Correcting Perceived Social Distancing Norms to Combat COVID-19[†]

James Allen IV[1,2,3], Arlete Mahumane[4], James Riddell IV[5], Tanya Rosenblat[1,6], Dean Yang[1,2,3], and Hang Yu[7,8]

[1]Department of Economics, University of Michigan
[2]Ford School of Public Policy, University of Michigan
[3]Population Studies Center, University of Michigan
[4]Beira Operational Research Center, National Institute of Health, Mozambique
[5]Division of Infectious Diseases, University of Michigan Medical School
[6]School of Information, University of Michigan
[7]National School of Development, Peking University
[8]Institute of South-South Cooperation and Development, Peking University

Version: March 26, 2021

Abstract

Can informing people of high rates of community support for social distancing encourage them to do more of it? Our Mozambican study population underestimated the rate of community support for social distancing, believing support to be only 69%, while the true share was 98%. In theory, informing people of high rates of community support has ambiguous effects on social distancing, depending on whether a perceived-infectiousness effect dominates a free-riding effect. We randomly assigned a "social norm correction" treatment, informing people of true high rates of community support for social distancing. We examine an improved measure of social distancing combining detailed self-reports with reports on the respondent by others in the community. The treatment increases social distancing where COVID-19 case loads are high (where the perceived-infectiousness effect dominates), but decreases it where case loads are low (where free-riding dominates). Separately, randomized local-leader endorsements of social distancing are ineffective. As COVID-19 case loads continue to rise, interventions such as the "social norm correction" treatment should show increased effectiveness at promoting social distancing.

JEL Classification: I12, D91, O12

Keywords: COVID-19, Social Distancing, Health Behavior, Mozambique

[†]Contacts: alleniv@umich.edu; deanyang@umich.edu. Faustino Lessitala provided top-notch leadership and field management. Patricia Freitag, Ryan McWay, and Maggie Barnard provided excellent research assistance. Julie Esch, Laura Kaminski, and Lauren Tingwall's grant management was world-class. This work is supported by the Abdul Latif Jameel Poverty Action Lab (J-PAL) Innovation in Government Initiative through a grant from The Effective Altruism Global Health and Development Fund (grant number IGI-1366), the UK Foreign, Commonwealth & Development Office awarded through Innovation for Poverty Action (IPA) Peace & Recovery Program (grant number MIT0019-X9), the Michigan Institute for Teaching and Research in Economics (MITRE) Ulmer Fund (grant number G024289), and the National Institute on Aging of the National Institutes of Health (award number T32AG000221). Our protocols were reviewed and approved by Institutional Review Boards (IRBs) at the University of Michigan (Health Sciences and Social and Behavioral Sciences IRB, approval number HUM00113011) and the Mozambique Ministry of Health National Committee on Bioethics for Health (CNBS reference number 302/CNBS/20). The study was submitted to the AEA RCT Registry on May 26, 2020, registration ID number AEARCTR-0005862: 10.1257/rct.5862-1.0. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the aforementioned institutions.

1 Introduction

Social distancing is one of the most important public health recommendations for reducing the spread of COVID-19 (CDC, 2020). Social norms in support of social distancing have changed rapidly during the pandemic (Janzwood, 2020; Reicher and Drury, 2021; Habersaat and Scheel, 2020). Because of this rapid change in norms, people often underestimate the level of support for social distancing in their communities. In our Mozambican sample, 98% thought that people should be social distancing, but estimated that only 69% of others in the community express similar support. This gap between perceived and actual support suggests a public health messaging strategy: simply inform people of high rates of community support for social distancing. What impact would such messaging have on social distancing behavior?

We implemented a randomized controlled trial testing the impact of informing people about high local support for social distancing. In practice, this treatment updated beliefs upwards or confirmed beliefs about high rates of support for social distancing. We analyze the impacts of the treatment through the lens of a simple theoretical model. Individuals weigh the benefits (avoiding infection) against the effort cost of social distancing. They have imperfect information on the level of community support for social distancing and on the infectiousness of the disease, and beliefs on both fronts adjust to be consistent with one another. Informing people that a higher proportion of their community supports social distancing has two opposing effects. When the local infection rate is low, *free-riding* effects dominate, and the treatment leads to *less* social distancing. By contrast, if local infection rates are high enough, a *perceived-infectiousness* effect dominates, leading to *more* social distancing.

Abiding by COVID-19 health protocols, we conducted all treatments and surveys by phone in our sample of 2,117 Mozambican households. The perceived social norm is a respondent's estimated share in the community supporting social distancing. The true social norm is the community average of respondents' stated support for social distancing. Alongside the social norm correction treatment, we also randomly assigned a "leader endorsement" treatment (an endorsement of social distancing by a community opinion leader).

We construct a novel measure of social distancing, improving on prior research on two fronts. Most prior studies ask respondents to self-report about general social distancing compliance. When we do so, 95% claim to be observing government social distancing recommendations. Our first improvement is to also ask respondents to self-report several specific social distancing actions; this alone leads the social distancing rate to drop to 36%. Second, we ask *others* in the community to report on the respondent's social distancing. This helps reduce biases in self-reported outcomes due to experimenter demand effects (Orne, 1962; Rosenthal, 1966; Zizzo, 2010; De Quidt et al., 2018).[1] We are aware of no prior study

[1] Jakubowski et al. (2021) find that self-reported mask wearing is overstated relative to measures based

that makes use of other-reports on a respondent's social distancing behavior. Other-reports also make a major difference, leading the share observing social distancing to fall further to just 8%. (See Figure 1 and Section 3.3 below for more detail.) This is the first finding of the paper: improved measurement leads the social distancing rate to fall by an order of magnitude, from 95% to 8%.

The average effect of the social norm correction treatment in the full sample is small and not statistically significantly different from zero. As theory predicts, there is substantial treatment effect heterogeneity: the treatment effect is statistically significantly more positive when local COVID-19 cases (per 100,000 population) are higher. In districts with few cases, the treatment effect is negative. In the district with the most COVID-19 cases, the treatment increases social distancing by 9.3 percentage points (statistically significant at the 5% level), a 75% increase over that district's control-group mean.

This pattern is consistent with the theoretical prediction that as infection rates rise, the perceived-infectiousness effect should increasingly dominate the free-riding effect of the social norm correction treatment, leading the treatment effect to become more positive. We also test a further implication of the model: expectations of future infection rates should show similar treatment effect heterogeneity. Empirical analyses confirm this prediction, providing additional support for the theoretical model.

The leader endorsement treatment has a very small effect on social distancing that is not statistically significantly different from zero. We also find no treatment effect heterogeneity for this treatment with respect to COVID-19 cases.

This paper contributes to research on interventions promoting social distancing to combat COVID-19. Related work studies public health messaging on self-reported social distancing behavior or intentions, randomizing messages framed as altruistic or selfish (Sasaki et al., 2020), deontological or consequentialist (Bos et al., 2020), emotional or rational (Capraro and Barcelo, 2020), individual- or group-oriented (Lunn et al., 2020), and self-interested or prosocial (Jordan et al., 2020; Banker and Park, 2020). These studies typically find that any kind of public health messaging increases self-reported social distancing, but differ on whether self- versus others-oriented language is more effective. No prior study has tested the impact of providing information on community support of social distancing.

This paper also contributes to understanding the impact of providing information about social norms in one's reference group (Benabou and Tirole, 2011; Bicchieri and Dimant, 2019). Banerjee et al. (2019) find that informing Nigerian young adults of their peers' attitudes on healthy sexual relationships did not change respondents' own attitudes. Yu (2020) and Yang et al. (2021) find (in an overlapping Mozambican sample) that correcting overestimates of stigmatizing attitudes promoted HIV testing.[2] Van Bavel et al. (2020)

on observations of others.

[2] Norm-based interventions have also been shown to change energy consumption (Schultz et al., 2007) and female labor force participation (Bursztyn et al., 2018).

recommend social-norm-based interventions to combat the COVID-19 pandemic, suggesting they are most effective when specific to an individual's social identity (Centola, 2011) or promoted by those central in a social network (Kim et al., 2015). Some studies measure respondent perceptions of social distancing norms (Papanastasiou et al., 2020; Xie et al., 2020; Arroyos-Calvera et al., 2021). Alsan et al. (2020) find that correcting social norms about COVID-19 mask-wearing leads to more correct norm perceptions.

Our emphasis on an interplay of free-riding and perceived-infectiousness effects is novel, but each effect has separately been the subject of prior research.[3] Free-riding has been studied in the context of vaccination decisions (Hershey et al., 1994; Lau et al., 2019) and COVID-19 social distancing (Cato et al., 2020; Paakkari and Okan, 2020).[4] Perceived COVID-19 infection risk (e.g., due to vaccine anticipation, Andersson et al. (2021)) has been shown to raise social distancing intentions. Our theoretical approach is also distinctive in that we do not assume that individuals value adherence to social norms *per se* (Bernheim, 1994; Benabou and Tirole, 2011; Krupka and Weber, 2013).

Our leader endorsement treatment relates to research on opinion leaders and health behavior change (Valente and Pumpuang (2007) provide a review). Opinion leaders have been recruited to promote healthy behaviors related to HIV/AIDS (Kelly et al., 1992), and are effective at diffusing public health information (Banerjee et al., 2019). Banerjee et al. (2020) find that endorsement videos by celebrity Nobel Laureate Abhijit Banerjee raise COVID-19 symptom reporting and have positive spillovers in West Bengal. In a different context, we find that community leaders' social distancing endorsements are ineffective.

2 Theory

We consider a community where people have random pairwise meetings. People believe that a share x of the population supports social distancing and that the probability of becoming infected from unprotected meetings is α. We assume that people treat x as given, but that they infer the infectiousness α from the current infection rate R in the community which they can observe (we describe this inference below). The true infectiousness of the disease is $\hat{\alpha}$.

Importantly, people in the community have *miscalibrated beliefs*: the true share of the population supporting social distancing is $\hat{x}$ (we are particularly interested in the case $\hat{x} > x$). People infer the true infectiousness $\hat{\alpha}$ of the disease only if they are correctly calibrated ($x = \hat{x}$).

[3] The mechanisms we highlight also differ from prior work that emphasizes norm conformity (Cialdini and Goldstein, 2004; Wood, 2000).

[4] Relatedly, positive health spillovers in the community have been shown to reduce demand for health goods (e.g., Dupas (2014) for anti-malarial bednets).

Individual Effort. A supporter engages in preventative effort e and assumes that other supporters choose effort e^* (in equilibrium we have $e = e^*$). Non-supporters choose effort $e = 0$.

When an agent supporting social distancing meets another person she escapes exposure with probability:

$$\begin{aligned} A(e, e_{other}) &= \sqrt{e + e_{other}} \\ &= \begin{cases} \sqrt{e + e^*} & \text{if other person is supporter} \\ \sqrt{e} & \text{if other person is non-supporter} \end{cases} \end{aligned} \tag{1}$$

Hence, the marginal benefit of effort decreases both with own effort e as well as the other person's effort e^*.

