MESSAGES ON COVID-19 PREVENTION IN INDIA INCREASED SYMPTOMS
REPORTING AND ADHERENCE TO PREVENTIVE BEHAVIORS AMONG 25
MILLION RECIPIENTS WITH SIMILAR EFFECTS ON NON-RECIPIENT MEMBERS
OF THEIR COMMUNITIES

Abhijit Banerjee
Marcella Alsan
Emily Breza
Arun G. Chandrasekhar
Abhijit Chowdhury
Esther Duflo
Paul Goldsmith-Pinkham
Benjamin A. Olken

Messages on COVID-19 Prevention in India Increased Symptoms Reporting and Adherence to Preventive Behaviors Among 25 Million Recipients with Similar Effects on Non-recipient Members of Their Communities
Abhijit Banerjee, Marcella Alsan, Emily Breza, Arun G. Chandrasekhar, Abhijit Chowdhury, Esther Duflo, Paul Goldsmith-Pinkham, and Benjamin A. Olken
NBER Working Paper No. 27496
July 2020
JEL No. D8,D83,I1,I15,I18,O1

ABSTRACT

During health crises, like COVID-19, individuals are inundated with messages promoting health-preserving behavior. Does additional light-touch messaging by a credible individual change behavior? Do the features of the message matter? To answer this, we conducted a large-scale messaging campaign in West Bengal, India. Twenty-five million individuals were sent an SMS containing a 2.5-minute clip, delivered by West Bengal native and 2019 Nobel laureate Abhijit Banerjee. All messages encouraged reporting symptoms to the local public health worker. In addition, each message emphasizes one health-preserving behavior (distancing or hygiene) and one motivation for action (effects on everyone or just on self). Further, some messages addressed concerns about ostracism of the infected. Messages were randomized at the PIN code level. As control, three million individuals received a message pointing them to government information. The campaign (i) doubled the reporting of health symptoms to the community health workers (p = 0.001 for fever, p = 0.024 for respiratory symptoms); (ii) decreased travel beyond one's village in the last two days by 20% (p = 0.026) (on a basis of 37% in control) and increased estimated hand-washing when returning home by 7% (p = 0.044) (67.5% in control); (iii) spilled over to behaviors not mentioned in the message – mask-wearing was never mentioned but increased 2% (p = 0.042), while distancing and hygiene both increased in the sample where they were not mentioned by similar amounts as where they were mentioned; (iv) spilled over onto nonrecipients within the same community, with effects similar to those for individuals who received the messages.

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

Marcella Alsan
Kennedy School of Government
Harvard University
79 John F. Kennedy St.
Rubenstein Bldg R403
Cambridge, MA 02138
and NBER
marcella_alsan@hks.harvard.edu

Emily Breza
Harvard University
Littauer Center, M28
1805 Cambridge Street
Cambridge, MA 02138
and NBER
ebreza@fas.harvard.edu

Arun G. Chandrasekhar
Department of Economics
Stanford University
579 Serra Mall
Stanford, CA 94305
and NBER
arungc@stanford.edu

Abhijit Chowdhury
John C. Martin Centre for Liver Research and Innov
Kolkata, West 7000150
India
abhijit@liverfoundation.in

Esther Duflo
Department of Economics, E52-544
MIT
50 Memorial Drive
Cambridge, MA 02142
and NBER
eduflo@mit.edu

Paul Goldsmith-Pinkham
Yale School of Management
165 Whitney Avenue
P.O. Box 208200
New Haven, CT 06520-8200
paulgp@gmail.com

Benjamin A. Olken
Department of Economics, E52-542
MIT
50 Memorial Drive
Cambridge, MA 02142
and NBER
bolken@mit.edu

1. INTRODUCTION

At the outset of the COVID-19 pandemic, governments faced the need to rapidly inform their population about the new disease: its symptoms, mechanisms of transmission, and behaviors that limit its spread. In particular, policymakers sought to increase hygienic behaviors, such as handwashing, social distancing, and mask-wearing (Lewnard and Lo, 2020; Prather et al., 2020; Kissler et al., 2020). In the absence of widespread testing, they also needed to rapidly put in place an effective system of reporting COVID-19 symptoms.

Very quickly, individuals were inundated with messages from numerous sources. In India, by the end of March and beginning of April 2020, individuals received messages about distancing and hygiene through television, radio, public signs, local government addresses, and even a short jingle accompanying outgoing mobile phone calls. Besides these official sources, they also receive messages via physical conversations, phone calls, WhatsApp, Facebook and Twitter.

In a survey in the Indian state of West Bengal, residents reported that on average in the two previous days, they heard about the importance of social distancing 20.2 times, about washing hands 16.9 times, and about wearing masks 17.2 times ($N = 408$) (see Appendix B, Table B2).

Absorbing and acting on multiple streams of information may be challenging and despite the ubiquity of messages, or perhaps because of it, compliance with basic recommendations is far from complete. Despite a nationwide lockdown in India during the time of this experiment, on average, 37% of our respondents in the control group left their village at least once every two days. Respondents report that typical villagers systematically wash their hands only 68% of the times when returning home. This suggests that the Indian government's public health messaging was either ignored, forgotten, misunderstood, or insufficiently disseminated within the community.

How can we create messages to have an impact in this setting of information overload? One potential avenue is celebrity messaging, a common public health tool in both normal times and the COVID-19 crisis. Our previous research on Twitter in Indonesia shows that there is a large premium for celebrity messaging in the context of vaccination, and the value is mostly driven by authentic messages coming from the celebrity's voice rather than echoing others such as the Ministry of Health (Alatas et al., 2019), but perhaps such messages get drowned in a context with such information overload? Further, assuming that messaging works, how should the message be

designed and diffused for effectiveness? In particular, should we provide comprehensive information about the best way to stay safe, which may be ignored (given the evidence that most people have limited attention span (Abaluck and Gruber, 2011; Beshears et al., 2013; Carvalho and Silverman, 2019)) or rely on the fact that the information is already out there and people mostly need a *nudge* to pay attention to it, in which case a pithy and memorable message might work better? And finally, is it important that we reach everyone directly or can we limit outreach and rely on the information diffusing through the community?

To answer these questions, we conducted a large-scale randomized controlled trial in West Bengal, a large state in India (population 91.3 million), with over two-thirds of the population living in rural areas.

We designed 2.5-minute-long video clips, delivered by Abhijit Banerjee, an author of this study and a public intellectual, who is native to West Bengal and grew up there. Since receiving the 2019 Sveriges Riksbank Prize in Economic Sciences in memory of Alfred Nobel (also known as the Nobel Prize in Economics), he has been covered widely in West Bengal media. He is also the chair of the West Bengal government's COVID-19 advisory board.

In the experiment, these video messages were sent via a link embedded in a text message to 25 million subscribers of the phone network Reliance Jio (henceforth Jio), randomly selected out of 28 million subscribers. The remaining three million received a text containing a link to the government website on COVID-19, similar to the many messages they have already received.

All video messages instructed individuals with cough and fever to contact their local frontline health worker. In addition, there were eight different variants emphasizing one practice (social distancing or hand-washing), one rationale for action (cost to self or cost to everyone including self), and a social problem (either an explicit statement that ostracism of COVID-19 victims is unacceptable and should be reported to the authorities, or no mention of the issue). Experimental message treatments were randomized at the PIN code (Postal Index Number) level across 1214 out of the state's 1264 PIN codes, covering an estimated 88 million people.[1]

A few days after the campaign, we conducted two surveys. First, we surveyed 677 frontline health workers from the sample communities to ask them about reporting

---

[1]India does not use PIN codes as an administrative unit, so we lack precise data about the population within each PIN code. We deliver messages in 1214 of the 1264 PIN codes in West Bengal, and estimate the total population in our study area by scaling the state's population of 91.3 million down by factor of 1214/1264. We present this estimate purely for illustrative purposes.

of symptoms. Second, we surveyed a random sample of 1,883 individuals drawn from a publicly-available database of former and current local village council (*gram panchayat*) members. We measured health-preserving behaviors, including whether individuals traveled outside their villages, the number of interactions within two arms' length (which was how the two meters mentioned in the message was interpreted in the local context), the estimated frequency of hand-washing upon returning home, and the use of masks or similar face covering when leaving home. We also measured the frequency of conversations that respondents had about COVID-19 (which included both face-to-face discussions, phone calls, or chats such as WhatsApp), beliefs about symptomatic and asymptomatic carriers, and knowledge of the virus. Our surveyed individuals included both those who received the message (Jio users) and those who were not messaged directly (non-Jio users).

