increasing access to HIV/AIDS medication.
- **Coding**: First, create exclude "increasing acces to HIV/AIDS medication." From the remaining list, score 1 if EFM was ranked first, 0.85 if ranked second, 0.71 if ranked third, 0.57 if ranked fourth, 0.42 if ranked fifth, 0.28 if ranked sixth, 0.14 if ranked seventh, and 0 if ranked eighth.

HIV/AIDS

- **Question**: Here is another set of cards, which show different goals for your village. Please put them in order, from most important to you to least important to you: Increasing agriculture and finshing equipment; Reducing crime in the village; Reducing early force marriage; Education / schools; Justice and resolving of problems eg land problems; Availabilty of electricity; Sanitation/waste management; Improvement of infrastructure and roads; Availability of health services like increasing access to HIV/AIDS medication.

- **Coding**: First, create exclude "increasing acces to HIV/AIDS medication." From the remaining list, score 1 if HIV/AIDS medication was ranked first, 0.85 if ranked second, 0.71 if ranked third, 0.57 if ranked fourth, 0.42 if ranked fifth, 0.28 if ranked sixth, 0.14 if ranked seventh, and 0 if ranked eighth.

- **Question 1**: Here is set of cards, which show different goals for your village. Please place the cards in order from least important to most important. Options: Reducing sexual violence; Access to water; Improved cell phone reception.

- **Coding**: 1 if sexual violence ranked first, 0.5 if ranked second, 0 if ranked third

- **Question 2**: Here is set of cards, which show different social problems in villages in Tanzania. Now, please put them in order, from biggest problem to smallest problem. Options: Sexual violence against young girls; Alcoholism; Not paying back loans; Kids not going to school and people not working.

- **Coding**: 1 if sexual violence against young girls ranked first, 0.66 if ranked second, 0.33 if ranked third, 0 if ranked last.

- **Index Coding**: Sum of two priority measures divided by 2.

Environment

- **Question 1**: Here is set of cards, which show different goals for your village: Caring for the environment; Health + clinics; Increasing agricultural equipment. Please place the cards in order from least important to most important.

- **Coding**: 1 if caring for the environment ranked first, 0.5 if ranked second, 0 if ranked third

- **Question 2**: Here is set of cards, which show different social problems in villages in Tanzania. Now, please put them in order, from biggest problem to smallest problem. Options: Hurting the environment; Men abandoning their family; Bribing of public officials; Not practicing religion much.

- **Coding**: 1 if hurting the environment ranked first, 0.66 if ranked second, 0.33 if ranked third, 0 if ranked last.

- **Index Coding**: Sum of two priority measures divided by 2.

G Tables

G.1 LASSO Adjusted Results

Table A4: Effects of Narrative Dramas on Voting Preferences and Political Priorities
LASSO-Adjusted, 2-4 weeks after exposure

	Pooled			Experiment 1						Experiment 2					
				HIV			EFM			Enviro			GBV		
	Index	Vote	Rank	Index	Vote	Rank	Index	Vote	Rank	Index	Vote	Rank	Index	Vote	Rank
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Treatment	0.078***	0.074***	0.079***	0.140***	0.175***	0.120***	0.025	-0.057	0.066**	0.049**	0.068**	0.049***	0.090***	0.098***	0.079***
Standard Error	0.009	0.015	0.008	0.022	0.031	0.020	0.023	0.030	0.024	0.015	0.033	0.011	0.013	0.025	0.009
RI p-value	<0.001	<0.001	<0.001	<0.001	0.001	0.001	0.217	0.937	0.049	0.016	0.040	0.002	<0.001	0.004	<0.001
Hypothesis	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Control Mean	0.38	0.45	0.35	0.25	0.39	0.18	0.36	0.56	0.26	0.41	0.38	0.43	0.48	0.49	0.47
Control SD	0.12	0.14	0.14	0.09	0.15	0.08	0.11	0.15	0.10	0.06	0.09	0.05	0.06	0.08	0.06
DV Range	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]
Blocked FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	65	44	72	16	14	17	26	8	22	19	0	14	12	6	16
Adj-R^2	0.10	0.04	0.16	0.09	0.09	0.06	0.08	0.06	0.08	0.05	0.00	0.08	0.05	0.03	0.05
Observations	4,504	3,822	4,462	1,001	660	1,001	1,001	660	1,001	1,251	1,251	1,230	1,251	1,251	1,230