The expected probability of escaping exposure is therefore:

$$\overline{A}(e, e^*) = (1 - x)\sqrt{e} + x\sqrt{e + e^*} \tag{2}$$

An agent becomes exposed with probability $1 - \overline{A}(e, e_{other})$. If exposed she gets infected with probability α and suffers disutility $-C$ from infection. If she is not exposed then she does not get infected. Her baseline utility from no infection equals $\overline{U}$. The cost of preventative effort is e. Hence, her total utility equals:

$$\overline{U} - \alpha(1 - \overline{A}(e, e_{other}))C - e \tag{3}$$

The agent chooses e to maximize her utility, giving us the following first-order condition:

$$\frac{\alpha C}{2\sqrt{e}} \left[1 - x(1 - \frac{1}{\sqrt{1 + \frac{e^*}{e}}}) \right] = 1 \tag{4}$$

In equilibrium it has to be the case that the population effort e^* equals e. Hence, we can characterize equilibrium effort as:

$$e = \left(\frac{\alpha C}{2} \left[1 - x(1 - \frac{1}{\sqrt{2}}) \right] \right)^2 \tag{5}$$

This demonstrates the basic *free-riding effect*: increasing the share x of supporters *decreases* effort because the marginal utility from own effort decreases. In addition, effort increases if the disease is more infectious (higher α) and if getting sick is more costly (higher C).

Infection Rate. People can observe the current infection rate in the community. Infections come from two sources: non-supporters become sick at rate $\alpha\,(1 - x\sqrt{e})$ while supporters become sick at rate $\alpha(1 - \overline{A}(e, e))$. Hence, people in the community assume that the current infection rate is generated by the following process:

$$\begin{aligned} R &= \alpha\left[\underbrace{(1-x)(1-x\sqrt{e})}_{\text{non-supporters}} + \underbrace{x\left(1-\sqrt{e}(1+(\sqrt{2}-1)x)\right)}_{\text{supporters}}\right] \\ &= \alpha\left[1-\sqrt{e}\,\underbrace{2x\left(1-x\left(1-\frac{1}{\sqrt{2}}\right)\right)}_{=G(x)}\right] \end{aligned} \tag{6}$$

However, the true process determining current infections is actually:

$$R = \hat{\alpha}\left[1 - \sqrt{e}G(\hat{x})\right] \tag{7}$$

In other words, the true infection process is driven by the same social distancing effort of supporters but different infectiousness $\hat{\alpha}$ and different $\hat{x}$.

2.1 Basic Equilibrium

Supporters initially assume that the disease has low infectiousness and they adjust their estimate of α upwards until the current infection rate R stabilizes.

Proposition 1 *In equilibrium, effort level e, the current infection rate R, and the assumed infectiousness α satisfy Equations 5, 6 and 7. Moreover, $\hat{\alpha} > \alpha$ if $\hat{x} > x$.*

In equilibrium, both the assumed infection process (captured in Equation 6) and the real infection rate (captured in Equation 7) have to produce the observed infection rate R. For the second part, note that $G(x)$ is increasing in $x \in [0, 1]$: hence, $\hat{x} > x$ implies $\hat{\alpha} > \alpha$ to generate the same infection rate R.

2.2 Treatment Impact

We now consider the impact of our treatment where people are informed that the share of people who support social distancing is really $\hat{x} > x$.

Proposition 1 implies that if supporters are informed that the true share of the population supporting social distancing is $\hat{x} > x$ then they have to infer that the disease is more infectious than they initially assumed (because their estimate of the infectiousness of the disease immediately jumps from α to the true $\hat{\alpha}$). This is the *perceived-infectiousness effect.*

Supporters of social distancing will adjust their effort level to a new level $\hat{e}$, but there are two countervailing effects:

1. Holding assumed infectiousness α constant, the free-riding effect *decreases* effort.

2. The perceived-infectiousness effect *increases* effort, because the agent now believes the disease is more infectious than initially thought (perceived α increases), increasing the gain from social distancing.

Intuitively, the perceived-infectiousness effect varies monotonically with R: when infections are low, supporters' effort is low, and both supporters and non-supporters get infected at similar rates. Hence, agents revise the estimate of infectiousness α only slightly upwards in response to the treatment. On the other hand, when infections are high, supporters' effort is high and the upward revision will be larger.

The following theorem makes this intuition precise. Instead of doing comparative statics on R (which is determined in equilibrium) we state the comparative statics results in terms of the infectiousness $\hat{\alpha}$ (for given x and $\hat{x}$). Note that R increases with $\hat{\alpha}$.

Theorem 1 *Assume an agent is informed that a share $\tilde{x} > x$ of the population supports social distancing. Then there is a threshold $\hat{\alpha}^*$ such that for any $\hat{\alpha} < \hat{\alpha}^*$ the free-riding effect dominates and equilibrium effort decreases, and for $\hat{\alpha} > \hat{\alpha}^*$ the perceived-infectiousness effect dominates and the equilibrium effort increases.*

The proof is in Appendix A.

To be clear, in our theoretical approach, there is no gain from social norm adherence *per se*, for example to signal conformity (Benabou and Tirole, 2011), avoid norm-violation social sanctions (Bernheim, 1994), or avoid direct utility losses from norm violation (Bicchieri, 2005; Krupka and Weber, 2013). Adding such motives would increase the likelihood that the social norm correction treatment increases social distancing, but would not by itself generate the heterogeneous treatment effects (with respect to current infection rates) that are central to our analysis.[5]

The interplay between free-riding and perceived-infectiousness effects also yields analogous predictions about a central belief about COVID-19: the future infection rate. In the endline survey, we ask respondents to estimate this rate. The expected future rate differs from the current infection rate R, because this study occurs at a point when infection rates are clearly evolving. The social norm correction treatment changes respondent beliefs about social distancing support and about infectiousness, and therefore should change expected future infection rates.

Recall that non-supporters are always infected with higher probability than supporters. The higher the infectiousness parameter $\hat{\alpha}$, the higher should be future infection rates for both groups.

[5] Bicchieri and Dimant (2019) discuss how norm interventions can backfire when descriptive and injunctive norms differ from each other. Our model can be seen as one way to provide micro-foundations for such a mechanism in the emerging infectious disease context.

When $\hat{\alpha}$ is currently small, the perceived-infectiousness effect is small. Simultaneously, the treatment corrects beliefs about the share of social-distancing supporters upwards, which should *reduce* estimates of future infection rates because supporters have lower infection rates. Thus, the expected future infection rate *decreases* when $\hat{\alpha}$ is currently small.

In contrast, when $\hat{\alpha}$ is currently large, the treatment leads to a large increase in perceived infectiousness, implying that the disease will infect higher shares of both supporters and non-supporters. This will tend to *increase* expected future infection rates.

To summarize, the social norm correction treatment effect on the expected future infection rate should show heterogeneity similar to that described in Theorem 1. The treatment effect on the expected future infection rate is strictly negative if the current local infection rate (R) (which moves monotonically with $\hat{\alpha}$) is small enough. The treatment effect on the expected future infection rate increases with the current infection rate, and can become positive if current infection rates are sufficiently high.

We now turn to our empirical analyses. We test the model's predictions regarding heterogeneity in the social norm correction treatment effect with respect to the current local infection rate.

3 Sample and Data

3.1 Context

The Mozambican government declared a State of Emergency due to the COVID-19 pandemic on March 31, 2020 (Republic of Mozambique, 3/31/2020) and shortly after recommended social distancing (at least 1.5 meters) and required it at public and private institutions and gatherings. The government also suspended schools, required masks at funerals and markets, banned gatherings of 20 or more, and closed bars, cinemas and gymnasiums (Republic of Mozambique, 4/1/2020). The government stopped short of implementing a full economic "lockdown" due to its economic costs (Siuta and Sambo, 2020; Jones et al., 2020). On August 5, 2020, the government renewed the State of Emergency (Republic of Mozambique, 8/5/2020), called for improved mask-wearing, and announced a schedule for loosening restrictions (Nyusi, 8/5/2020). On September 7, 2020, the government downgraded its State of Emergency to a State of Public Calamity, further loosening some restrictions including resuming religious events and ceremonies at 50% capacity (Nyusi, 9/5/2020; U.S Embassy in Mozambique). Throughout this period, the government's social distancing recommendation remained constant.

3.2 Data

We collected survey data in three rounds between July 10 and November 18, 2020. Following COVID-19 research protocols, we conducted all surveys over the phone. Respondents were from households with phones in the sample of a separate study of a health program (Yang et al., 2021) in central Mozambique.[6] We surveyed one adult respondent per study household. Appendix B provides details on the study communities.

Appendix Figure A.2 depicts the study timeline below a rolling average of new Mozambican COVID-19 cases. We piloted surveys in Round 1. Immediately before the Round 2 survey, we randomly assigned households to treatments and submitted our pre-analysis plan (PAP) to the AEA RCT Registry. The Round 2 survey served as a baseline, and was immediately followed by treatments. Round 3 was our endline survey. There was a minimum of 3.0 weeks and average of 6.3 weeks between Rounds 2 and 3 surveys for any given respondent. While the Round 1 survey occurred when new COVID-19 cases remained relatively steady, both the Round 2 and Round 3 surveys occurred when cases were rising rapidly.

The Round 3 sample size is 2,117 respondents, which followed a sample size of 2,226 in Round 2 and 2,412 in Round 1. The retention rate between Round 2 (baseline) and Round 3 (endline) is 95.1% overall, at least 94.4% in each of the seven districts surveyed, and balanced across treatment conditions.

Table 1 presents summary statistics. 99% of respondents support social distancing, but respondents underestimate the share of others in their community expressing such support, on average estimating 69% in the Round 1 survey and 80% in Round 2.

3.3 Primary Outcome

The primary outcome is an indicator that the respondent practiced social distancing. We completely pre-specified its definition prior to Round 2. It is constructed from self-reports of social distancing as well as others' reports of the respondent's social distancing. Incorporation of others' reports on social distancing yields an improvement over sole reliance on self-reported social distancing, which is subject to experimenter demand effects (Orne, 1962; Rosenthal, 1966; Zizzo, 2010; De Quidt et al., 2018). The primary outcome is equal to one if the respondent is practicing social distancing according to both self-reports and other-reports, and zero otherwise.

Respondents are social distancing according to their self-report if both of the following are true: 1) they answer "yes" to "In the past 14 days, have you observed the government's recommendations on social distancing?", and 2) they report doing at least seven out of eight "social distancing actions" (listed below) in the past seven days (higher than the sample

[6] The AEA RCT Registry record for Yang et al. (2021) is here: https://doi.org/10.1257/rct.3990-5.1

median number, six).

Social Distancing Actions: Is this something your household has been doing for the last seven days? (Answers indicating social distancing in parentheses. Summary statistics presented in Appendix F.)

1. Shop in crowded areas like informal markets (No)
2. Gather with several friends (No)
3. Help the elderly avoid close contact with other people, including children (Yes)
4. If show symptoms of coronavirus, immediately inform my household and avoid people (Yes)
5. Drink alcohol in bars (No)
6. Wear a face mask if showing symptoms of coronavirus (Yes)
7. Instead of meeting in person, call on the phone or send text message (Yes)
8. Allow children to build immunity by playing with children from other households (No)

To collect others' reports on a respondent's social distancing, study participants were asked about their social interactions with ten other study participants in their community. These ten other study participants were identified first from prior data (from Yang et al. (2021)) on social network contacts within our sample, and then based on closest geographic proximity. Additionally, community leaders identified and surveyed for the leader endorsement intervention were also asked about their social interactions with all study participants in their communities. Each respondent household was known at baseline by 0.98 community leaders and 3.21 neighboring survey respondents on average. Other-reports were collected at baseline and endline.