Our results demonstrate large effects of treatment with significant spillovers both from message recipients onto non-recipients and also onto behaviors that were not emphasized in the message.

First, reporting of symptoms to the frontline health workers, critical to the tracking of the epidemic, doubled. Since the measurement was done within five days of the intervention, this almost certainly reflects reporting, rather than an increased level of sickness, given lag times for transmission and the time it takes for someone to become symptomatic. Second, rates of not leaving the village, washing hands and wearing masks improved significantly in treatment communities. Critically, the effects are of similar magnitudes for behaviors targeted in the videos (not traveling out, washing hands) and those not even mentioned (mask use). Third, distancing (hygiene) went up in the sample where only hygiene (distancing) was mentioned by more or less the same amount as when distancing (hygiene) was explicitly mentioned. Both these spillovers are consistent with the view that the message acted as a nudge. Finally, magnitudes of the effects are also similar between Jio users and others, suggesting significant community diffusion of health behaviors.

Overall, the results show that even against a background of a high level of messaging, an additional message by a respected public figure can still have large direct and indirect effects.

2. Context, Experiment, and Data

2.1. **Context.** This study took place in the state of West Bengal, India. West Bengal has a population of 91.3 million, with 62.2 million living in rural areas. The literacy

rate is 77.08% (Office of the Registrar General and Census Commissioner, India, 2011). As of March 19, 2019, around 57.72 million had access to mobile devices (Department of Telecommunications, India, 2019).

Community health workers (Accredited Social Health Activists, known as ASHAs) serve an important role in West Bengal's COVID-19 response as frontline health workers. Introduced as a part of the National Rural Health Mission in 2005, ASHAs are women selected from within the community itself, and each ASHA is responsible for serving the health needs of approximately 1000 residents. Prior to the pandemic, ASHAs primarily focused on maternal and newborn health, family planning, child vaccinations, and monitoring TB patients (further details are in Appendix A.1.2). During the pandemic, ASHAs have been re-purposed to become the frontline health workers responsible for connecting communities to the formal healthcare system for COVID-19 concerns (Ministry of Health and Family Welfare, India, 2020). They are tasked with community outreach, screening for symptoms, tracking return of migrants via door-to-door checks, and referring cases to higher healthcare facilities.

In response to COVID-19, India began a nationwide lockdown on March 24, 2020. At the time of our survey, India was in "Phase 3" of this lockdown, which began on May 4, 2020 and was scheduled to last for the next two weeks. During this period the country was divided into red, orange, and green zones. Red zones were those with high COVID-19 cases that were increasing steadily, orange zones were those with comparatively fewer cases, and green zones were those without any cases in the past 21 days. At the start of our information campaign, out of the 23 districts in West Bengal, 10 districts were red zones, five were orange, and eight were green. Prior to Phase 3, states prohibited all non-essential services, inter-district travel, and suspended all public transport. Phase 3 saw relaxations on some activities primarily in green and orange zones, including inter-district travel with limits on the number of passengers and resumption of public transport at 50 percent capacity.

Through all phases of the lockdown, individuals received messages on precautions, hygiene and distancing through public service announcements on television, radio, newspapers, and via text messages and jingles from their telecom service providers. State governments also ran information campaigns (including using celebrity videos) that were shared on social media. Apart from this, individuals received messages through their social networks via platforms like Facebook, WhatsApp and TikTok. In a separate survey we ran in West Bengal, respondents reported hearing about

social distancing 20.2 times, washing hands 16.9 times and wearing masks 17.2 times in the last 2 days ($N = 408$, Appendix B, Table B2).

2.2. **Experiment.** We delivered eight treatment messages in a cross-randomized $2 \times 2 \times 2 + 1$ design, described below. The exact scripts were crafted under the guidance of a physician member of the West Bengal COVID-19 advisory board, Abhijit Chowdhury, as well as physician Marcella Alsan, both co-authors of this study and are included in Appendix A.3. Each message emphasized a health behavior and a motivation for acting, and some messages addressed concerns of ostracism of the ill:

  T1: *Health Behavior:* Social distancing vs. Hygiene

  The message varied in whether it emphasized that people must maintain 2-meter distance (as mandated by the government), or the steps individuals must take to maintain hygiene (e.g., hand-washing with soap). Some other behaviors (e.g., not spitting) were mentioned in both and others (like wearing a mask) came up in neither.

  T2: *Motivation:* Externality + Internality vs. Internality

  The message varied in whether it emphasized that the illness could be damaging to others in one's community, especially the elderly and vulnerable (externality) or not. Both types of messages mentioned the effect on oneself (internality).

  T3: *Ostracism:* Anti-ostracism vs. No Mention

  The message varied in whether it explicitly stated that ostracism of COVID-19 victims is unacceptable and should be reported to the authorities, or stayed silent on that issue.

Every message concluded by encouraging the individual to contact a health worker if exhibiting symptoms of fever or coughing.

The messages were recorded by Abhijit Banerjee in separate 2.5-minute video clips posted privately on YouTube. Banerjee jointly won the Nobel Prize in Economics in October 2019 and is also the chair of the West Bengal COVID-19 advisory board. Both events, his winning the Nobel Prize and his role in advising in the state during the pandemic, have received widespread media coverage in West Bengal.

To deliver messages we partnered with Reliance Jio, one of the largest telecom operators in India. We randomized each of the 1214 study PIN codes to receive one

of the messages, and stratified our randomization at the district level [2]. The telecom partner then sent SMS messages to each of the 28 million subscribers on their network in the study PIN codes. All SMS messages were sent on May 4, 2020 and May 5, 2020.

Subscribers in 1085 treatment PIN codes received an SMS with a link to one of the eight video messages. The message contained the text: "Nobel laureate Abhijit Vinayak Banerjee's appeal on the subject of Coronavirus" and the unique YouTube link to the treatment's message. Those in 129 control PIN codes received an SMS with a link to a government website with COVID-19 information, similar to typical government messaging. The message contained the text: "An appeal on the subject of Coronavirus" and a link to a website with COVID-19 information.[3] In total, 24,585,927 Jio users received an SMS message with a link to a video message and 3,358,819 received the control message.

2.3. **Data.** We collected two datasets to measure the impact of the campaign. First, we surveyed 677 ASHAs by phone within 5 days of the messaging intervention to measure symptom reporting in their local communities. Second, from May 8, 2020 to May 19, 2020, we surveyed by phone a random sample of individuals drawn from a publicly available phone directory of previous and current village council (*gram panchayat*) members. This generated a cross-sectional dataset spanning nine days.

In the council member survey, we asked respondents whether they had traveled outside their village in the last 2 days. We also asked them the total number of individuals they interacted with within two arms' length, both within their own village as well as outside of their own village, over the last 2 days. To measure the level of conversation, we asked them the total number of people from whom they received information, or gave information about COVID-19 over the last 2 days, be it in person, over phone, on chat, or by other means. Hygiene practices were measured by asking respondents whether they wear a mask when they step out of the house, and how many times (out of 10) a typical person living in their village washes their hands after coming back home (we deliberately ask about community handwashing to

---

[2]India has 36 States and Union territories. Each of these is further divided into administrative units called districts. There are 19 districts in our study area. We use districts as specified in the most recent Indian Census.

[3]These messages contained the link: https://wb.gov.in/COVID-19.aspx

avoid demand effects). Respondents were also assessed for their knowledge of COVID-19 symptoms and precautions.[4] For non-binary variables (number of COVID-related conversations and number of in-person interactions), we winsorize at the 95%-ile level to deal with unreasonable outliers and further show robustness to the 97.5%-ile cut off in Appendix B, Table B3.

