OXFORD

Religious Leaders' Compliance with State Authority: Experimental Evidence from COVID-19 in Pakistan

Kate Vyborny 🆔

Abstract

A randomized controlled trial in Pakistan tests whether one-on-one engagement with community religious leaders can encourage them to instruct congregants to follow government regulations. Treated religious leaders are 25 percent more likely to comply with government requirements to tell congregants they should wear a mask to prevent COVID transmission when attending prayers. Treatment effects do not depend on the religious content of the message. Effects are driven by respondents who already understand the mechanics of COVID transmission at baseline, suggesting the treatment does not work by correcting basic knowledge about the disease, but rather through a mechanism of persuasion.

JEL classification: I18, Z12

Keywords: State capacity, state authority, religious institutions, imams, RCT, Pakistan

Weak state capacity is a central challenge for economic development ([Wade 1990](#); [Besley and Persson 2009](#); [Acemoglu, Ticchi, and Vindigni 2011](#)). Among other challenges, states in developing countries often have limited ability to enforce laws and regulations to collect resources, maintain public safety, or protect public health. Regulations promulgated by the state may be challenged by informal or non-state leaders, such as ethnic, tribal, or religious leaders. Thus, states must often engage with non-state leaders, or contend with their opposition. Yet there is surprisingly limited evidence addressing how non-state leaders decide whether to comply with or oppose state regulation.

This study investigates whether and how local informal leaders can be persuaded to comply, and encourage their followers to comply, with state institutions. Specifically, it examines local religious leaders' compliance with government mandates in the context of the COVID-19 pandemic, a situation which has made painfully clear the limited ability of state capacity to directly enforce regulations such as masking.

A randomized controlled trial is used to test whether local religious leaders can be persuaded to follow a government mandate to ask followers to follow such measures when attending congregational prayer. The intervention tested is low cost and light touch: a one-on-one, interactive persuasive telephone call, focusing on the participants' importance as community leaders in guiding followers and protecting vulnerable members of the congregation.

The outcome of interest is a change in religious leaders' behavior in asking congregants to comply with government mandates for masking and other preventative measures while attending prayer. In separate "mystery-shopper" calls, a prospective congregant calls the leader to ask about service timings and requirements for masking. The religious leaders do not know that the caller is associated with the study. This approach has a substantial advantage over survey-based measures, in that behaviors may correlate only weakly to stated attitudes (Druckman 2021; Lau 2020), and the treatment itself could have induced differential social desirability bias in survey questions on beliefs or behaviors; recent evidence points to substantial over-reporting of COVID-19 prevention measures, including by religious leaders in the context of Pakistan (Grépin, Mueller, and Wu 2023; Pattan Development Organization 2020). The intervention leads to a 25 percent increase in the proportion of respondents who instruct congregants to comply with the government mandates when attending prayer.

Variation in baseline knowledge about the basics of COVID transmission allows one to test whether the communication treatment takes effect through an *information/education* channel or a *persuasion* channel (Lau 2020). The results are stronger among those who answered baseline questions about COVID transmission *correctly*. This is not consistent with a pure information updating effect. Rather, this result suggests that the interactive conversation has an effect through a persuasion or commitment mechanism (Lau 2020; Druckman 2021; Dellavigna and Gentzkow 2010), connecting that knowledge to respondents' sense of identity as leaders and motivation to protect the vulnerable in their community. Testing whether explicit religious framing moderates the effects on leaders' decisions sheds further light on this mechanism. Religious framing might be particularly important in contexts in which religious institutions and the state are potentially competing sources of authority. However, the results show no differential impact of adding explicit religious content to the script.

This study contributes to three main strands of literature. First, an extensive literature in economics and political science has explored influences on decisions made by leaders and officials, in particular bureaucrats and state employees (reviewed by Finan, Olken, and Pande (2015); Besley et al. (2022)) or elected officials (reviewed by Grose (2014)). Within this literature, this study contributes most closely to a nascent literature conducting information and communication experiments with bureaucrats or elected officials (Callen et al. 2020; Vivalt and Coville 2020; Banuri, Dercon, and Gauri 2019; Hjort et al. 2021; Nakajima 2021; Grundler and Potrafke 2022; Rogger and Somani 2018; Mehmood, Naseer, and Chen 2022; Lee 2022). However, these studies do not examine the decisions of non-state or informal leaders, and tend to focus on mechanisms that would not apply to such leaders. Studies of bureaucrats typically focus on extrinsic incentives applicable only to employees. Similarly, studies of elected officials focus on direct or indirect electoral incentives, which are inapplicable to informal leaders. Thus it is unclear to what extent findings from studies of state leaders might generalize to the behavior of non-state or informal leaders.

A second literature explores the role of non-state institutions in supporting or hindering the capacity of the state. This literature focuses on understanding the sources of legitimacy and power of institutions, using historical quasi-experiments (Dell, Lane, and Querubin 2018; Acemoglu and Robinson 2017; Banerjee and Iyer 2005) or using natural or field experiments to understand citizens' decisions on compliance with state versus non-state authorities or trust in state services (Acemoglu et al. 2020; Balán et al. 2022; Khan et al. 2021; Blattman, Hartman, and Blair 2014; Sandefur and Siddiqi 2015; McCauley 2014; Wilke

2023; Cavgias et al. 2023; Blair et al. 2022; Martinez-Bravo and Stegmann 2021).[1] This literature has highlighted that non-state leaders may either directly influence the political system or act as an alternative source of authority or legitimacy to that of the state, and are thus particularly important in influencing compliance with state authority. However, these studies do not explore the determinants of decisions made by informal or non-state leaders.