In collecting other-reports, we first asked others whether they had seen anyone from the respondent household in the last 14 days. If so, we then asked: 1) Did he/she come closer than 1.5 meters to you or others not of his/her household at any point in the last 14 days?; 2) Did he/she shake hands, try to shake hands, or touch you or others not of his/her household in the last 14 days?; and 3) In general, did he/she appear to be observing the government's recommendations on social distancing (avoid large gatherings and keep at least 1.5 meters distance from people not of his/her household)? Respondents are considered to be social distancing according to others if all others responded "no", "no", and "yes" (respectively) to these three questions, reported having not seen the respondent in the past 14 days, or reported not knowing the respondent.[7]

[7] At baseline, 90.55% of respondent households were known by at least one other respondent or community leader.

Figure 1 displays how we use these questions to construct the social distancing outcome. 95% of respondents say "yes" to the self-reported question on general social distancing compliance. When we then consider whether respondents self-report doing at least seven out of the eight specific social distancing actions, this leads the social distancing rate to fall to 36%. Finally, we incorporate others' reports, leading the social distancing rate to drop further to 8%. Incorporating other types of information into the social distancing measure – using self-reports of more specific social distancing behaviors as well as other-reports – leads to substantially lower social distancing rates.

4 Research Design

4.1 Treatments

We implement a randomized controlled trial estimating impacts on social distancing of two treatments: 1) social norm correction, and 2) leader endorsement. Before initiating the Round 2 survey, we randomly assigned households who completed the Round 1 survey to one of two treatments or a control group; probabilities were 30% for each of the treatment conditions and 40% for the control condition. Randomization was carried out on the computer of one of the co-authors. Survey staff implemented treatments after completing the Round 2 baseline survey, at the end of the same phone call.

To implement the social norm correction treatment, we first asked in Round 1 whether individuals themselves support social distancing, to calculate the share in the community supporting social distancing. We then asked individuals in Round 2 to estimate that share (reported as an integer out of 10). Individuals underestimating the share were told the true share supporting social distancing, also as an integer out of 10. Individuals correctly estimating the share were told that they were correct. In practice, 92.4% of treated respondents received this treatment, 53.2% of whom underestimated community support for social distancing and 46.8% of whom correctly estimated it. The small minority overestimating the share were not provided additional information.

To implement the leader endorsement treatment, we first identified and surveyed community opinion leaders prior to the Round 2 survey. These were local leaders serving in an official capacity (e.g., village chief), as well as individuals nominated during data collection for Yang et al. (2021) as community members who were best at circulating information within the community.[8] During our survey, we requested their permission to tell others in their community that they "support social distancing, are practicing social distancing, and encourage others to do the same". Then, in this treatment, we reported this endorsement of social distancing to respondents, mentioning the community leader(s) by name.

Complete treatment implementation protocols and scripts can be found in Appendix C.

[8]We followed the "gossips" methodology of Banerjee et al. (2020).

We also randomly assigned treatments to improve COVID-19 knowledge in the same study population. Randomization of the social norm correction and leader endorsement treatments were stratified within 76 communities and within the separate knowledge treatment conditions. Further details are in Appendix D, where we also present regression results showing that there are no interactions between the knowledge treatments and this paper's treatments of interest.

Sample sizes by treatment condition were as follows: social norm correction (N=628, 29.7% of sample), leader endorsement (N=637, 30.1%), and control group (N=852, 40.3%). Attrition between Rounds 2 (baseline) and 3 (endline) is low (4.9%). In Appendix E, we show that attrition between Rounds 2 and 3 and key baseline variables are balanced across treatment conditions.

4.2 Regressions

We estimate intent-to-treat effects using the following ordinary-least-squares regression specification:[9]

$$Y_{ijd} = \beta_0 + \beta_1 T1_{ijd} + \beta_2 T2_{ijd} + \eta B_{ijd} + \delta_{ijd}^{others} + \delta_{ijd}^{leaders} + \gamma_{jd} + \varepsilon_{ijd} \tag{8}$$

where Y_{ijd} is the social distancing indicator for respondent i in community j and district d; $T1_{ijd}$ and $T2_{ijd}$ are indicator variables for the social norm correction and leader endorsement treatment groups, respectively; B_{ijd} is the baseline value of the dependent variable; γ_{jd} are community fixed effects; and ε_{ijd} is a mean-zero error term. We report robust standard errors.

We also control for the number of other respondents and community leaders who report knowing the respondent at baseline. δ_{ijd}^{others} is a vector of dummy variables for the number of other respondents who report knowing the respondent's household from 0 to 8. $\delta_{ijd}^{leaders}$ is a vector of dummy variables for the number of community leaders who report knowing the respondent's household from 0 to 4.[10] Including these controls reduces residual variance, because they are predictive of social distancing.

Coefficients β_1 and β_2 represent the impacts of the social norm correction and leader endorsement treatments (respectively) on social distancing.

We modify Equation 8 to estimate heterogeneity in treatment effects with respect to local COVID-19 case loads:

$$\begin{aligned} Y_{ijd} = \beta_0 + \beta_1 T1_{ijd} + \beta_2 T2_{ijd} + \beta_3 (T1_{ijd} * Covid_d) + \beta_4 (T2_{ijd} * Covid_d) \\ + \eta B_{ijd} + \delta_{ijd}^{others} + \delta_{ijd}^{leaders} + \gamma_{jd} + \varepsilon_{ijd} \end{aligned} \tag{9}$$

[9] We show in Appendix G that all conclusions are robust to estimating logit and probit regressions instead.

[10] As pre-specified, we cap δ_{ijd}^{others} at the first integer that covers over 90% of the sample, and $\delta_{ijd}^{leaders}$ at the maximum number of leaders found in any community.

Equation 9 adds interactions between treatment indicators and the cumulative number of COVID-19 cases per 100,000 population at the district level at the start of the endline survey.[11] Coefficients β_1 and β_2 in Equation 9 now represent the impacts of the treatments in districts where COVID-19 cases are zero (slightly out of sample); β_3 and β_4 represent the change in the respective treatment effect for a one-unit increase in COVID-19 cases per 100,000 population.

4.3 Hypotheses

In our PAP, we hypothesized that effects of both treatments (β_1 and β_2 in Equation 8) would be positive. Subject-matter experts (surveyed without knowing results) concurred with this expectation.[12] The mean expert predictions were that the social norm correction and leader endorsement treatments would increase social distancing by 5.23 and 5.56 percentage points, respectively.

We also test hypotheses about treatment effect heterogeneity: the impact of the social norm correction treatment on social distancing and on the expected future infection rate will be more positive the higher the current COVID-19 infection rate (β_3 in Equation 9 will be positive). We did not specify these hypotheses in advance, but put them forward on the basis of our theoretical model.

5 Results

5.1 Average Treatment Effects

Estimates of average treatment effects are represented by coefficients β_1 and β_2 in Equation 8.

We first examine whether the treatments affected perceptions of the social norm. We did not pre-specify this analysis in our PAP. We present these results in Appendix H. Respondents report their perceived social norm in whole numbers out of 10, which we convert to a 0-1 scale. The social norm correction treatment has a positive effect on this measure that is marginally statistically significantly different from zero (p-value 0.12). Examination of the full CDF of the perceived social norm (Figure A.3) shows that the treatment effect is concentrated on the lower end of the distribution, reducing the frequency of perceptions below 50%. In regression analysis, the treatment has a positive and statistically significant effect on perceiving at least 50% of households in their community support social distanc-

[11] The main effect of $Covid_d$ is absorbed by γ_{jd}.

[12] 71 individuals provided predictions at https://socialscienceprediction.org/ (survey closing date January 2, 2021).

ing.[13] By contrast, the leader endorsement treatment has no statistically significant effect on the perceived social norm in any specification.

This analysis likely understates the effect of the social norm correction treatment on actual perceived social norms. Stated perceived social norms may be biased upwards due to experimenter demand effects (to make communities appear compliant). Notably, half of respondents in all rounds estimate 100% support of social distancing. If some share of these respondents overstate their perceptions of the social norm before treatment, and social norm correction led them to raise their actual social norm perceptions, this would not be observable to us.[14]

We now turn to treatment effect estimates in Table 2. In the regression for our primary social distancing outcome (Column 1), both treatment coefficients are small in magnitude and neither is statistically significantly different from zero at conventional levels. These findings diverge from expert predictions of treatment effects, which were each positive and roughly 5-6 percentage points.[15]

5.2 Treatment Effect Heterogeneity

Our theory predicts that the social norm correction treatment may either increase or decrease social distancing, depending on the local infection rate. In Table 2 Column (2), we present regression estimates of treatment effect heterogeneity (Equation 9) with respect to the local infection rate (COVID-19 cases per 100,000 population in the respondent's district).

The social norm correction treatment effect is heterogeneous with respect to local COVID-19 cases. The coefficient on the interaction term with $T1_{ijd}$ is positive and statistically significant at the 1% level. The coefficient on the $T1_{ijd}$ main effect is the predicted effect of social norm correction in a district with zero cases (slightly out of sample), and suggests that the social norm correction would reduce social distancing by 3.4 percentage points in such a location (statistically significant at the 5% level).

Figure 2 graphically depicts the social norm correction treatment effect heterogeneity. We plot district-specific treatment effects (estimating Equation 8 separately in each of seven districts) on the y-axis (with 95% confidence intervals) against district case counts on the x-axis. In the six districts with the lowest case counts, coefficients are negative. By contrast, in Chimoio, the district with the most cases (39.08/100,000) and that accounts for one-quarter of the sample, we estimate a large positive effect: 9.3 percentage points (a 75% increase over that district's control group, statistically significant at the 5% level).[16]

[13]There is also a positive and statistically significant effect on whether respondents increase their perceived community support between baseline and endline.

[14]The social norm correction treatment could also have raised confidence in high estimates of perceived social norms, even if stated perceptions remained constant.

[15]We reject at conventional significance levels that our T1 and T2 treatment effect estimates are equal to the mean expert predictions (p-value$<$0.001 in each case).

[16]This heterogeneous treatment effect finding does not depend on inclusion of Chimoio in the sample.

By contrast, the leader endorsement treatment effect is not heterogeneous with respect to local case loads. The coefficient on the corresponding interaction term in Column (1) is small in magnitude and not statistically significantly different from zero at conventional levels.

The interplay between the free-riding and perceived-infectiousness effects is the distinctive feature of our theoretical model. When the perceived-infectiousness effect is large enough, it overcomes the countervailing free-riding effect, and the social norm correction treatment leads to more social distancing. An additional implication of the theory is that the treatment should have similar heterogeneous effects on the expected future infection rate.

We conduct this additional test of the theory, examining treatment effects on the expected future infection rate.[17] In Columns (3) and (4) of Table 2, the outcome is the share of the community the respondent thinks will get sick from COVID-19 (responses were integers out of 10; we divide by 10 to yield a 0-1 scale). In Column (3), we estimate average treatment effects. Each coefficient is small in magnitude and not statistically significantly different from zero.