To keep the survey short, we collected minimal demographic information: gender, age, current council member status and smartphone access. 57.6% of our sample individuals belonged to a household where at least one member had a Jio connection. Henceforth, we refer to these individuals as the Jio sample. Compared to the non-Jio population, Jio users are significantly and substantially more likely to be male (21.5% vs. 14.3%, $p < 0.01$) and own a smartphone (81.1% vs. 61.5%, $p < 0.01$), and perhaps slightly younger as well (mean age 44.3 vs 45.8, t-test $p < 0.01$).

The rates of health-preserving behavior in the control villages are often low, but varied. On average over the past two days, despite the lockdown, 37% of individuals had travelled to other villages, the average respondent had interacted with 11 people within 2 arms' distance and reported washing their hands when they return home only 65.7%, of the time. However, 97.8% wear a mask or cover their face when going outside their house.

Finally, we have direct view counts of each video from YouTube (displayed in Appendix B, Figure B6). All treatments had similar viewing rates, on average 1.14%. The low viewing rate is consistent with the literature on click-through-rates, which documents low rates (0.3%-2.6%) (Richardson et al., 2007; Kanich et al., 2008). Despite this, given the large reach of the messaging campaign, 279,800 unique individuals eventually clicked on the link to watch the video. In what follows, we focus on treatment as receiving our treatment SMS (intention-to-treat from the perspective of the YouTube video) and the spillovers from the Jio to the non-Jio individuals within a PIN code.

3. Results

3.1. **Reporting Symptoms to Health Workers.** Because every message encouraged individuals to report symptoms to the ASHA, we examine how reporting rates

---

[4]We used the WHO list of symptoms to distinguish between correct and incorrect symptoms. Correct symptoms include cough, shortness of breath, sore throat, feverloss of taste, loss of smell, diarrhea, body aches, headaches, covid-toes, conjunctivitis, tiredness, chest pain, and rash. Incorrect symptoms include unny nose, vomiting, dizziness, itching,chills, and swelling in the legs and feet

TABLE 1. ASHA Regressions

VARIABLES	(1) Number of Fever Cases	(2) Number of Respiratory Cases
TREATMENT	0.272	0.187
	(0.082)	(0.083)
	[0.001]	[0.024]
Observations	677	677
District FE	✓	✓
Total Rounds FE	✓	✓
Smartphone FE	✓	✓
Control Mean	0.247	0.198

Notes: Both columns look at cases reported within a window of 5 days after broadcasting. We include fixed effects for district, total number of survey rounds, and smartphone access, and further control for the number of households the ASHA supervises. Standard errors are clustered at the PIN code level and reported in parentheses, and p-values are reported in brackets.

for fevers and respiratory issues responded to messaging. To isolate a *reporting* effect, as opposed to possible effects on disease transmission itself, we look at reports within a 5-day horizon of the intervention. Given that the average incubation period is 5.1 days (Li et al., 2020; Guan et al., 2020; Chan et al., 2020; Lauer et al., 2020), as well as time lags for disease transmission, limiting ourselves to reports within 5 days of the intervention means that is likely almost entirely capturing a change in reporting. To show robustness, we study 3 and 4 day window periods (Appendix B, Tables B4 and B5) and find similar results there as well.

We estimate treatment effects using the following regression:

$$(3.1) \qquad y_j = \beta \text{Treatment}_j + \delta' X_j + \epsilon_j$$

where y_j is, for ASHA worker j, the number of reports of fever or of respiratory symptoms. Reports include all means through which an ASHA might find out about a case: through regular home visits, over the phone, or via the patient or their household members approaching the ASHA in person. X_j is a vector of controls including smartphone access and the number of households supervised by the ASHA and fixed effects including district and total rounds of surveys done with the ASHA within a 5 day horizon of the intervention. All standard errors are clustered at the PIN code level. We find large effects on reporting. Reporting of fevers increases by

TABLE 2. Phone Survey Regressions

VARIABLES	(1) Did you travel outside your village?	(2) Number of interactions with people within 2 arms length	(3) Estimated % time washing hands upon returning home	(4) Did you use a mask?	(5) Number of conversations in person /online /mobile about COVID-19	(6) COVID-19 Knowledge Index
Panel A: Pooled						
TREATMENT	-0.074 (0.033) [0.026]	-1.473 (1.164) [0.206]	0.047 (0.023) [0.044]	0.019 (0.009) [0.042]	-2.099 (1.303) [0.108]	0.097 (0.123) [0.435]
Observations	1,883	1,875	1,821	1,883	1,881	1,883
District FE	✓	✓	✓	✓	✓	✓
Survey Day FE	✓	✓	✓	✓	✓	✓
Jio FE	✓	✓	✓	✓	✓	✓
Control Mean	0.370	11.052	0.675	0.978	11.678	5.135
Panel B: Jio						
TREATMENT	-0.061 (0.043) [0.155]	0.421 (1.527) [0.783]	0.012 (0.030) [0.694]	0.023 (0.014) [0.088]	-0.539 (1.604) [0.737]	0.227 (0.140) [0.106]
Observations	1,082	1,076	1,046	1,082	1,082	1,082
District FE	✓	✓	✓	✓	✓	✓
Survey Day FE	✓	✓	✓	✓	✓	✓
Control Mean	0.376	10.096	0.708	0.976	11.368	5.088
Panel C: Non Jio						
TREATMENT	-0.094 (0.057) [0.100]	-3.869 (2.073) [0.063]	0.088 (0.036) [0.015]	0.011 (0.012) [0.349]	-3.794 (2.121) [0.074]	-0.077 (0.199) [0.699]
Observations	801	799	775	801	799	801
District FE	✓	✓	✓	✓	✓	✓
Survey Day FE	✓	✓	✓	✓	✓	✓
Control Mean	0.362	12.190	0.636	0.981	12.048	5.190
Treat(Jio) = Treat(non-Jio)	0.645	0.112	0.100	0.492	0.221	0.174

Notes: All columns control for district and survey day fixed effects. Panel A also includes a Jio access fixed effect. The last row presents p-values for a test of equality between treatment effects in the Jio and non-Jio samples. Respondent level controls also include age, gender and smartphone access. Standard errors are clustered at the PIN code level and reported in parentheses, and p-values are reported in brackets.

110% relative to the control mean ($p = 0.001$) and reporting of respiratory symptoms increases by 94% ($p = 0.024$). In the absence of widespread community testing to track disease spread, this is an extremely important and promising result.

3.2. **Effects of Messaging.** Next we report on self-reports of behavior in the community. Table 2 presents estimated treatment effects, pooling all eight messages. Panel A presents results across the entire sample, including both Jio recipients and non-Jio community members. Panels B and C restrict results to Jio recipients and to non-Jio community members, respectively. In all cases, for respondent i in PIN code

p and on survey day t, we estimate

$$(3.2) \qquad y_{ipt} = \beta \text{Treatment}_p + \delta' X_{ipt} + \epsilon_{ipt}$$

where y_{ipt} is one of the following outcomes: traveling outside the village, interactions within two arms' length, percent of times washing hands upon returning home, mask usage when leaving home, conversations about COVID-19, and knowledge about COVID-19. Treatment$_p$ indicates whether the PIN code p was assigned to any of the eight treatment messages, and X_{ipt} is a vector of controls including a district fixed effect, a survey day fixed effect, a dummy for i being a Jio user, and controls for age, gender and smartphone access. All standard errors are clustered at the PIN code level.

We begin by examining behaviors mentioned explicitly in messaging (columns 1-3). In column 1, we see that being in a treated PIN code decreased travel outside one's village by 20% ($p = 0.026$). These effects are statistically indistinguishable in the Jio and non-Jio samples and if anything, appear to be smaller in magnitude in the Jio sample. As noted already, the Jio sample is quite different from the non-Jio sample so we focus on the fact that there is an effect in the non-Jio sample, rather than the difference between the two samples.