Note: * p < 0.1, ** p < 0.05, and *** p < 0.01. Standard errors clustered at the village level. RI p-value come from one-tailed tests in the hypothesized direction. "Vote" columns report results for the *Vote Preferences* outcome as described in the text. The responses are scored as 1 for voting for the targeted topic platform, 0 otherwise. "Rank" columns report results for the *Community Priorities* outcome as described in the text. The responses are scored between 1 when the Target topic is ranked first, and 0 when it is ranked last.

Table A5: Effects of Narrative Dramas on Voting Preferences and Political Priorities
LASSO-Adjusted, 16-17 months after exposure

	Pooled			Experiment 1						Experiment 2					
				HIV			EFM			Enviro			GBV		
	Index	Vote	Rank	Index	Vote	Rank	Index	Vote	Rank	Index	Vote	Rank	Index	Vote	Rank
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Treatment	0.032***	0.041**	0.023**	0.048**	0.085***	0.012	0.024	0.016	0.028	0.026	0.034	0.020	0.039*	0.044	0.030
Standard Error	0.009	0.013	0.008	0.022	0.025	0.017	0.018	0.023	0.017	0.019	0.031	0.014	0.020	0.025	0.018
RI p-value	0.005	0.017	0.040	0.018	0.006	0.285	0.170	0.298	0.100	0.165	0.217	0.152	0.092	0.119	0.133
Hypothesis	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Control Mean	0.49	0.50	0.47	0.55	0.50	0.59	0.49	0.56	0.41	0.41	0.38	0.43	0.51	0.55	0.48
Control SD	0.09	0.13	0.09	0.07	0.12	0.05	0.07	0.10	0.06	0.07	0.09	0.06	0.09	0.14	0.06
DV Range	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]
Blocked FE	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Covariates	53	49	72	0	4	1	14	15	5	12	9	21	16	17	15
Adj-R^2	0.06	0.04	0.08	0.00	0.02	0.01	0.05	0.03	0.05	0.06	0.03	0.06	0.08	0.04	0.14
Observations	4,450	4,450	4,386	1,003	1,003	971	1,003	1,003	971	1,222	1,222	1,222	1,222	1,222	1,222

Note: * p < 0.1, ** p < 0.05, and *** p < 0.01. Standard errors clustered at the village level. RI p-value come from one-tailed tests in the hypothesized direction. "Vote" columns report results for the *Vote Preferences* outcome as described in the text. The responses are scored as 1 for voting for the targeted topic platform, 0 otherwise. "Rank" columns report results for the *Community Priorities* outcome as described in the text. The responses are scored between 1 when the Target topic is ranked first, and 0 when it is ranked last.

G.2 Spillovers - Friends

Table A6: Effects of Narrative Dramas on Voting Preferences and Political Priorities of Friends
16-17 months after exposure

	Pooled			Experiment 1					
				HIV			EFM		
	Index	Vote	Rank	Index	Vote	Rank	Index	Vote	Rank
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment	0.004	-0.008	0.018	0.016	0.002	0.031	-0.008	-0.018	0.005
Standard Error	0.008	0.016	0.012	0.012	0.025	0.017	0.012	0.019	0.016
RI p-value	0.375	0.637	0.173	0.186	0.469	0.132	0.683	0.742	0.422
Hypothesis	+	+	+	+	+	+	+	+	+
Control Mean	0.56	0.60	0.52	0.59	0.58	0.61	0.53	0.63	0.43
Control SD	0.07	0.09	0.11	0.05	0.08	0.06	0.07	0.08	0.08
DV Range	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]
Blocked FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj-R^2	0.03	0.01	0.11	0.01	0.01	0.02	0.01	0.01	0.00
Observations	1,924	1,921	1,924	962	961	962	962	960	962