In Column (4), we estimate heterogeneity in treatment effects with respect to local cases, and find the same pattern as in Column (2). The social norm correction decreases the expected future infection rate in districts with no cases, and this impact becomes more positive as current cases rise (the $T1_{ijd}$ main effect and interaction term coefficients are both statistically significant at the 5% level).

These treatment effect heterogeneity findings (Columns 2 and 4, Table 2) jointly support the theoretical model. When current infection rates are low, the social norm correction treatment does not change perceived infectiousness much, but leads to realizations that social distancing support is higher than previously thought. People therefore reduce estimates of the future infection rate, and also reduce their own social distancing (choosing to free-ride). By contrast, when current infection rates are high, the treatment causes larger increases in perceived infectiousness. Notwithstanding an increase in the share of social distancing supporters, people increase their estimate of the future infection rate, and increase their social distancing.

When estimating Equations 8 and 9 for the sample excluding Chimoio, results are very similar. See Appendix I for details.

[17]The question is "For every 10 people in your community, how many do you think would get sick from coronavirus?" Sample sizes in these regressions are smaller. We implemented this question midway through the endline survey, after finding preliminary evidence suggesting the need to explore mechanisms behind treatment effect heterogeneity.

6 Conclusion

Support for social distancing has increased substantially during the COVID-19 pandemic. If people are unaware of the extent to which norms have changed, would revealing true high rates of such support lead to more social distancing? In theory, the impact of providing such information is ambiguous: it could reduce social distancing, if free-riding effects dominate, but could have a positive effect on social distancing if perceived-infectiousness effects dominate. Perceived-infectiousness effects are more likely to dominate when the current local infection rate is higher.

We implemented a randomized controlled trial testing the impact of a "social norm correction" treatment revealing high community support for social distancing. The treatment effect on social distancing exhibits the spatial heterogeneity predicted by theory: negative in areas with low infection rates (reflecting the dominance of free-riding effects), and more positive in areas with higher rates (as perceived-infectiousness effects become increasingly prominent). In the area with the most cases, amounting to one-quarter of our sample, the treatment effect is positive and large in magnitude. The treatment effect on the expected future infection rate shows similar heterogeneity, confirming an additional theoretical prediction.

Our results suggest that when local infection rates are high, interventions shifting perceptions of community social distancing support upwards could help promote social distancing. These findings may also help predict the impacts of analogous public health messaging revealing social norms of support for preventive measures against other infectious diseases.

References

Alsan, M., F. Cody Stanford, A. Banerjee, E. Breza, A. G. Chandrasekhar, S. Eichmeyer, P. Goldsmith-Pinkham, L. Ogbu-Nwobodo, B. A. Olken, C. Torres, A. Sankar, P. Vautrey, and E. Duflo (2020). Comparison of Knowledge and Information-Seeking Behavior After General COVID-19 Public Health Messages and Messages Tailored for Black and Latinx Communities: A Randomized Controlled Trial. *Annals of internal medicine*.

Andersson, O., P. Campos-Mercade, A. Meier, and E. Wengstrom (2021). Anticipation of COVID-19 Vaccines Reduces Social Distancing. *SSRN Working Paper 3765329*.

Arroyos-Calvera, D., M. Drouvelis, J. Lohse, and R. McDonald (2021). Improving Compliance with COVID-19 Guidance: A Workplace Field Experiment. *SSRN*, 19.

Banerjee, A., M. Alsan, E. Breza, A. G. Chandrasekhar, A. Chowdhury, E. Duflo, P. Goldsmith-Pinkham, and B. A. Olken (2020). 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. *NBER Working Paper #2747*.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2019). Using Gossips to Spread Information: Theory and Evidence from Two Randomized Controlled Trials. *The Review of Economic Studies 86*, 2453–2490.

Banerjee, A., E. La Ferrara, and V. H. Orozco-Olvera (2019). The Entertaining Way to Behavioral Change: Fighting HIV with MTV. *NBER Working Paper #26096*.

Banker, S. and J. Park (2020). Evaluating Prosocial COVID-19 Messaging Frames: Evidence from a Field Study on Facebook. *Judgment and Decision Making 15*, 7.

Benabou, R. and J. Tirole (May 2011). Identity, Morals, and Taboos: Beliefs as Assets. *Quarterly Journal of Economics 126*, 805–855.

Bernheim, B. D. (1994). A Theory of Conformity. *Journal of Political Economy 102*, 841–877.

Bicchieri, C. (2005). *The grammar of society: The nature and dynamics of social norms.* Cambridge University Press.

Bicchieri, C. and E. Dimant (2019). Nudging with care: The risks and benefits of social information. *Public choice*, 1–22.

Bos, B., M. A. Drupp, J. Meya, and M. Quaas (2020). Moral Suasion and the Private Provision of Public Goods: Evidence from the COVID-19 Pandemic. *SSRN*, 24.

Bursztyn, L., A. L. Gonzalez, and D. Yanagizawa-Drott (2018). Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia. *American Economic Review 110*, 2997–3029.

Capraro, V. and H. Barcelo (2020). Priming Reasoning Increases Intentions to Wear a Face Covering to Slow Down COVID-19 Transmission. *arXiv*.

Cato, S., T. Iida, K. Ishida, A. Ito, K. M. McElwain, and M. Shoji (2020). Social Distancing as a Public Good Under the COVID-19 Pandemic. *Public Health 188*, 51–53.

CDC (2020). COVID-19 and Your Health: Social Distancing. *Center for Disease Control and Prevention*.

Centola, D. (2011). An Experimental Study of Homophily in the Adoption of Health Behavior. *Science 334*, 1269–1272.

Cialdini, R. B. and N. J. Goldstein (February 4, 2004). Social Influence: Compliance and Conformity. *Annual Review of Psychology 55*, 591–621.

De Quidt, J., J. Haushofer, and C. Roth (2018). Measuring and Bounding Experimenter Demand. *American Economic Review 108*, 3266–3302.

Dupas, P. (2014). Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence From a Field Experiment. *Econometrica 82*, 197–228.

Habersaat, K. B. and A. E. Scheel (November 2020). *Pandemic Fatigue: Reinvigorating the Public to Prevent COVID-19: Policy Framework for Supporting Pandemic Prevention and Management.* Copenhagen, Denmark: World Health Organization.

Hershey, J. C., D. A. Asch, T. Thumasathit, J. Meszaros, and V. W. Waters (1994). The Roles of Altruism, Free Riding, and Bandwagoning in Vaccination Decisions. *Organizational Behavior and Human Decision Processes 59*, 177–187.

Jakubowski, A., D. Egger, C. Nekesa, L. Lowe, M. Walker, and E. Miguel (2021). Self-Reported Mask Wearing Greatly Exceeds Directly Observed Use: Urgent Need for Policy Intervention in Kenya. *medRxiy Preprint*.

Janzwood, S. (April 27, 2020). *The Social Distancing Norm Cascade: The Role of Belief Systems in Accelerating Normative Change during the COVID-19 Pandemic*. Canada: Cascade Institute.

Jones, S., E. Egger, and R. Santos (2020). Is Mozambique Prepared for a Lockdown During the COVID-19 Pandemic? *UNU-WIDER Blog*.

Jordan, J., E. Yoeli, and D. Rand (2020). Don't Get it or Don't Spread it? Comparing Self-Interested Versus Prosocially Framed COVID-19 Prevention Messaging. *PsyArXiv Preprint*.

Kelly, J. A., J. St.Lawrence, Y. S. ad Allan Hauth, S. Kalichman, Y. Diaz, T. Brasfield, J. Koob, and M. Morgan (1992). Community AIDS/HIV Risk Reduction: The Effects of Endorsements by Popular People in Three Cities. *American Journal of Public Health 82*, 1483–1489.

Kim, D. A., A. R. Hwong, D. Stafford, A. D. Hughes, J. A. O'Malley, J. H. Fowler, and N. A. Christakis (2015). Social Network Targeting to Maximise Population Behaviour Change: A Cluster Randomised Controlled Trial. *Lancet*, 145–153.

Krupka, E. L. and R. A. Weber (2013, 06). Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary? *Journal of the European Economic Association 11*(3), 495–524.

Lau, K., M. Miraldo, M. M. Galizzi, and K. Hauck (2019). Social Norms and Free-riding in Influenza Vaccine Decisions in the UK: An Online Experiment. *The Lancet 394*, 65.

Lunn, P. D., S. Timmons, C. A. Belton, M. Barjakova, H. Julienne, and C. Lavin (2020). Motivating Social Distancing During the Covid-19 Pandemic: An Online Experiment. *Social Science and Medicine*.

Ministry of Health (2020). *COVID-19 Epidemiological Situation, Republic of Mozambique.* Maputo, Mozambique: Government of Mozambique.

National Institute of Statistics (INE) (2017). *Mozambique Population and Housing Census 2017.* Maputo, Mozambique: Government of Mozambique.

Nyusi, F. (August 5, 2020a). *Communication to the Nation of His Excellency Philip Jacinto Nyusi, President of Republic of Mozambique, on the New State of Emergency, within the Scope of the Coronavirus Pandemic COVID-19.* Maputo: Maputo Mozambique.

Nyusi, F. (September 5, 2020b). *Communication to the Nation of His Excellency Philip Jacinto Nyusi, President of Republic of Mozambique, on the New State of Emergency, within the Scope of the Coronavirus Pandemic COVID-19.* Maputo: Maputo Mozambique.

Orne, M. T. (1962). On the Social Psychology of the Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications. *American Psychologist 17*, 776–783.

Paakkari, L. and O. Okan (2020). COVID-19: Health Literacy is an Underestimated Problem. *The Lancet Public Health 5.*

Papanastasiou, A., B. J. Ruffle, and A. Zheng (2020). Compliance with Social Distancing: Theory and Empirical Evidence from Ontario during COVID-19. *SSRN*, 38.

Reicher, S. and J. Drury (January 18, 2021). Pandemic Fatigue? How Adherence to COVID-19 Regulations has been Misrepresented and Why it Matters. *British Medical Journal 372*, 137.

Republic of Mozambique (April 2, 2020c). *"Bulletin of the Republic", I Series, No. 64.*

Republic of Mozambique (August 5, 2020a). *"Bulletin of the Republic", I Series, No. 149.*

Republic of Mozambique (March 31, 2020b). *"Bulletin of the Republic", I Series, No. 62.*

Rosenthal, R. (1966). *Experimenter Effects in Behavioral Research.* New York, USA: Appleton-Century-Crofts.

Sasaki, S., H. Kurokawa, and F. Ohtake (2020). Short-term Responses to Nudge-based Messages for Preventing the Spread of COVID-19 Infection: Intention, Behavior, and Life Satisfaction. *Osaka University Discussion Papers In Economics And Business*, 1–31.

Schultz, W. P., J. M. Nolan, R. B. Cialdini, N. J. Goldstein, and V. Griskevicius (2007). The Constructive, Destructive, and Reconstructive Power of Social Norms. *Psychological Science 18*, 429–434.

Siuta, M. and M. Sambo (April 1, 2020). *COVID-19 Em Mocambique: Dimensao e possiveis impactos. Boletim No. 124p.* Maputo: Instituto de Estudos Socias e Economicos.

U.S Embassy in Mozambique (2020). COVID-19 Information.

Valente, T. and P. Pumpuang (2007). Identifying Opinion Leaders to Promote Behavior Change. *Health Education and Behavior 34*, 881–896.