Treatment had no statistically detectable effect on social interactions that were not appropriately socially distanced, with a 13.3% decline ($p = 0.206$, column 2) relative to a base of 11.05 interactions with others within two arms' distance in control villages. This effect is larger among non-Jio users (31.7% decline, $p = 0.063$), but dampened by an insignificant 4.2% ($p = 0.783$) increase in interactions among Jio users.

Column 3 shows treatment increased the rate of estimated hand washing upon returning home from outside by 7% ($p = 0.044$) relative to the control village mean of 67.5%. Again, the effect is indistinguishable from zero for Jio users ($p = 0.694$) but large for non-Jio users, who report increases in handwashing by 13.8% ($p = 0.015$).

Column 4 reports impact on a behavior that was not directly targeted. Despite a high baseline use rates of masks (97.8%), treatment spilled over to that behavior: mask usage increased by 1.9pp ($p = 0.042$). We find this increase in both the Jio and non-Jio samples, though less precisely than in the pooled sample. In Appendix B, Table B3, we look at respondents' estimates of mask usage by others in their community and find a 4.0% ($p = 0.037$) increase relative to the control village mean of 77.2%.

The key result is thus that the intervention induced behavioral changes both in reporting of symptoms and in prevention behavior and spilling over on to behaviors that were not mentioned in the video, and these changes were at least as large in the sample that was not directly treated, indicating large-scale community diffusion.

The next two columns investigate whether direct communication was a likely mechanism for this community diffusion. Column (7) looks at knowledge about COVID-19. Knowledge also slightly improved in the Jio treated group (4.5%, $p = 0.106$), though not in the non-Jio group (the point estimate is a reduction of 1.4%, $p = 0.699$ in the knowledge index of this group). We don't find evidence that the intervention sparked conversations: in column 5 we estimate an 18% decline in the number of conversations (in-person/online/on mobile) regarding COVID-19 ($p = 0.108$). The drop is particularly large among non-Jio community members (31.4%, $p = 0.074$), while for Jio recipients it is indistinguishable from zero ($p = 0.737$). These findings are consistent with people having fewer in-person interactions within 2 arms' distance–as we report above, there is a large effect on this measure of distancing though it is not significant in the pooled sample–and not fully compensating for the missed conversations, be it in person, over phone, on chat, or by other means. This interpretation explains both the large drop in conversations for the non-Jio sample, who also report far fewer in-person interactions, and the small drop in conversations in the Jio sample, for whom there is no drop in in-person interactions. The finding that mass messaging doesn't necessarily spark conversations on a complicated topic is also consistent with other recent empirical work in India (Banerjee et al., 2019). Overall it does not appear likely that it is through conversations about COVID-19 that the intervention led to diffusion: it seems more likely to have come from direct observations and imitation of behavior.

3.3. **Effects by Content.** We next separately analyze the three message topics and show results in Figure 1. In Appendix B, Figures B4 and B5, we further disaggregate this analysis for both the Jio and non-Jio samples. Here, we fit a separate regression for each of the three topics (behavior, motivation, and ostracism), including dummies for the two message variants within a topic and omitting a dummy for the control group. We adjust for the same covariates as in Table 2 and use the specification

$$(3.3) \qquad y_{ipt} = \alpha + \beta_{m1} V_{pm1} + \beta_{m2} V_{pm2} + \delta' X_{ipt} + \epsilon_{ipt}.$$

Here m indexes each of the three regressions conducted, one for each of behavior, motivation, and ostracism. V_{pm1} and V_{pm2} are dummies for the message variants
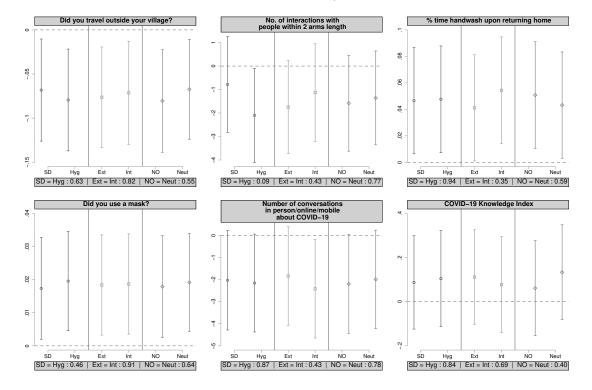
FIGURE 1. Effects by Content

Notes: This figure presents estimated treatment effects by message content, with each panel presenting the treatment effect on a separate outcome. For all regressions, we use district and survey day, as well as controls for age, gender, smartphone access, and Jio access. Within each panel, the point and whiskers give point estimate and 90% confidence intervals, and vertical lines separate estimates from different regressions. The note below each pane presents the p-value for a test of equality between the noted pairs of coefficients. Standard errors are clustered at the PIN code level. SD - "Social Distancing", Hyg - "Hygiene", Ext - "Externality + Internality", Int - "Internality Only", NO - "No Ostracism", Neut - "Neutral". Sample size varies from 1821 observations (estimated handwash) to 1883 observations (travel outside village).

(e.g., for the behavior topic, these are dummies for social distancing and hygiene, respectively), and β_{m1} and β_{m2} are the respective effects. Standard errors are clustered at the PIN code level. Each pane in Figure 1 plots estimates for β_{m1} and β_{m2} for a common outcome, across topics and the standard errors are clustered at PIN code level. To emphasize that we fit separate regressions for each topic, we use vertical bars within each pane to distinguish estimates from different regressions.

The main takeaway is that, for the most part, we cannot distinguish the impact of any particular emphasis within each topic. For example, we cannot reject that a message that spent more time on hygiene had the same impact on travel outside the

village as one explicitly geared to encourage social distancing, with minimal focus on hygiene ($p = 0.63$).

Taken together, the results suggest that messaging on any subtopic is sufficient to spillover not only to other people, but also onto other behaviors. This, combined with the fact that this is an information-rich environment where it is difficult to get people to pay attention to messages, suggests that short messages are likely to be most effective.

4. DISCUSSION

Public health behaviors and messages that inform and encourage them are thought to be crucial. Yet the marginal effect of a message in the setting of a rich information environment is unclear. This is particularly true in the context of the novel respiratory virus COVID-19, where individuals were bombarded with information that was initially evolving rapidly, and may have been confusing and frightening.

India has deployed a variety of different strategies to try and inform individuals about the health-promoting behaviors. These strategies include replacing mobile caller tune with an informational jingle on COVID-19 across telecom operators, getting popular celebrities to talk about COVID-19 on television and promoting COVID-19 related videos on the official YouTube channel (cal, 2020).[5] However, even in this rich information environment, there is still scope to improve public health behaviors since social distancing often fails and hand-washing rates are low.

We find large effects with large spillovers in a light-touch campaign. Messaging 27% of community members leads to large improvements in distancing and hygiene in the entire community and mask use, which was not part of the campaign, also went up.

Consistent with prior work (rates of 0.3%-2.6%), our SMS blast caused only a small fraction to watch the actual video messages (Richardson et al., 2007; Kanich et al., 2008). The share of videos watched on average is 1.14% and does not vary by message content. Combined with the effect on non-targeted behavior and non-targeted community members, this suggests that the video had indirect impacts on many who did not watch them.

This rapid diffusion of behavior is particularly intriguing in a context where social interactions were generally discouraged, and where our intervention itself aimed to

---

[5]A video from the government YouTube campaign can be found - here

reduce social interactions (and did to an extent, by reducing trips outside the village, and for the non-Jio population, intimate contact even with people inside the village). Indeed, we find reductions in conversations (in-person/online/on mobile) about COVID-19. This highlights that one cost of improved adherence to social distancing guidelines is a decrease in information flow in the community.

But despite this decreased information flows, reporting of symptoms and adherence to preventive behaviors increased. This could be due to a shift in social norms–people observe others adhering to the new rules and do the same.