Note: * p < 0.1, ** p < 0.05, and *** p < 0.01. Standard errors clustered at the village level. "Vote" columns report results for the *Vote Preferences* outcome as described in the text. In summary, they report results for responses to the question: *"Imagine a village about one day's walk from here is having an election for village chairperson. There are two candidates giving speeches. Let me tell you about each one and you can tell me which of the two you think should be elected. The first candidate is named [randomize: Mr. Salim, Mr. John, Mrs. Mwanahidi, Mrs. Nema] and [he / she] promises to [Target topic]. Their slogan is [Target topic slogan]. The second candidate is named [randomize: Mr. Salim, Mr. John, Mrs. Mwanahidi, Mrs. Nema] and [he/she] promises to [randomize: 2 other issues]. Their slogan is [2 other slogans]. Which of these two candidates do you think should be elected?"* The responses are scored as 1 for voting for the targeted topic platform, 0 otherwise. "Rank" columns report results for the *Community Priorities* outcome as described in the text. In summary, the report results for responses to the question: *"Here is a set of cards, which show different goals for your village. Now, please rank them in order from the most important to the least."*. The responses are scored between 1 when the target topic is ranked first, and 0 when it is ranked last.

G.3 Spillovers - Children

Table A7: Effects of Narrative Dramas on Voting Preferences and Political Priorities of Children 16-17 months after exposure

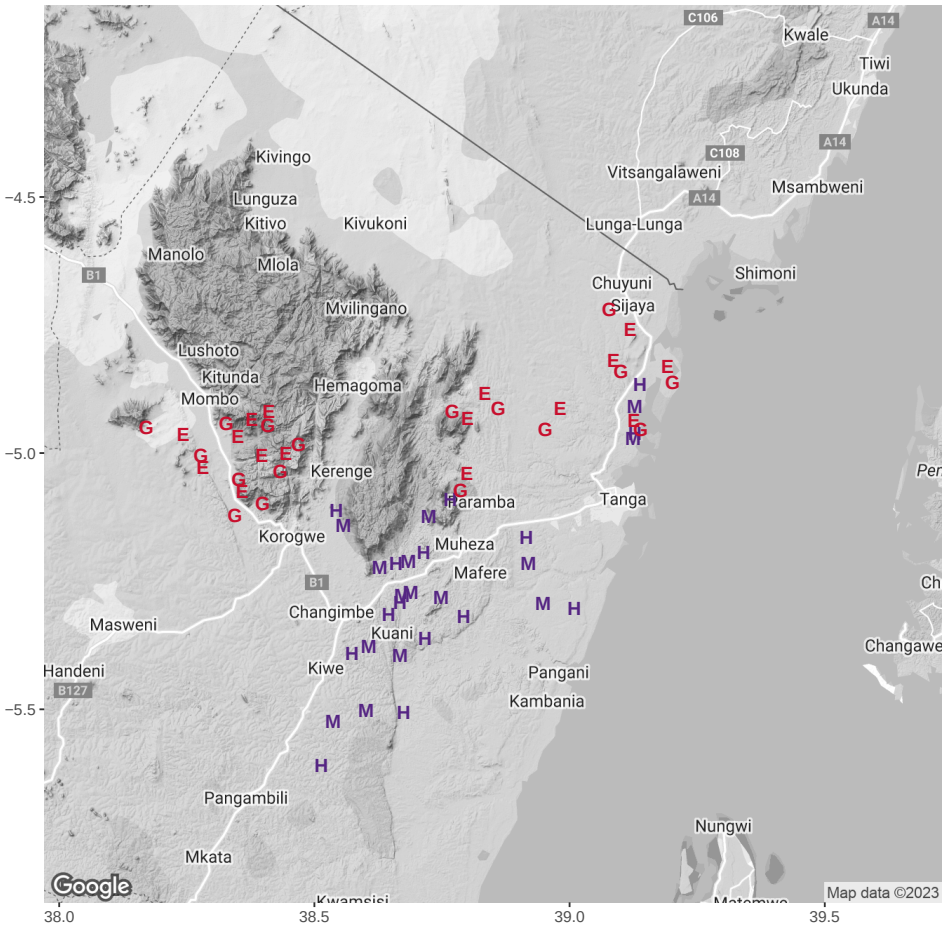
		Experiment 1		Experiment 2	
	Pooled	HIV	EFM	Enviro	GBV
	Vote	Vote	Vote	Vote	Vote
	(1)	(2)	(3)	(4)	(5)
Treatment	-0.026	-0.026	-0.069	-0.066	0.074*
Standard Error	0.017	0.025	0.028	0.043	0.038
RI p-value	0.858	0.774	0.952	0.878	0.070
Hypothesis	+	+	+	+	+
Control Mean	0.60	0.66	0.74	0.36	0.57
Control SD	0.21	0.14	0.13	0.24	0.14
DV Range	[0-1]	[0-1]	[0-1]	[0-1]	[0-1]
Blocked FE	Yes	Yes	Yes	Yes	Yes
Adj-R^2	0.10	0.01	0.03	0.06	0.03
Observations	1,626	482	482	331	331