Van Bavel, J. J., K. Baicker, P. S. Boggio, V. Capraro, A. Cichocka, M. Cikara, M. J. Crockett, A. J. Crum, K. M. Douglas, J. N. Druckman, J. Drury, O. Dube, N. Ellemers, E. J. Finkel, J. H. Fowler, M. Gelfand, S. Han, S. Haslam, J. Jetten, S. Kitayama, D. Mobbs, L. E. Napper, D. J. Packer, G. Pennycook, E. Peters, R. E. Petty, D. G. Rand, S. D. Reicher, S. Schnall, A. Shariff, L. J. Skitka, S. S. Smith, C. R. Sunstein, N. Tabri, J. A. Tucker, S. van der Linden, P. van Lange, K. A. Weeden, M. J. A. Wohl, J. Zaki, S. R. Zion, and R. Willer (2020). Using Social and Behavioural Science to Support COVID-19 Pandemic Response. *Nature human behaviour 4*, 460–471.

Wood, W. (February 2000). Attitude Change: Persuasion and Social Influence. *Annual Review of Psychology 51*, 539–570.

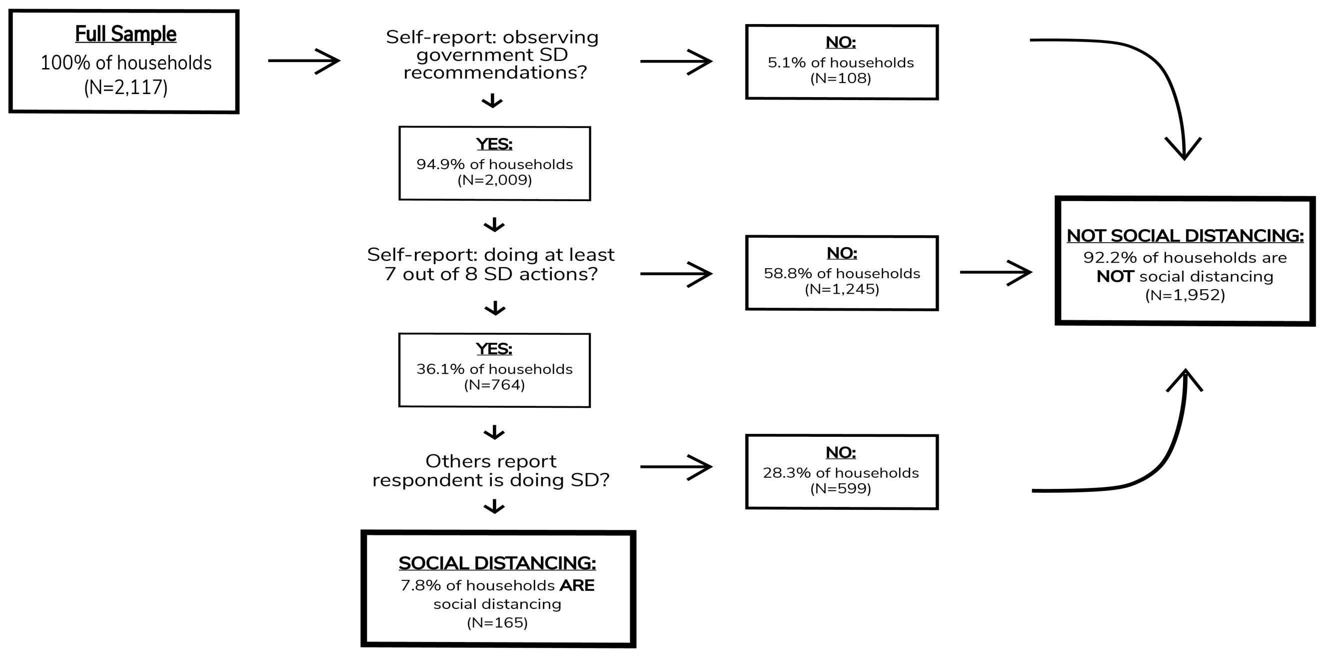
Xie, K., B. Liang, M. A. Dulebenets, and Y. Mei (2020). The Impact of Risk Perception on Social Distancing During the COVID-19 Pandemic in China. *International journal of environmental research and public health 17*.

Yang, D., J. Allen IV, A. Mahumane, J. Riddell IV, and H. Yu (2021). Knowledge, Stigma and HIV Testing: An Analysis of a Widspread HIV/AIDS Program. *Working Paper*.

Yu, H. (2020). *Three Essays in Development Economics. Dissertation.* Ann Arbor, MI: University of Michigan.

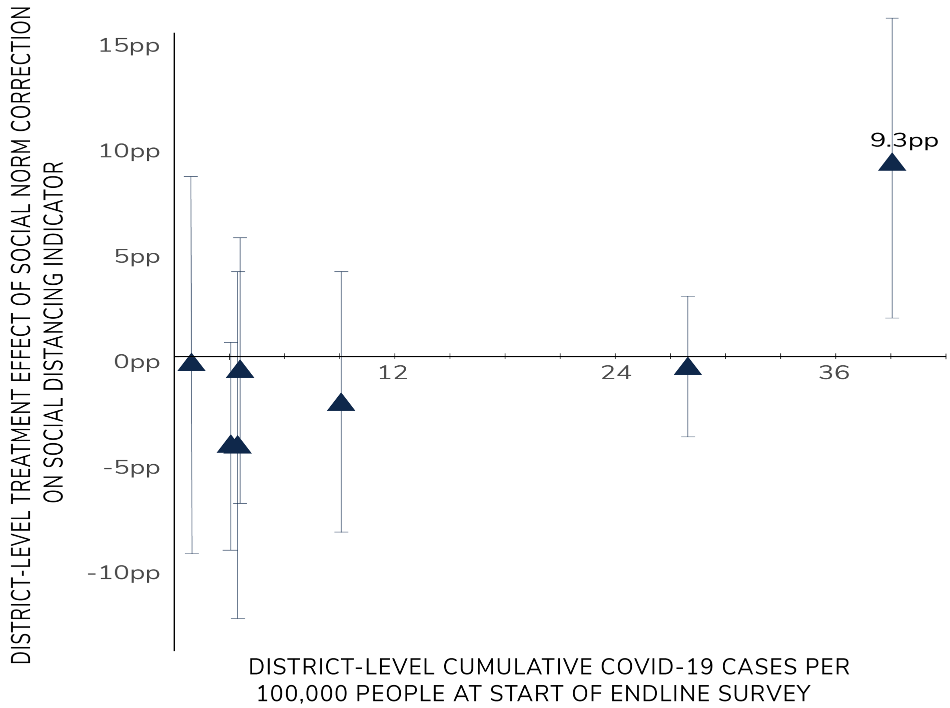
Zizzo, D. J. (2010). Experimenter Demand Effects in Economic Experiments. *Experimental Economics 13*, 75–98.

Figure 1: **The Social Distancing Measure**



Notes: Definition of primary social distancing (SD) outcome measure. Definition was pre-specified in pre-analysis plan (PAP) and submitted to AEA RCT Registry on Aug. 25, 2020. Data for constructing the measure were then collected in our Rounds 2 and 3 phone surveys (from August through November 2020). To be considered social distancing (SD), respondents must: 1) self-report they are following government SD recommendations, 2) self-report they are doing at least seven out of eight SD actions, and 3) be reported by others in community to be doing SD. Percentages reported are all shares of full sample (N=2,117). See Table 1 and Section 3.3 of main text for social distancing question definitions.

Figure 2: **District-Level Social Norm Correction Treatment Effects by Cumulative COVID-19 Cases**



Notes: Social norm correction treatment effects (triangles) estimated separately for each of seven study districts (with 95% confidence intervals). District-level treatment effects plotted on vertical axis against district COVID-19 case loads at start of endline survey (per 100,000 population) on horizontal axis.

Table 1: **Summary Statistics of Social Distancing Support, Norms, and Behavior**

	VARIABLES	N	Mean	SD	Min	Max
(1)	Pre-baseline: Respondent supports social distancing (SD)	2,117	0.976	0.153	0	1
(2)	Pre-baseline: Perceived share of community supporting SD	2,109	0.689	0.313	0	1
(3)	Respondent supports SD	2,117	0.989	0.104	0	1
(4)	Perceived share of community supporting SD	2,114	0.800	0.262	0	1
(5)	Primary SD indicator	2,117	0.0784	0.269	0	1
(6)	Self-report of SD indicator	2,117	0.355	0.479	0	1
(7)	Self-report: Followed govt rules in past 14 days	2,117	0.949	0.219	0	1
(8)	Self-report: SD behaviors above median	2,117	0.361	0.481	0	1
(9)	Others' report of SD indicator	2,117	0.232	0.422	0	1
(10)	Other households' report of SD	2,117	0.378	0.485	0	1
(11)	Leaders' report of SD	2,117	0.519	0.500	0	1

Notes: "Pre-baseline" refers to data collected during the Round 1 phone survey from July 10 to August 16, 2020. All other summary statistics are baseline data collected during the Round 2 telephone survey from August 26 to October 4, 2020. See study timeline in Figure A.2. Variables are as follows. Rows 1 & 3: indicator equal to one if respondent answers "yes" to supporting "the practice of social distancing to prevent the spread of coronavirus" and zero otherwise. Rows 2 & 4: perceived share of households (asked as "for every 10 households") in community that support social distancing (SD). Row 5: indicator for SD equal to one if respondent is SD according to self (Row 6) and others' reports (Row 9), and zero otherwise. Row 6: indicator for SD according to self if respondent answered "yes" to observing the government's recommendations on SD in the last 14 days (Row 7) and the report carrying out at least seven out of eight (above the sample median) social distancing behaviors (Row 8), and zero otherwise. Row 9: indicator for SD according to others if all other respondents (Row 10) and community leaders (Row 11) reported not knowing the respondent household, not seeing the respondent household in the past 14 days, or—if seen—that the respondent household 1) did NOT come closer than 1.5 meters to others outside their household; 2) did NOT shake hands, try to shake hands, or touch others outside their household; and 3) appeared to be observing the government's recommendations on SD, and zero otherwise.