Our results point to several important policy implications. First, there is value in persistent messaging, at least by celebrities. Though there is a lot of existing information, a new message can act as a useful nudge and reminder and this may be particularly important if the conditions change and updating the information is necessary. Second, given that governments have a fixed budget for messages and since within-community spillovers are large, it is important to disseminate the messages broadly across many communities, rather than instead using the same budget to target more people within a particular community. Third, the exact content of the message may not matter in such cases, since the information is out there and what is needed is a nudge to pay attention to it. Indeed, the literature finds that very long messages hurt comprehension so the message should be short rather than comprehensive (Abaluck and Gruber, 2011; Beshears et al., 2013; Carvalho and Silverman, 2019). Fourth, behavior and knowledge are not one-to-one. We have observed large shifts in distancing, hygiene, and mask-wearing despite little-to-no shifts in knowledge. Emphasis on behavior seems to be paramount. On net, the results suggest an important role for messaging by credible individuals.

References

(2020): "India TV, Jio, Airtel spreading coronavirus precautions through caller tunes: Here's how it works," https://www.indiatvnews.com/technology/news-jio-airtel-coronavirus-precautions-caller-tunes-how-it-works-596078, accessed on 16 June 2020.

ABALUCK, J. AND J. GRUBER (2011): "Choice inconsistencies among the elderly: Evidence from plan choice in the Medicare Part D program," *American Economic Review*, 101, 1180–1210.

ALATAS, V., A. G. CHANDRASEKHAR, M. MOBIUS, B. A. OLKEN, AND C. PAL-
ADINES (2019): "When Celebrities Speak: A Nationwide Twitter Experiment Pro-
moting Vaccination in Indonesia," Tech. rep., National Bureau of Economic Re-
search.

BANERJEE, A. V., E. BREZA, A. G. CHANDRASEKHAR, AND B. GOLUB (2019):
"When less is more: Experimental evidence on information delivery during India's
demonetization," *Available at SSRN 3163930.*

BESHEARS, J., J. J. CHOI, D. LAIBSON, AND B. C. MADRIAN (2013): "Simplifi-
cation and saving," *Journal of Economic Behavior & Organization*, 95, 130–145.

CARVALHO, L. AND D. SILVERMAN (2019): "Complexity and Sophistication," Tech.
rep., National Bureau of Economic Research.

CHAN, J. F.-W., S. YUAN, K.-H. KOK, K. K.-W. TO, H. CHU, J. YANG,
F. XING, J. LIU, C. C.-Y. YIP, R. W.-S. POON, ET AL. (2020): "A familial
cluster of pneumonia associated with the 2019 novel coronavirus indicating person-
to-person transmission: A study of a family cluster," *The Lancet*, 395, 514–523.

DEPARTMENT OF TELECOMMUNICATIONS, INDIA (2019): "Telecom Statistics India-
2019," https://dot.gov.in/sites/default/files/Telecom%20Statistics%
20India-2019.pdf?download=1, accessed 04 June 2020.

GUAN, W.-J., Z.-Y. NI, Y. HU, W.-H. LIANG, C.-Q. OU, J.-X. HE, L. LIU,
H. SHAN, C.-L. LEI, D. S. HUI, ET AL. (2020): "Clinical characteristics of
coronavirus disease 2019 in China," *New England journal of medicine*, 382, 1708–
1720.

KANICH, C., C. KREIBICH, K. LEVCHENKO, B. ENRIGHT, G. M. VOELKER,
V. PAXSON, AND S. SAVAGE (2008): "Spamalytics: An empirical analysis of spam
marketing conversion," in *Proceedings of the 15th ACM conference on Computer
and communications security*, 3–14.

KISSLER, S. M., C. TEDIJANTO, M. LIPSITCH, AND Y. GRAD (2020): "Social
distancing strategies for curbing the COVID-19 epidemic," *medRxiv.*

LAUER, S. A., K. H. GRANTZ, Q. BI, F. K. JONES, Q. ZHENG, H. R. MEREDITH,
A. S. AZMAN, N. G. REICH, AND J. LESSLER (2020): "The incubation period
of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases:
Estimation and application," *Annals of internal medicine*, 172, 577–582.

LEWNARD, J. A. AND N. C. LO (2020): "Scientific and ethical basis for social-
distancing interventions against COVID-19," *The Lancet. Infectious diseases.*

Li, Q., X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. Lau, J. Y. Wong, et al. (2020): "Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia," *New England Journal of Medicine*.

Ministry of Health and Family Welfare, India (2020): "Micro Plan for Containing Local Transmission of Coronavirus Disease (COVID-19)," https://www.mohfw.gov.in/pdf/ModelMicroplanforcontainmentoflocaltransmissionofCOVID19.pdf, accessed on 04 June 2020.

Office of the Registrar General and Census Commissioner, India (2011): "STATE OF LITERACY," https://censusindia.gov.in/2011-prov-results/data_files/india/Final_PPT_2011_chapter6.pdf, accessed 04 June 2020.

Prather, K. A., C. C. Wang, and R. T. Schooley (2020): "Reducing transmission of SARS-CoV-2," *Science*.

Richardson, M., E. Dominowska, and R. Ragno (2007): "Predicting clicks: Estimating the click-through rate for new ads," in *Proceedings of the 16th international conference on World Wide Web*, 521–530.

Appendix A. Study Design and Details

A.1. **Data Collection and Sample Characteristics.**

A.1.1. *Survey of Present and Ex Village Council Members.* India has a well-established system of local government with village councils (small clusters of villages with a common council, known as Gram Panchayat) at its core. The councils are composed of elected representatives from each administrative unit (known as ward) in the cluster of villages. The term of a village council member lasts for 5 years. They are charged with ensuring basic amenities in the cluster of villages, implementation of government schemes, as well as other functions related to planning, and upkeep of welfare. Village councils derive legal authority from the 73rd amendment to the Indian Constitution, passed in 1992. The amendment also mandated a quota-based affirmative action system wherein a fraction of council member seats is reserved on a rotating basis for historically disadvantaged groups of Schedules Castes (SCs), Scheduled Tribes (STs) and women. The fraction of reserved seats for SCs and STs depends on each group's population share. [6] At least one-third of all council member seats have to be reserved for women and this proportion is maintained even within the seats reserved for SCs and STs. Moreover, in 2010, the West Bengal government decided to increase the share of seats reserved for women from one-third to half, to be operational from the 2013 elections.

Our sampling frame consisted of a publicly available directory of 44,312 present and ex village council members from 3,340 village councils across 19 districts in West Bengal. The directory included village council members from 2017, and 19.64% of our respondents were active village council members.

Postal code information was extracted where possible, for each council, and mapped onto the list of PIN codes that received treatment and control SMS. We were able to map 11,614 council members onto our PIN code database.

We started surveying the random sample of previous and current village council members two days after the broadcast, on May 8th, 2020, and continued until May 19th, 2020. The surveys were conducted by phone.

Enumerators were randomly assigned approximately 50 phone numbers to attempt per day, stratifying on treatment status. In order to obtain an appropriate number of observations each day for inference, numbers were attempted only once, and respondents that did not receive the call after a single attempt were tried again at the end

---

[6]The exact implementation varies by state.

of the day, and not contacted further at any point in the experiment if they did not pick up at the second attempt. If the respondent asked to be called again at a later time, we would do that. However, if in their second call they asked to be called at a later time again, we did not call a third time.

For the same reason, the survey was conducted with the person who received the call, provided they consented and were willing. We asked the respondent whether they currently held the post of a village council member, in order to track which of our respondents currently held council positions.

We collected our primary outcomes as follows:

(1) **Travel outside village**

Respondents were asked the following question:

"Including yesterday and the day before yesterday, where did you go outside your village? Think about the time when you went to the market which may have been outside of your village or when you went to a bank or a hospital outside your village or for any other reason whatsoever". Respondents were given the option of choosing if they went to another village, another town, another city, any combination of these 3 options, or they did not go out of their village in the last 2 days. If the respondent said they visited either another village, town, city, or any combination of these options, they were considered to have exited from the village in the last 2 days, for our purposes.