Note: * p < 0.1, ** p < 0.05, and *** p < 0.01. Standard errors clustered at the village level. "Vote" columns report results for the *Vote Preferences* outcome as described in the text. In summary, they report results for responses to the question: *"Imagine a village about one day's walk from here is having an election for village chairperson. There are two candidates giving speeches. Let me tell you about each one and you can tell me which of the two you think should be elected. The first candidate is named [randomize: Mr. Salim, Mr. John, Mrs. Mwanahidi, Mrs. Nema] and [he / she] promises to [Target topic]. Their slogan is [Target topic slogan]. The second candidate is named [randomize: Mr. Salim, Mr. John, Mrs. Mwanahidi, Mrs. Nema] and [he/she] promises to [randomize: 2 other issues]. Their slogan is [2 other slogans]. Which of these two candidates do you think should be elected?"* The responses are scored as 1 for voting for the targeted topic platform, 0 otherwise.

H Intervention Details

H.1 Map

Figure A5: **Geographic Location of the Two Experiments**
Experiments 1 and 2



Experiment 1 (Red): M = EFM Drama, H = HIV Drama

Experiment 2 (Blue): G = GBV Drama, E = Environment Drama

H.2 Timeline

- Experiment 1 (July 2019 - March 2021)
  - July-August 2019: Baseline survey
  - July-August 2019: Audio screenings (2-3 days after baseline survey in each village)
  - August-September 2019: Midline survey (2-3 weeks after intervention in each village)
  - December 2020: Endline survey and partner survey (16 months after intervention in each village)

  - February 2021: Friends and children survey (20 months after intervention in each village)

- Experiment 2 (April 2022 - September 2023)
  - April-May 2022: Baseline survey
  - April-May 2022: Audio screenings (2-3 days after baseline survey in each village)
  - May-June 2022: Midline survey (4 weeks after intervention in each village)
  - August-September 2023: Endline survey (16 months after intervention in each village)
  - August-September 2023: Partner and children survey (2 weeks after endline survey in each village)

I Variables Included in Generalized Random Forest (GRF)

(1) Gender: Male (0) or Female (1)
(2) Age: Integer
(3) Education: None (0) to Graduate (20)
(4) Household size: Integer
(5) Religion: Christian (0) or Muslim (1)
(6) Listen to news daily: No (0) or Yes (1)
(7) Ever listen to radio in last 3 months: No (0) or Yes (1)
(8) Frequency of radio listening last two weeks: Never (0) or Multiple times per day (4)
(9) Ever consume media: No (0) or Yes(1)
(10) Frequency of following news:
(11) Born in village where respondent lives: No (0) or Yes (1)
(12) Have never visited a city: No (0) or Yes (1)
(13) Own cell phone: No (0) or Yes (1)
(14) Frequency of visiting city
(15) Gender equality attitudes index
(16) Number of others known in the village