Table 2: **Treatment Effects on Social Distancing and Expected COVID-19 Illnesses**

VARIABLES	(1) Primary SD Indicator	(2) Primary SD Indicator	(3) Perceived share of people in community that will get sick from Covid	(4) Perceived share of people in community that will get sick from Covid
T1: Social Norm Correction	0.00425	-0.0466**	0.0418	-0.194**
	(0.0140)	(0.0191)	(0.0322)	(0.0944)
T2: Leader Endorsement	-0.00541	-0.0258	-0.0209	-0.0598
	(0.0137)	(0.0198)	(0.0308)	(0.0944)
T1 × District Covid Cases		0.00304***		0.00725**
		(0.00106)		(0.00291)
T2 × District Covid Cases		0.00122		0.00127
		(0.000992)		(0.00287)
Observations	2,117	2,117	812	812
R-squared	0.158	0.163	0.146	0.152
Control Mean DV	0.0857	0.0857	0.359	0.359
Control SD DV	0.280	0.280	0.369	0.369

Notes: Dependent variable in Columns 1-2 defined in Table 1. Dependent variable in Columns 3-4 is the expected future infection rate: "For every 10 people in your community, how many do you think would get sick from coronavirus?" (converted to share from 0 to 1). "T1: Social Norm Correction" is equal to one if respondent was randomly assigned to the social norm correction treatment, and zero otherwise. "T2: Leader Endorsement" is equal to one if respondent was randomly assigned to the leader endorsement treatment, and zero otherwise. "T1 x District Covid Cases" & "T2 x District Covid Cases" are the respective treatment indicators interacted with district-level cumulative COVID-19 cases per 100,000 population at the start of the endline survey (as detailed in Appendix J). All regressions control for a baseline measure of the dependent variable, a vector of indicators for number of community leaders knowing the respondent at baseline (0 through 4), and a vector of indicators for number of other respondents knowing the respondent at baseline (0 through 8). All regressions also include community fixed effects. Robust standard errors in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Online Appendix

A Proofs

A.1 Proof of Theorem 1

The agent will adjust her effort level in response to the treatment to $\sqrt{\hat{e}} = \frac{\hat{\alpha}C}{2}H(\hat{x})$ where $H(x) = 1 - x(1 - \frac{1}{\sqrt{2}})$. Hence, the prior and posterior effort levels satisfy:

$$\frac{\sqrt{\tilde{e}}}{\sqrt{e}} = \frac{\tilde{\alpha}}{\alpha}\frac{H(\tilde{x})}{H(x)} \tag{A.1}$$

We take the ratios of Equations 6 and 7:

$$\frac{\hat{\alpha}}{\alpha} = \frac{1 - \sqrt{e}G(x)}{1 - \sqrt{e}G(\hat{x})} \tag{A.2}$$

We therefore obtain:

$$\frac{\sqrt{\tilde{e}}}{\sqrt{e}} = \frac{H(\hat{x})(1 - \sqrt{e}G(x))}{H(x)(1 - \sqrt{e}G(\hat{x}))} \tag{A.3}$$

Effort increases iff $\frac{\sqrt{\tilde{e}}}{\sqrt{e}} > 1$:

$$\begin{aligned} \frac{H(\hat{x})(1 - \sqrt{e}G(x))}{H(x)(1 - \sqrt{e}G(\hat{x}))} &> 1 \\ \sqrt{e}\left[H(x)G(\hat{x}) - H(\hat{x})G(x)\right] &> H(x) - H(\hat{x}) \end{aligned}$$

Now note that $G(x) = 2xH(x)$ such that:

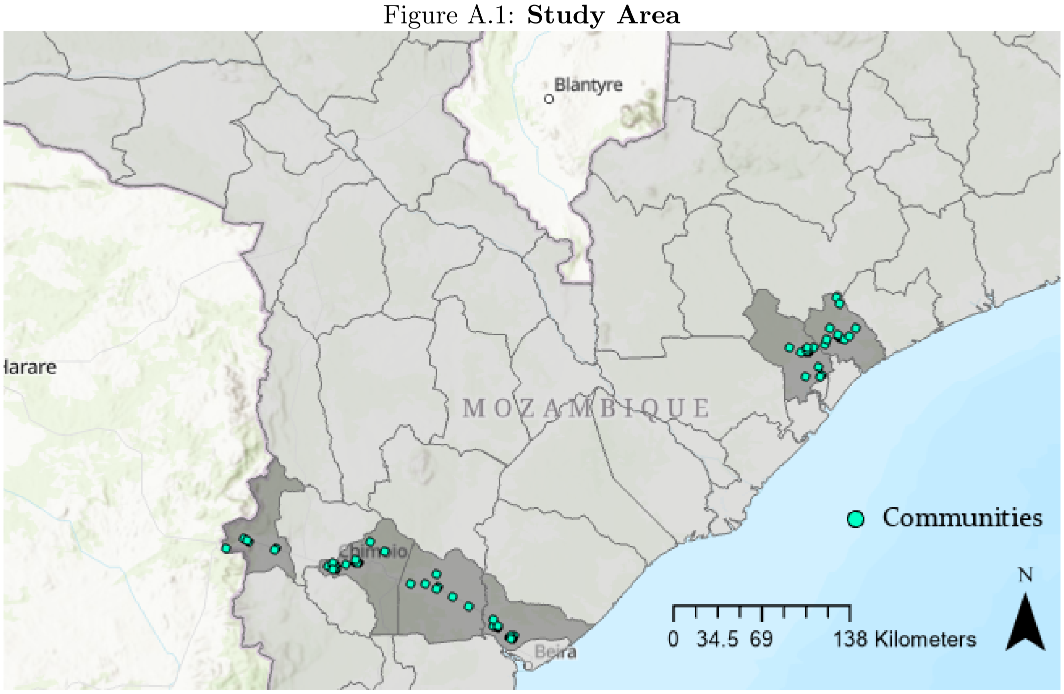
$$\begin{aligned} \sqrt{e}2H(x)H(\hat{x})(\hat{x} - x) &> (1 - \frac{1}{\sqrt{2}})(\hat{x} - x) \qquad \text{(A.4)} \\ \sqrt{e} &> \frac{1 - \frac{1}{\sqrt{2}}}{H(x)H(\hat{x})} \end{aligned}$$

This shows that the perceived-infectiousness effect dominates if the initial effort level e is high enough. Effort is determined by Equation 5 and increases with α (which increases with $\hat{\alpha}$). Therefore, for sufficiently large $\hat{\alpha}$ the perceived-infectiousness effect dominates.

B Study Area and Timeline

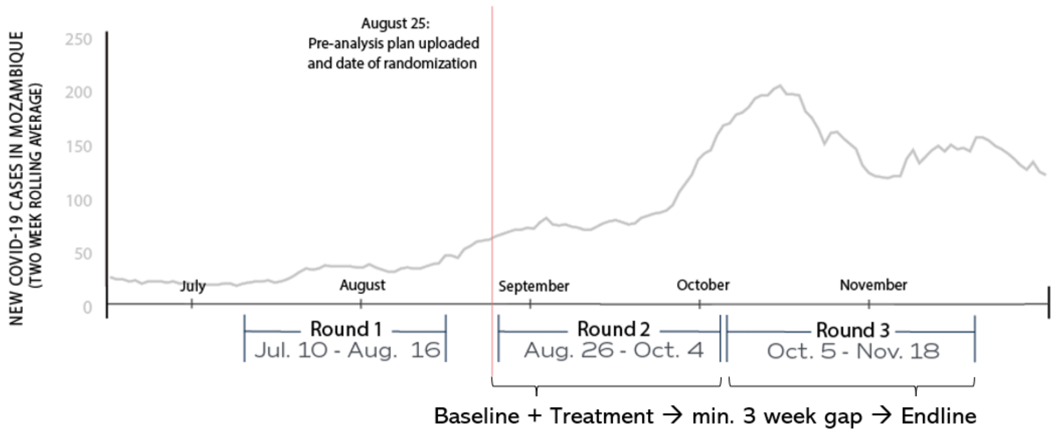
Study participants come from 76 communities in central Mozambique. The study communities are in seven districts of three provinces: Dondo and Nhamatanda in Sofala province; Gondola, Chimoio and Manica in Manica province; and Namacurra and Nicoadala in Zambezia province. These 76 communities are mapped in Figure A.1. Compared to other communities in Mozambique, the study areas are relatively accessible to main transport corridors (highways and ports), and are thus important geographic conduits for infectious disease.

Figure A.1: **Study Area**



We collected survey data in three rounds between July 10 and November 18, 2020. Appendix Figure A.2 depicts the study timeline below a rolling average of new Mozambican COVID-19 cases. We piloted surveys in Round 1. Immediately before the Round 2 survey, we randomly assigned households to treatments and submitted our pre-analysis plan to the AEA RCT Registry. The Round 2 survey served as a baseline, and was immediately followed (on the same phone call) by our treatment interventions. Round 3 was our endline survey. Surveys collected data on COVID-19 knowledge, beliefs, and behaviors. While data collection for Round 3 began only one day after completion of Round 2, there was a minimum of 3.0 weeks and average of 6.3 weeks between Rounds 2 and 3 surveys for any given respondent. While the Round 1 survey occurred when new COVID-19 cases remained relatively steady, both the Round 2 and Round 3 surveys occurred during a period of substantial growth in new COVID-19 cases.

Figure A.2: **Study Timeline**



Notes: Round 1 is pre-baseline survey to collect social distancing support data, Round 2 is baseline survey, and Round 3 is endline survey. There is at least a three week gap between baseline and endline survey for any given study participant. Pre-analysis plan uploaded and treatments randomly assigned immediately prior to start of Round 2 baseline survey, on Aug. 25, 2021. Treatments implemented immediately following baseline survey on same phone call. Baseline measures reported in Table 1 come from Round 2 surveys and endline measures come from Round 3 surveys.

C Treatment Scripts

Both treatments were implemented directly following the baseline survey, on the same phone call. If a respondent was randomly assigned to a treatment, the corresponding intervention text would appear on the enumerator's tablet. Enumerators read a script aloud exactly as shown below. Following the treatment, respondents were asked if they would like the information repeated.

The social norm correction treatment involves sharing the level of "actual community support for social distancing" with the respondent if they underestimate or correctly estimate that level. We express this as integer units out of 10, rounded to the nearest integer based on actual shares of respondents in the community expressing support for social distancing in the Round 1 (pre-baseline) survey. In 63 out of 76 communities (82.9%) the number we convey to respondents is 10 out of 10, and in 13 communities (17.1%) the number is 9 out of 10.

Script for T1: Social Norm Correction – "Now I want to give you some information about social distancing. In this survey, you indicated that you think *<insert respondent's answer here>* of every 10 households in your community support the practice of social distancing."

- *If response UNDERESTIMATES community support for social distancing:* "However, more households support social distancing than you think! Based on the results of our first COVID-19 survey, approximately *<insert actual community support for social distancing here>* of every 10 households in your community support social distancing to prevent the spread of the coronavirus."

- *If response CORRECTLY ESTIMATES community support for social distancing:* "You are correct! Based on the results of our first COVID-19 survey, approximately *<insert actual community support for social distancing here>* of every 10 household in your community support social distancing to prevent the spread of the coronavirus."

- *If response OVERESTIMATES community support for social distancing: (no information given)*

Script for T2: Leader Endorsement – "Our research team recently called and talked to your *<list leaders' titles and names here>*. They said that they support social distancing, are practicing social distancing themselves, and encourage others to do the same."

D Populated Pre-analysis Plan

On August 25, 2020, prior to baseline data collection, we uploaded our pre-analysis plan (PAP) to the American Economic Association's RCT Registry, registration ID number AEARCTR-0005862: https://doi.org/10.1257/rct.5862-1.0.