(2) **Interactions**

In order to measure interactions, we asked respondents the number of people they came within 2 arms distance of, excluding members of their household, within their village, in the last 2 days. Enumerators related the following script to the respondents, to help them think about people they met intentionally, as well as those they came into contact with unintentionally.

"Think of both the people you purposefully met with as well as close contact that was unintentional. Like, if you went to a shop, count both the shopkeeper as well as other people who were there to buy things and were within 2 arms distance of you. Think about anybody you ate a meal or had tea with, anybody you stopped to talk to (be it at work, at the tea stalls, medicine shops or on way to buy ration etc.), anybody you socialized with (spent time with, outside your family-watching TV together, played games with), anybody you worked closely with, anybody who's home you visited or anybody who visited your home."

Those who visited other villages, towns, or cities were also asked the number of people they came within 2 arms distance of, in each of these locations. Total interactions were measured by summing up the number of people respondents had interacted with, within their own village, as well as outside over the past 2 days.

(3) **Conversations about COVID-19**

We asked respondents the following questions to assess levels of conversation and learning about COVID-19 in the community:

(a) Whether they had given any information or advice to anyone in their village about COVID-19 over the phone, on WhatsApp, in person, or through any other means, over the last 2 days.

(b) The number of people in their village they gave information or advice to, if they did, over the last 2 days.

(c) Whether they had received any information or advice from anyone in their village about COVID-19 over the phone, on WhatsApp, in person, or through any other means, over the last 2 days.

(d) The number of people in their village that they received advice from, if they did, over the last 2 days.

To calculate the total number of conversations in Table 1, we summed the number of persons the respondent gave advice to, (point b. above) and the number they received advice from (point d. above). Conversations are measured over 2 days.

(4) **Frequency of washing hands**

To measure the level of hand washing, respondents were asked to think about a typical person in their village. If this person went out 10 times a day, they were asked the number of times upon returning home a typical person would wash their hands with soap, for over 20 seconds. Respondents were asked to think about a typical person in their community, in order to eliminate any demand effects that may arise as a result of asking them about their own hand washing behavior.

(5) **Mask usage**

In order to measure mask usage, respondents were asked the following:

(a) Out of 100 people in their village, how many are wearing masks.

(b) Whether they themselves wear a mask or anything else such as handkerchief to cover their face when they go out.

(c) Whether they agree or disagree with the statement "If you wear a mask, you can meet and interact with people as you like". This was asked to only a subset of respondents (randomized to appear in the digital survey 25 percent of the time).

(d) Whether they agree or disagree with the statement "If I wear a mask and go out in my location, I will not feel judged or people will not look at me differently". This was asked to only a subset of respondents (randomized to appear in the digital survey 25 percent of the time).

Items 4. and 5. were asked on a 5-point Likert scale, ranging from "Agree Strongly" to "Disagree Strongly", with a neutral midpoint. In table 2., we report on the respondent's own mask usage (point b. from the list above).

(6) **Knowledge of symptoms and precautions**

We asked respondents to list the symptoms of coronavirus, as well as the precautions that people could take to avoid contracting the virus. The list of symptoms consisted of correct, as well as incorrect symptoms, and the list of precautions consisted of correct as well as incorrect precautions.

- The correct symptoms included in the list were: cough, shortness of breath, sore throat, fever, loss of taste, loss of smell, diarrhea, body aches, headaches, rashes on one's feet or body, conjunctivitis, tiredness, and chest pain. We used the WHO list of symptoms to differentiate between correct and incorrect symptoms.

- The incorrect symptoms included in the list were: runny nose, vomiting, dizziness, itching, chills and swelling in legs and feet.

- The correct precautions included in the list were: not going to work, wearing a mask, washing hands, using hand-sanitizer, social distancing, wearing clean clothes, not spitting, covering one's mouth when sneezing or coughing, wearing gloves and not touching one's face.

- The incorrect precautions included in the list were: having herbs, drinking hot water, drinking chemical substances, drinking alcohol, taking antibiotics, taking anti-malarial medication and eating hot food.

Respondents were first asked to enumerate the symptoms of coronavirus, to the best of their ability, without the enumerator reading out any options from the list. We then read out items from the list that the respondent had not mentioned previously, and asked them if they believed it was a symptom of

the virus or not. This was recorded in a separate variable. The same process was followed for precautions.

We then calculated the total correct symptoms listed by the respondent, by summing the number of correct symptoms they were able to tell us, without priming. The total correct precautions were also calculated in the same manner. Total incorrect symptoms were calculated by summing the number of incorrect symptoms mentioned by the respondent, without priming. Total incorrect precautions were also calculated in the same manner.

We then created a "knowledge index" for each respondent, (reported in Table 2), using the following formula:

total number of correct symptoms - total number of incorrect symptoms + total number of correct precautions - total number of incorrect precautions

We collected demographic information and other variables as follows:

(1) **Demographic Variables**
    We collected basic demographic information on respondents including:
    (a) Gender
    (b) Age
    (c) District: The name of the administrative division within the state of West Bengal, in which they lived.
    (d) Subdistrict: The name of the administrative division within the district, in which they lived.
    (e) Village: The name of the village, within the subdistrict, in which they lived.
    (f) Gram Panchayat (Village Council)- The name of the village council under which their village was governed.
    (g) Location: Whether they live in a village, town or city. For our purposes, a city was defined as a district headquarter or bigger than that. A town is a subdistrict headquarter. Everything smaller than that is a village.
    (h) PIN Code: The Postal Index Number (PIN) of the village in which they live. We used this information to then merge with the list of PIN Codes that Jio had delivered the different messages to, in order to obtain the treatment status of the respondent.
    (i) Village council membership: Whether they were currently holding office as a village council member.

(j) Smartphone access: Whether they, or anyone in their household possessed a smartphone.

(k) Jio access: Whether they, or anyone in their household, possessed a Jio sim card.

(2) **Other Secondary data**

(a) Population: We asked respondents to estimate the population of their village.

(b) Probability of Travel: We asked respondents to estimate the number of people from their village, that exited the village, in the last 2 days. Specifically, we asked: "Think about your friends, relatives or other people in your village you know, between yesterday and the day before yesterday how many people went out of the village?". Based on the answer they provided to the above question, we asked them to estimate the number of people who exited to go to another village (a), the number that exited to go to another town (b), and the number who exited to go to another city (c), out of the total that exited in the last 2 days (x).

The proportion of the village that exited (y) was calculated by dividing the total number of people that exited the village (x), by the total village population estimated by the respondent in 1.

Probabilities of traveling from the village to another village, town or city was estimated as per the below:

(i) Probability of traveling from the village to another village = proportion of the village that exited (y)*the number that exited to go to another village(a) / the total number that exited (x)

(ii) Probability of traveling from the village to another town = proportion of the village that exited (y)*the number that exited to go to another town (b) / the total number that exited (x)

(iii) Probability of traveling from the village to another city = proportion of the village that exited (y)*the number that exited to go to another city (c) / the total number that exited (x)

(c) Distances: We asked respondents to think of the last village that they or someone they know had visited, and to estimate the distance from their current village to that village. We similarly asked them to estimate the distance to the last town and the last city that they or someone they know had visited. Questions on distances were asked to only a subset

of respondents (randomized to appear in the digital survey 25 percent of the time).

(d) Migration: Respondents were asked the following questions on persons who had recently returned back to their village from other places in the country:

(i) Whether any members of their household who lived and worked outside the village had come back home since the beginning of March, and how many such household members had returned. This was asked to a subset of respondents (randomized to appear in the digital survey 25 percent of the time).

(ii) Whether any other persons had visited/come to stay with them since the beginning of March, and how many such persons had done so. This included friends, relatives, new brides etc., who stayed in their home for a night or longer. This was asked to a subset of respondents (randomized to appear in the digital survey 25 percent of the time).

(iii) We also asked them to estimate the number of households in their village, that had any new persons coming to stay with them since the beginning of March.

(e) YouTube Views: All of our treatment videos were shared with respondents via a private YouTube link. We were able to capture the total number of views on each video, as well as the number of unique views, as per data provided by YouTube.

A.1.2. *Survey of Health Workers.* Accredited social health activists (ASHAs) are community level health workers that are generally responsible for serving the health needs of approximately 1000 residents of a single village. Instituted by the government of India's Ministry of Health and Family Welfare (MoHFW) as a part of the National Rural Health Mission (NRHM) in 2005, ASHAs are selected from amongst the female members of the village, and are required to live in the same village they serve in. Their primary tasks include overseeing immunization of children, catering to the needs of pregnant and lactating women by providing basic pre and post natal care, spreading awareness on good health practices and sanitation, treating basic illnesses, keeping demographic records, and monitoring those suffering from non-communicable diseases. ASHA workers are meant to be the link between the community and the

public health system. At the onset of the COVID-19 pandemic, ASHAs were repurposed to be the first point of contact for COVID-related concerns in villages (Ministry of Health and Family Welfare, India, 2020). They were required to keep track of the incidence of COVID related symptoms by conducting house visits, referring suspected cases to higher hospitals, and keeping a check on the number of migrants returning from other parts of the country, to the village.

We built a database of health worker contact information across 19 districts in West Bengal using our village council directory. We randomly sampled 5,253 individuals from this directory, and contacted them to obtain the contact details of the health worker in their village. This list of respondents was separate from the 11,614 members that were sampled for the outcomes survey.

Enumerators were randomly assigned 15 phone numbers to attempt each day. For making initial contact, health workers that did not receive the call were tried again the following day, after which they were dropped. For follow up surveys, respondents were tried again for 2 more days, after they did not receive their follow-up call, before being dropped from further data collection.

On contacting a health worker for the first time, we received information on demographics, the area and number of households they were in charge of, the number of households they visited, the number of recent migrants that had returned to their areas, and the number of cases of fever, cough, respiratory illness, anosmia and rashes around the feet in their area for the previous 3 days. We then obtained permission to conduct follow up surveys with them at 3-day intervals. For the health workers that agreed to follow-ups, we called them 3 days later to obtain data on cases of fever, cough, respiratory illness, anosmia and rashes on the feet again.

A.2. **Data Quality Assurance Procedures.**

- Data Checks: Incoming data was scrutinized on a daily basis for enumerator productivity, respondent refusal rates, successful call rates, total survey duration, section wise survey duration, the rate of responses that were coded as "Do Not Know", and outliers greater than 2 standard deviations from the mean. The above items were observed at an enumerator level as well as at an aggregate level. Unreasonable outliers were cross verified with the enumerator who had conducted the survey.
- Accompaniments: 10 percent of outcome surveys, and 30 percent of health worker surveys were also observed by senior field staff. Senior enumerators

were on the same phone call with the respondent, along with the enumerator conducting the survey. The senior field staff members were required to record responses, and submit a separate survey form, for all calls they observed. Responses from the senior enumerator's form and the enumerator's form were cross checked, to ensure minimal discrepancies across all members of the survey team.

A.3. **Randomization and Intervention.** The list of candidate PIN codes comes from two sources: a dataset of PIN codes from data.gov.in dated May 2019, and an unofficial repository of PIN codes mirrored from the "PIN code Search" tool on the official India Post website. We classify each PIN code as either a city PIN code or a non-city PIN code.

PIN codes do not have precise geographical boundaries, so we use a geocoding procedure to identify city pin codes. First, we query the Google Maps API to locate the centroid of each PIN code as depicted on Google Maps. Then, we overlay these centroids onto a shapefile of West Bengal and define city PIN codes as those whose centroids fall inside a polygon with a population of at least 100,000. We exclude 50 PIN codes that we could not classify as city or non-city using this procedure, resulting in a sample of 1214 PIN codes.

We assign treatment at the PIN code level using stratified randomization, stratifying at the district level outside cities, and at the district and city level inside cities. We present basic demographics for our treatment and control units as well as results from balance tests in Appendix Table S1.

Treatment videos use the following script:

> "*This is Abhijit Banerjee. I am stuck very far away but my heart is with you in West Bengal. I am sending you a set of suggestions and requests from the government of West Bengal and myself.*
>
> *West Bengal, like the rest of the world, is fighting the spread of an unseen and unknown enemy. It won't be won in a few days, it's going to take some months at least. On the other hand, while it is very contagious, most people get better in a matter of days. We need to be careful, not petrified.*
>
> *Please forward this message to others on the phone via WhatsApp or text about it. Do not physically meet others to share it, you may pass the virus or get sick yourself.*

- (Only in Hygiene-focused): *To prevent getting the virus, wash your hands frequently and well with soap and water, especially after going out of the house. Without soap the virus will still stay on your hand. Count 30 while washing. Avoid touching your face as far as possible.*
- (Only in Social Distancing-focused): *To prevent getting the virus avoid going to any crowded places—markets, hospitals, meetings, as far possible. Try to stay at least 2 meters from the nearest person when you are outside the house, even when you are doing allowed activities such as groceries, bank, work or even your next door neighbors.*

Also please do not spit. If you have to cough or sneeze don't cough or sneeze into your hands. Use your elbow or your clothes or a handkerchief or a towel, but remember to boil them after use before hanging them to dry

- (Only in Internality + Externality-focused): *By doing so you protect the community and especially the elderly and otherwise vulnerable people in your neighborhood.*
- (Only in No-Ostracism): *If you or your family members have a cough and fever, it does not have to be corona. No one has the right to ostracize you for having the disease. If that happens let the local ASHA know.*
- (Ostracism not mentioned): *If you or your family members have a cough and fever, it does not have to be corona.*

And/However if you or your family members have a cough and fever, just to be safe, please report it to the ASHA or on West Bengal government's Shondhane app so that the government can track the spread of the disease and try to control it. Be well, don't lose hope, and please accept my greetings. "

Appendix B. Tables and Figures

Table B1. Balance Table

		(1) Control		(2) Treatment	T-test P-value
Variable	N	Mean/SE	N	Mean/SE	(1)-(2)
Respondent age	242	44.996 (0.674)	1747	44.858 (0.253)	0.849
Respondent gender	242	0.202 (0.026)	1747	0.170 (0.009)	0.212
Respondent or at least someone in family has a Jio connection	230	0.543 (0.033)	1653	0.580 (0.012)	0.292
Has a smartphone?	242	0.740 (0.028)	1747	0.726 (0.011)	0.650
Are you currently a gram panchayat member?	214	0.196 (0.027)	1573	0.196 (0.010)	0.995

Notes: The value displayed for t-tests are p-values.

TABLE B2. Background Conversations Survey

	Mean	SD	N
Outgoing phone calls made	26.432	29.443	405
Social Distancing messages received	20.248	19.030	408
Mask usage messages received	17.176	17.135	408
Messages received on washing hands	16.863	16.359	408

Notes: Respondents were asked the total number of outgoing phone calls they had made, and the number of messages on social distancing, washing hands, and using masks they had heard or received in the last 2 days.

TABLE B3. Phone Survey Regressions

	In the last two days (Winsorized at 97.5%-ile)		Belief		
	(1) Number of interactions with people within 2 arms length	(2) Number of conversations in person /online /mobile about COVID-19	(3) Belief about mask usage in community	(4) About Symptomatic Rate	(5) About Asymptomatic Rate
Panel A: Pooled					
TREATMENT	-1.091 (1.458) [0.455]	-3.056 (1.765) [0.084]	0.031 (0.015) [0.037]	0.005 (0.020) [0.783]	-0.002 (0.003) [0.377]
Observations	1,875	1,881	1,872	1,580	1,686
district FE	✓	✓	✓	✓	✓
survey day FE	✓	✓	✓	✓	✓
jio FE	✓	✓	✓	✓	✓
Control Mean	11.891	14.104	0.772	0.114	0.011
Panel B: Jio					
TREATMENT	1.264 (1.873) [0.500]	-1.847 (2.211) [0.404]	0.040 (0.021) [0.055]	0.054 (0.018) [0.003]	-0.002 (0.004) [0.519]
Observations	1,076	1,082	1,075	928	987
district FE	✓	✓	✓	✓	✓
survey day FE	✓	✓	✓	✓	✓
Control Mean	10.640	14.112	0.762	0.072	0.010
Panel C: Non-Jio					
TREATMENT	-4.048 (2.573) [0.117]	-4.210 (2.936) [0.153]	0.020 (0.023) [0.390]	-0.055 (0.039) [0.162]	-0.002 (0.004) [0.614]
Observations	799	799	797	652	699
district FE	✓	✓	✓	✓	✓
survey day FE	✓	✓	✓	✓	✓
Control Mean	13.381	14.095	0.784	0.166	0.013

Notes: Columns 1 and 2 are winsorized at 97.5%, as compared to 95% in table of main paper. All columns control for district and survey day fixed effects and in panel A we have Jio access fixed effects as well. Respondent level controls also include age, gender and smartphone access. Standard errors (clustered at PIN code level) are reported in parentheses and p-values are reported in brackets.