In our PAP, we specify the following regression for our primary analysis, which is the same as Equation 8 in the main text:

$$Y_{ijd} = \beta_0 + \beta_1 T1_{ijd} + \beta_2 T2_{ijd} + \eta B_{ijd} + \delta_{ijd}^{others} + \delta_{ijd}^{leaders} + \gamma_{jd} + \varepsilon_{ijd} \tag{D.1}$$

where Y_{ijd} is the social distancing indicator for household i in community j and district d; $T1_{ijd}$ and $T2_{ijd}$ are indicator variables for the social norm correction and leader endorsement treatment groups, respectively; B_{ijd} is the baseline value of the dependent variable; γ_{jd} are community fixed effects; and ε_{ijd} is a mean-zero error term. We report robust standard errors. The regression also controls for the number of other survey respondents and community leaders who report knowing the survey respondent at baseline (in Round 2). Specifically, δ_i^{others} is a vector of dummy variables for the distinct number of other surveyed study respondents who report knowing the household (0, 1, 2..., 7, 8 or more; where 8 is the first integer where over 90% of the sample is represented by previous non-negative integers), and $\delta_i^{leaders}$ is a vector of dummy variables for the distinct number of community leaders who report knowing the household (0, 1, 2, 3, 4; where 4 is maximum number of leaders found within one of the 76 sample communities). Including this control variable helps reduce residual variance in the dependent variable, because respondents who are known by more others in the community will also have more reports of social interactions with others. These results are presented in the main paper in Table 2 column (1) and are also replicated in column (1) of Table A.1.

Additionally, we pre-specified the following secondary analyses. First, we analyze impacts of the social distancing treatments on the separate components of the social distancing index—the others' and self-report. These results are presented in Table A.1 columns (2) & (3), respectively. Treatment effects on these outcomes are very similar to those in column (1).

Second, we also pool SD1 and SD2 together to examine the effect of some endorsement of social distancing (whether by other community members or by community leaders) on the primary social distancing outcome. These coefficient in Table A.1 column (4) is small in magnitude and not statistically significantly different from zero at conventional levels.

Table A.1: **Additional Prespecified Analyses**

VARIABLES	(1) Primary SD Indicator	(2) Others' Report of SD	(3) Self-Report of SD	(4) Primary SD Indicator
T1: Social Norm Correction	0.00425	0.000950	0.0134	
	(0.0140)	(0.0181)	(0.0238)	
T2: Leader Endorsement	-0.00541	0.0145	-0.0189	
	(0.0137)	(0.0183)	(0.0234)	
Pooled SD Treatments				-0.000642
				(0.0116)
Observations	2,117	2,117	2,117	2,117
R-squared	0.158	0.333	0.211	0.158
Control Mean DV	0.0857	0.211	0.406	0.0857
Control SD DV	0.280	0.408	0.491	0.280

Notes: Dependent variables are defined in Table 1. "T1: Social Norm Correction" is an indicator equal to one if respondent was randomly assigned to the social norm correction treatment, and zero otherwise. "T2: Leader Endorsement" is an indicator equal to one if respondent was randomly assigned to the leader endorsement treatment, and zero otherwise. "Pooled SD Treatments" is an indicator equal to one if respondent was randomly assigned to the social norm correction treatment or leader endorsement treatment, and zero otherwise. Controls are as defined in Table 2. All regressions also include community fixed effects. Robust standard errors in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

We also randomly assigned a family of treatments to improve COVID-19 knowledge in the same study population.[1] Randomization of the social norm correction and leader endorsement treatments were stratified within 76 communities and within the separate knowledge treatment conditions (i.e., the knowledge and social distancing treatments were cross-randomized). As pre-specified, we run a regression on the primary social distancing outcome with indicators for social distancing treatments, the cross-randomized knowledge treatments and their interaction terms. Results are presented in Table A.2, and show no large or statistically significant interaction effects between the social distancing and knowledge treatments.

[1] The pre-analysis plan (PAP) for the knowledge study can be found here: https://fordschool.umich.edu/mozambique-research/combatting-covid-19.

Table A.2: **Interactions between Social Distancing and Knowledge Treatments**

	(1)
VARIABLES	Primary SD Indicator
T1: Social Norm Correction	-0.0237
	(0.0214)
T2: Leader Endorsement	-0.0210
	(0.0222)
K1: Incentive	-0.0218
	(0.0241)
K2: Feedback	-0.00250
	(0.0251)
K3: Incentive & Feedback	-0.0144
	(0.0238)
T1 × K1	0.0545
	(0.0390)
T2 × K1	0.0249
	(0.0372)
T1 × K2	0.0467
	(0.0397)
T2 × K2	0.0139
	(0.0385)
T1 × K3	0.0404
	(0.0382)
T2 × K3	0.0374
	(0.0372)
Observations	2,117
R-squared	0.160
Control Mean DV	0.0857
Control SD DV	0.280

Notes: Dependent variable is defined in Table 1. Social distancing treatments are defined in Table 2. "K1 Incentive", "K2 Feedback", and "K3 Incentive & Feedback" are indicators equal to one if respondent was randomly assigned to one of these knowledge treatments, and zero otherwise. Remaining regressors represent interactions between social distancing treatments and the knowledge treatments. Controls are as defined in Table 2. Regression also includes community fixed effects. Robust standard errors in parentheses. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

E Attrition and Balance

Appendix Table A.3 presents regressions examining whether attrition and baseline variables are balanced with respect to treatment assignment[2]. Attrition between Round 2 (baseline) and Round 3 (endline) is only 4.9% and is less than 5.6% in each of the seven districts surveyed. Balance in attrition is confirmed in column 1, which starts with the Round 2 (baseline) sample and regresses treatments on an indicator equal to one if the respondent was not reached for the Round 3 (endline) survey. Balance in baseline social distancing outcomes is confirmed in columns 2-4, which examines the Round 2 social distancing outcomes. Balance in baseline household characteristics is confirmed in columns 6-8, which examines the final Round 3 sample and regresses treatments on Round 1 measures of household income, an index of food insecurity, and an indicator for presence of an older adult over 60 years. In not a single regression in the table is a coefficient on a treatment indicator statistically significant at conventional levels.

[2]Figure A.2 shows the study timeline for the three survey rounds collected. Round 1 is a pre-baseline measure, Round 2 measures baseline values and Round 3 measures endline outcomes.

Table A.3: **Treatment Effect on Attrition and Balance**

VARIABLES	(1) Attrition	(2) Primary SD Indicator	(3) Others' Report of SD	(4) Self-Report of SD	(5) Perceived Social Norm	(6) Hh Income	(7) Food Insecurity	(8) Older Adult in Hh
T1: Social Norm Correction	-0.0127	-0.0176	-0.000450	-0.00956	-0.0101	-159.5	0.00107	-0.00293
	(0.0111)	(0.0134)	(0.0203)	(0.0247)	(0.0138)	(181.7)	(0.0191)	(0.0250)
T2: Leader Endorsement	-0.00154	-0.00324	0.00897	0.00420	-0.0201	-39.95	-0.0240	0.0240
	(0.0113)	(0.0143)	(0.0206)	(0.0249)	(0.0137)	(181.8)	(0.0193)	(0.0252)
Observations	2,226	2,117	2,117	2,117	2,114	1,873	2,117	2,096
R-squared	0.030	0.096	0.199	0.076	0.047	0.043	0.090	0.058
Control Mean DV	0.0533	0.0833	0.229	0.356	0.810	1176	0.842	0.342
Control SD DV	0.225	0.277	0.420	0.479	0.262	4029	0.365	0.475

Notes: Dependent variables are as follows. Column 1: indicator if respondent attrited from the sample between baseline and endline. Columns 2-4: baseline SD outcomes defined in Table 1. Column 5: baseline perceived share of community supporting SD, defined further in Table 1. Column 6: at pre-baseline, self-reported total income for the previous week (in Mozambican meticais). Column 7: indicator if, in the last 7 days, household has 1) lacked food; 2) reduced number of meals/portions; or was unable to buy their usual amount of food due to 3) market shortages, 4) high prices, 5) reduced income. Column 8: indicator if adult age 60 or older is present in the household. Controls are as defined in Table 2. All regressions also include community fixed effects. Robust standard errors in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

F Summary Statistics for Social Distancing Index

Below are the summary statistics for the questions that comprise the self-reported social distancing index at baseline and endline. Respondents were asked “Is this something your household has been doing for the last seven days?” about a randomly determined four social distancing actions at baseline and all eight social distancing actions at endline. Responses were coded as indicators equal to one if indicative of social distancing (answers that indicate social distancing shown in parentheses), and zero otherwise.

Table A.4: **Summary Statistics for Components of Social Distancing Index**

	Baseline			Endline		
VARIABLES	N	Mean	SD	N	Mean	SD
Shop in crowded areas like informal markets (No)	1,032	.642	.479	2,115	.678	.467
Gather with several friends (No)	1,047	.349	.477	2,113	.414	.493
Help the elderly avoid close contact with other people, including children (Yes)	1,094	.877	.329	2,114	.923	.266
If show symptoms of coronavirus, immediately inform my household and avoid people (Yes)	1,050	.836	.370	2,113	.859	.347
Drink alcohol in bars (No)	1,082	.226	.419	2,113	.272	.445
Wear a face mask if showing symptoms of coronavirus (Yes)	1,034	.902	.297	2,114	.885	.319
Instead of meeting in person, call on the phone or send text message (Yes)	1,039	.935	.247	2,112	.930	.255
Allow children to build immunity by playing with children from other households (No)	1,070	.439	.497	2,113	.456	.498

Notes: Variables are coded as indicators equal to one if indicative of social distancing (answers that indicate social distancing shown in parentheses), and zero otherwise. Respondents were asked “Is this something your household has been doing for the last seven days?” about a randomly determined four social distancing actions at baseline and all eight social distancing actions at endline. The baseline sample was asked a subset of these questions which explains the smaller number of observations at baseline.

G Treatment Effect Estimates from Logit and Probit Regressions

The primary social distancing indicator is a binary variable that is analyzed using an ordinary least-squares (OLS) regression, as prespecified. As a robustness check, we adapt Equation 8 to be run using logit and probit regression.

Table A.5 presents results from the logistic regression on the primary outcomes, while Table A.6 presents corresponding probit regression results. Regression coefficients are presented as marginal effects. Results in both tables are consistent with the results from OLS linear probability models presented in Table 2.

Table A.5: **Treatment Effects Estimated Using Logistic Regression**

VARIABLES	(1) Primary SD Indicator	(2) Primary SD Indicator	(3) Perceived share of households in community that will get sick from Covid	(4) Perceived share of households in community that will get sick from Covid
T1: Social Norm Correction	0.0100	-0.0756**	0.0270	-0.403***
	(0.0221)	(0.0376)	(0.0395)	(0.138)
T2: Leader Endorsement	-0.00695	-0.0398	-0.0274	-0.300**
	(0.0222)	(0.0349)	(0.0394)	(0.135)
T1 × District Covid Cases		0.00384***		0.0132***
		(0.00131)		(0.00397)
T2 × District Covid Cases		0.00162		0.00842**
		(0.00129)		(0.00397)
Observations	1,285	1,285	806	806
Control Mean DV	0.141	0.141	0.356	0.356
Control SD DV	0.349	0.349	0.368	0.368

Notes: Dependent variables are defined in Tables 1 and 2. Coefficients presented are marginal effects from logit regression. Social distancing treatments and controls are as defined in Table 2. All regressions also include community fixed effects. Robust standard errors in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A.6: **Treatment Effects Estimated Using Probit Regression**

VARIABLES	(1) Primary SD Indicator	(2) Primary SD Indicator	(3) Perceived share of households in community that will get sick from Covid	(4) Perceived share of households in community that will get sick from Covid
T1: Social Norm Correction	0.00892	-0.0709**	0.0288	-0.401***
	(0.0212)	(0.0347)	(0.0390)	(0.132)
T2: Leader Endorsement	-0.00837	-0.0356	-0.0298	-0.306**
	(0.0214)	(0.0330)	(0.0392)	(0.135)
T1 × District Covid Cases		0.00369***		0.0132***
		(0.00126)		(0.00384)
T2 × District Covid Cases		0.00139		0.00851**
		(0.00124)		(0.00396)
Observations	1,285	1,285	806	806
Control Mean DV	0.141	0.141	0.356	0.356
Control SD DV	0.349	0.349	0.368	0.368

Notes: Dependent variables are defined in Tables 1 and 2. Coefficients presented are marginal effects from probit regression. Social distancing treatments and controls are as defined in Table 2. All regressions also include community fixed effects. Robust standard errors in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

H Effect on Perceived Social Norm

Using answers to the question "For every 10 households in your community, how many support social distancing?", we estimate each respondent's perceived share of households in the community who support social distancing. We note that this measure may be an upward-based estimate of true perceptions of the social norm, since experimenter demand effects may lead respondents to report higher shares of support for social distancing in order to make their communities look favorable.