TABLE B4. ASHA Case Counts (Reports from May 7th to 8th)

VARIABLES	(1) Number of Fever Cases	(2) Number of Respiratory Cases
TREATMENT	0.272	0.245
	(0.111)	(0.101)
	[0.015]	[0.016]
Observations	406	406
district FE	✓	✓
total rounds FE	✓	✓
smartphone FE	✓	✓
Control Mean	0.208	0.189

Notes: The case counts for both columns look at the window from 7th to 8th May. We account for district, total rounds and smartphone fixed effects. We also control for total households under the ASHA. Standard errors (clustered at PIN code level) are reported in parentheses and p-values are reported in brackets.

TABLE B5. ASHA Case Counts (Reports from May 7th to 9th)

VARIABLES	(1) Number of Fever Cases	(2) Number of Respiratory Cases
TREATMENT	0.300	0.196
	(0.091)	(0.088)
	[0.001]	[0.027]
Observations	602	602
district FE	✓	✓
total rounds FE	✓	✓
smartphone FE	✓	✓
Control Mean	0.229	0.214

Notes: The case counts for both columns look at the window from 7th to 9th May. We account for district, total rounds and smartphone fixed effects. We also control for total households under the ASHA. Standard errors (clustered at PIN code level) are reported in parentheses and p-values are reported in brackets.

FIGURE B1. Effects over Time (Pooled Sample)

- For all plots, we have district and Jio access fixed effects, as well as controls for age, gender and smartphone access. The dot represents the point estimate and whiskers represent 90% confidence intervals. Standard errors are clustered at the PIN code level.
- The x-axis represents number of days post the broadcast message was sent.
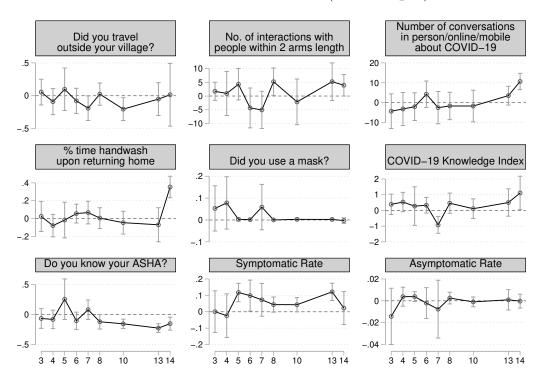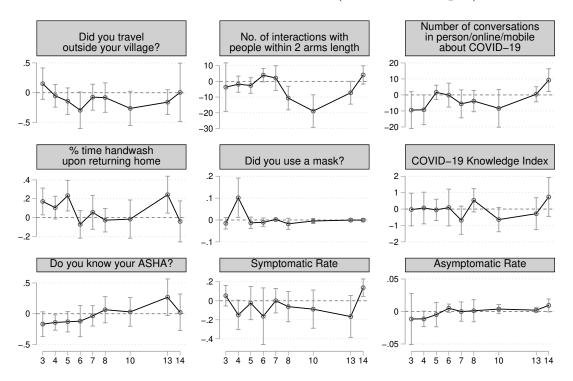
FIGURE B2. Effects over Time (Jio Sample)

- For all plots, we have district fixed effects, as well as controls for age, gender and smartphone access. We keep only the sample which had access to Jio connection. The dot represents the point estimate and whiskers represent 90% confidence intervals. Standard errors are clustered at the PIN code level.
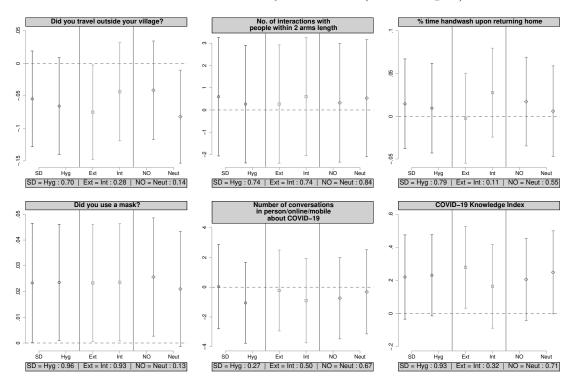- The x-axis represents number of days post the broadcast message was sent.

FIGURE B3. Effects over Time (Non-Jio Sample)

- For all plots, we have district fixed effects, as well as controls for age, gender and smartphone access. We keep the sample which did not have access to a Jio connection. The dot represents the point estimate and whiskers represent 90% confidence intervals. Standard errors are clustered at the PIN code level.
- The x-axis represents number of days post the broadcast message was sent.

FIGURE B4. Effects by Content (Jio Sample)

- For all plots, we have district and survey day fixed effects, as well as controls for age, gender and smartphone access. We keep only the sample which had access to Jio connection. The dot represents the point estimate and whiskers represent 90% confidence intervals. Standard errors are clustered at the PIN code level. Sd - "Social Distancing", Hyg - "Hygiene", Ext - "Externality", Int - "Internality", NO - "No Ostracism", Neut - "Neutral". Externality here implies both "Internality and Externality" was in the message.
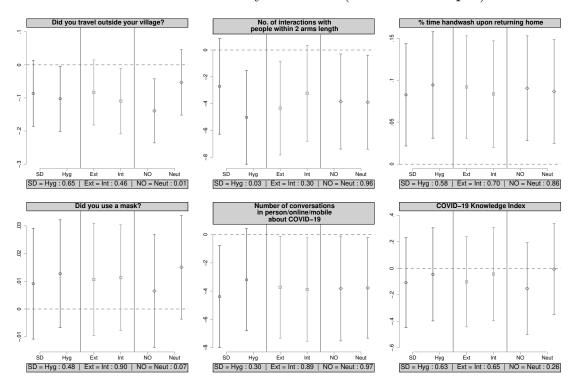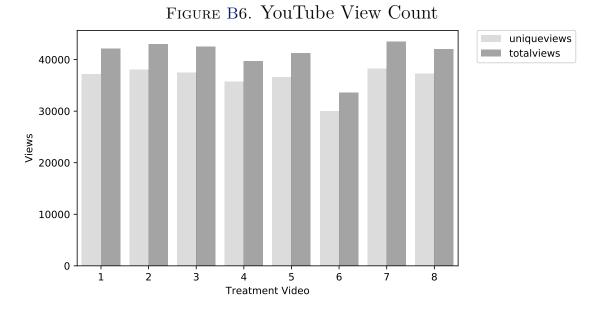
FIGURE B5. Effects by Content (Non Jio Sample)

- For all plots, we have district and survey day fixed effects, as well as controls for age, gender and smartphone access. We keep only the sample which did not have access to a Jio connection. The dot represents the point estimate and whiskers represent 90% confidence intervals. Standard errors are clustered at the PIN code level. Sd - "Social Distancing", Hyg -"Hygiene", Ext - "Externality", Int - "Internality", NO - "No Ostracism", Neut - "Neutral". Externality here implies both "Internality and Externality" was in the message.

FIGURE B6. YouTube View Count

- For each of the 8 treatment videos we provide both "Unique viewers" and "Total views".