Table A.8 presents the cumulative distribution of this perceived social norm measure in the full samples at baseline and endline, and subdivided by treatment arm at endline. Even at baseline, the distribution is skewed upwards with over 90% of the sample reporting that the majority (50% or greater) of households in their community support social distancing and over half of the sample reporting that 100% of households do the same. Therefore, this measure may be limited in its ability to measure the effect of a social norm correction given that there is "little room to improve" for many respondents in the sample. If some high estimates are due to experimenter demand effects, then the social norm correction may still boost respondents' true perception of the social norm in a way not captured by our measure.

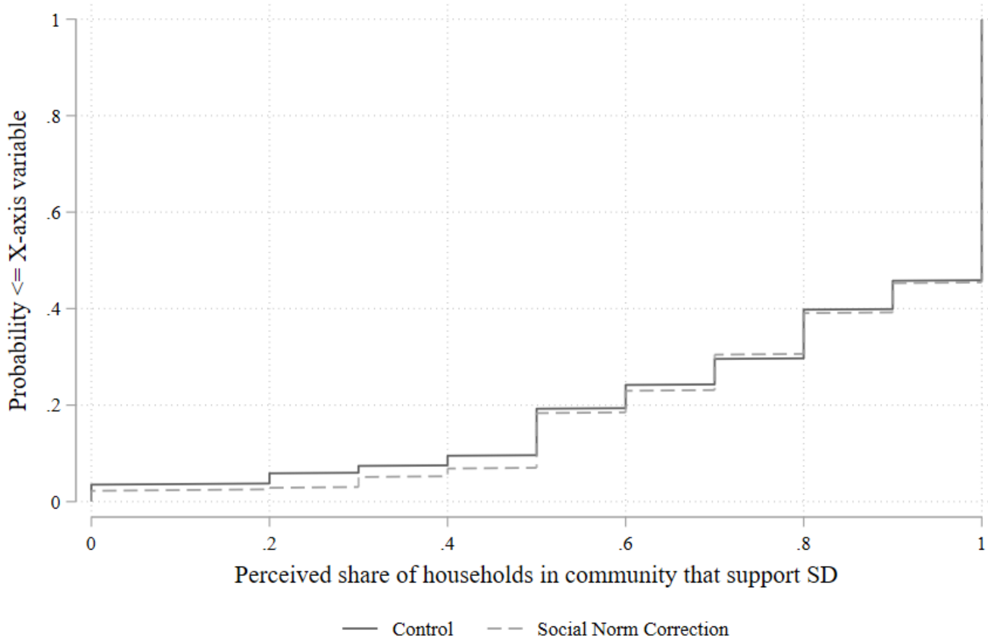
Table A.7: **Sample Distribution (Cumulative %) by Perceived Social Norm**

	Baseline	Endline			
Perceived Share	Total	Total	Control	T1	T2
0%	2.7	2.8	3.5	2.2	2.5
10%	3.1	3.1	3.6	2.4	3.0
20%	4.4	4.4	5.9	2.9	3.9
30%	6.5	6.5	7.4	5.1	6.6
40%	9.6	8.8	9.5	6.9	9.9
50%	21.1	19.0	19.3	18.3	19.2
60%	27.1	23.9	24.2	23.0	24.5
70%	33.4	30.3	29.6	30.5	31.1
80%	43.4	40.8	39.8	39.1	44.0
90%	48.9	46.8	45.8	45.3	49.6
100%	100.0	100.0	100.0	100.0	100.0

Notes: Perceived share of households in the community who support social distancing is estimated by dividing responses to the question "For every 10 households in your community, how many support social distancing?" by 10, and hence has 11 categories from 0%, 10%... 90%, 100%. Cells report cumulative percentages from 0% up to the row in question. At baseline, "Total" refers to the Round 2 responses from the whole sample. At endline, "Total" refers to Round 3 responses from the whole sample, "Control" from the control group, "T1" from the social norm correction treatment group, and "T2" from the leader endorsement treatment group.

We find that the social norm correction treatment did increase respondents' perceived social norm, particularly for those at the lower end of the distribution. Figure A.2 shows the cumulative distribution function for the perceived social norm measure at endline. Relative to the control group, those receiving the social norm correction treatment were less likely to report that fewer than 50% of households in their community supported social distancing, instead reporting higher social norms. Further, Table A.9 presents three regressions estimating the treatment effects on the perceived social norm. In column (1), the dependent variable is the perceived share of households in the community who support social distancing. The coefficient is positive and marginally statistically significant (p-value 0.12). Regressions in columns (2) and (3) find that the social norm correction treatment has a positive effect on an indicator for the respondent believing the majority (50% or more) of households in their community support social distancing, and an indicator that the respondent's perceived social norm increased between baseline and endline (both coefficients are statistically significantly different from zero at the 5% level).

Figure A.3: **Cumulative Distribution of Perceived Social Norm by Treatment**



Notes: Perceived share of households in the community who support social distancing is estimated by dividing responses to the question "For every 10 households in your community, how many support social distancing?" by 10, and hence has 11 categories from 0%, 10%... 90%, 100%. Figure depicts the cumulative distribution function of this variable for the "Control" group and "Social Norm Correction" treatment arm. The leader endorsement treatment is excluded for clarity.

Table A.8: **Treatment Effects on Perceived Social Norm (PSN)**

VARIABLES	(1) Continuous PSN	(2) Indicator if PSN $\geq$ 50%	(3) Indicator if PSN increased
T1: Social Norm Correction	0.0196	0.0291**	0.0507**
	(0.0128)	(0.0138)	(0.0241)
T2: Leader Endorsement	0.00405	0.00357	0.0358
	(0.0128)	(0.0149)	(0.0236)
Observations	2,116	2,116	2,113
R-squared	0.164	0.118	0.043
Control Mean DV	0.812	0.905	0.255
Control SD DV	0.268	0.293	0.436

Notes: Dependent variables are defined as follows. Column 1 is the perceived share of households in community that support social distancing, which takes on the values shown in Table A.7. Column 2 is an indicator equal to one if respondent reports that majority (50% or more) of households in community support social distancing, and zero otherwise. Column 3 is an indicator equal to one if the respondent's perceived social norm increased between the baseline (pre-treatment) and endline (post-treatment) surveys. "T1: Social Norm Correction" & "T2: Leader Endorsement" and controls are as defined in Table 2, except column 3 does not include a baseline value of the outcome as a control as it was used to calculate the outcome. All regressions also include community fixed effects. Robust standard errors in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

I Excluding Chimoio District

A central finding of the paper is the heterogeneity in the treatment effect of the social norm correction treatment with respect to local COVID-19 cases per 100,000 population (Column 2 of Table 2). A question that arises is whether this heterogeneity is entirely driven by the Chimoio district, which has the highest case loads in the sample by a fair margin (see Figure 2 and Appendix J). We therefore test the robustness of our findings to excluding from the sample the 524 respondents in Chimoio district (one-quarter of the sample), thereby only exploiting the more limited variation in district-level case loads across the remaining six districts.

The table below presents coefficient estimates from estimating Equations 8 and 9 in this restricted sample. First of all, Column 1 reveals that the coefficient on the social norm correction treatment is negative and statistically significant at the 10% level. Because this sample drops the district with the highest case loads, this result is consistent with theoretical predictions and previous findings that at lower case loads, the social norm correction treatment effect is more likely to be negative.

In Column 2, where we test for heterogeneity in the treatment effect, results are quite similar to the findings in Column 2 of Table 2 in the main text. The T1 main effect and interaction term coefficients are of similar magnitudes to those in Column 2 of Table 2, and maintain statistical significance at conventional levels (the T1 interaction term coefficient is now significant at the 10% instead of 5% level).

In sum, our central findings regarding heterogeneity in the treatment effect of the social norm correction treatment are robust to excluding from the sample respondents from the district (Chimoio) with the highest COVID-19 case loads.

Table A.9: **Treatment Effects on Social Distancing, Excluding Chimoio District**

	(1)	(2)
VARIABLES	Primary SD Indicator	Primary SD Indicator
T1: Social Norm Correction	-0.0237*	-0.0410**
	(0.0131)	(0.0194)
T2: Leader Endorsement	-0.0150	-0.0263
	(0.0141)	(0.0208)
T1 $\times$ District Covid Cases		0.00186*
		(0.000989)
T2 $\times$ District Covid Cases		0.00120
		(0.00103)
Observations	1,593	1,593
R-squared	0.141	0.142
Control Mean DV	0.0710	0.0710
Control SD DV	0.257	0.257

Notes: Regressions exclude 524 respondents from Chimoio district. Dependent variable in Columns 1-2 defined in Table 1. "T1: Social Norm Correction" is equal to one if respondent was randomly assigned to the social norm correction treatment, and zero otherwise. "T2: Leader Endorsement" is equal to one if respondent was randomly assigned to the leader endorsement treatment, and zero otherwise. "T1 x District Covid Cases" & "T2 x District Covid Cases" are the respective treatment indicators interacted with district-level cumulative COVID-19 cases per 100,000 population at the start of the endline survey (as detailed in Appendix J). All regressions control for a baseline measure of the dependent variable, a vector of indicators for number of community leaders knowing the respondent at baseline (0 through 4), and a vector of indicators for number of other respondents knowing the respondent at baseline (0 through 8). All regressions also include community fixed effects. Robust standard errors in parentheses. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

J COVID-19 Cases

Data on district-level population come from Mozambique's 2017 Census (National Institute of Statistics (INE), 2017). District COVID-19 case counts come from the government's COVID-19 Mozambique dashboard (Ministry of Health, 2020) and correspondence with provincial health offices. Each district's case count is as of the start date of the endline survey in the district (ranging from October 5 to November 1, 2020). We also show the number of respondents in our study sample in each district.

Table A.10: **Covid Cases by District**

	(1)	(2)	(3)	(4)
	Cumulative Covid	Cases per	Population	Number of Study
DISTRICT	Cases	100,000 people		Respondents
Sofala Province				
Dondo	8	4.136	193,382	323
Nhamatanda	12	4.299	279,081	214
Manica Province				
Gondola	3	3.553	84,429	224
Chimoio	142	39.082	363,336	524
Manica	20	9.292	215,239	290
Zambezia Province				
Namacurra	4	1.652	242,126	244
Nicoadala	52	28.779	180,686	298