The third literature examines the behavior of local leaders in interacting with the state. These studies have focused on the role of such leaders as intermediaries in the electoral process (Kitschelt and Wilkinson 2007; Stokes, Dunning, and Brusco 2012; Nichter and Peress 2017; Nichter 2018; Baldwin 2013, 2014; De Kadt and Larreguy 2018; Acemoglu, Reed, and Robinson 2014), or the delivery of state or donor assistance (Basurto, Dupas, and Robinson 2020; Casey et al. 2012; Kilic, Whitney, and Winters 2015; Alatas et al. 2012, 2019; Conning and Kevane 2002; Mansuri and Rao 2004; Baldwin 2018; Casey et al. 2023; Voors et al. 2018). Quantitative researchers have not focused on the question of how informal leaders make decisions on supporting or undermining compliance with the state. In addition, this literature has focused on leaders such as landlords and village chiefs; there has been little work on decisions by religious leaders, despite a wealth of evidence on the influence of religious institutions and leaders on individuals' economic, social, and political decisions.[2] Two recent exceptions use quasi-experiments or survey experiments to study how religious leaders engage in electoral politics and rent seeking (Mehmood and Seror 2023; Smith 2016).

This study thus contributes to the existing literature by rigorously investigating how informal leaders decide whether to support or undermine the state, an area that has received limited attention in quantitative research. It also innovates by applying experimental methods to study the decision-making of religious leaders; previously such methods have been used primarily with individual citizens and to a more limited extent with politicians. By examining how targeted communication can influence informal and religious leaders in making decisions that support or undermine state capacity, this study bridges a gap between the literatures on the influences on the decisions of leaders and the relationship between the state and informal institutions.

A secondary contribution of this study relates to the literature investigating how communication may influence pro-social public-health behavior. Information campaigns often focus on inducing individuals to update their baseline beliefs about the private returns to health behaviors, even when the targeted behaviors have large positive spillovers: convincing individuals to wash their hands (Bennett, Naqvi, and Schmidt 2018), use a bednet to prevent the spread of malaria (Rhee et al. 2005), have themselves or their children vaccinated (Nyhan et al. 2014; Alsan and Eichmeyer 2023), or avoid risky sexual behavior (de Walque 2007; Kerwin 2020; Dupas 2011; Duflo, Dupas, and Kremer 2015). However, approaches emphasizing private returns may face limitations in situations where the positive spillover of a health

---

1 Some of these studies examine interventions by the state or non-state actors, and examine citizen behavior in response (Cavgias et al. 2023; Sandefur and Siddiqi 2015; Balán et al. 2022; Martinez-Bravo and Stegmann 2021; Blattman, Hartman, and Blair 2014), while others use lab or survey experiments to elicit citizens' preferences over state versus non-state institutions (Blair et al. 2022; McCauley 2014). This study is methodologically most closely related to Acemoglu et al. (2020) and Khan et al. (2021); in both these studies a research team member provides information to participants in a survey or lab-in-field experiment about state effectiveness, and examines shifts in survey responses or behavior in incentivized choices. However, this study differs from these studies in that it tests the role of persuasion as distinct from information.

2 Economists and political scientists have demonstrated this in a range of settings across developed and developing countries (Auriol et al. 2020; Bassi and Rasul 2017; Bryan, Choi, and Karlan 2020; Murphy, Nourani, and Lee 2020; Gruber 2005; Moreno-Medina 2021; Torgler 2006; Boyer et al. 2022; McClendon and Riedl 2019, 2015; Mcclendon and Riedl 2016; McClendon 2019; McClendon and Riedl 2021; Blair et al. 2021; Clingingsmith, Khwaja, and Kremer 2009; Rahman 2021; Sharma 2018; Sperber, McClendon, and Kaaba 2021; Chhibber and Sekhon 2014; Condra, Isaqzadeh, and Linardi 2019; Hsiung and Djupe 2019; Freedman 2020; Hong and Paik 2021; Ben-Nun Bloom, Arikan, and Courtemanche 2015; Rink 2018).

behavior is large relative to the private returns, such as mask use by young, healthy people in the case of COVID, or the take-up of childhood vaccinations. Appealing to pro-social motivations may have potential in such situations, particularly given recent evidence on the importance of pro-social motivation in incentivizing health sector workers (Ashraf, Bandiera, and Jack 2014; Deserranno 2018; Khan 2020) and the relevance of social signalling concerns in motivating vaccination take-up (Karing 2021). However, few studies have examined health information treatments that appeal to respondents' pro-social motivations in changing their behavior, and several recent studies have found no detectable effects of such variations in information delivered to the general public (Guiteras et al. 2016; Banerjee et al. 2020; Khan et al. 2021; Dai et al. 2021). Abaluck et al. (2021) report the results of a package of interventions to reduce COVID transmission in Bangladesh. They find no effect of cross-randomized household-level variation in altruism-based or self-protection-based promotion.[3] The message in the intervention in the present study has a strong pro-social focus: respondents are asked in their capacity as leaders to take safety measures in the mosque to protect the elderly and vulnerable in their congregations, and to influence others to do the same. The strong response to this treatment suggests the potential of pro-social messaging in information campaigns for public-health behaviors with large positive spillovers, and those targeted to community leaders, who may be positively selected for pro-social motivation, or feel that communities expect them to take greater responsibility for protecting members as part of their role.

The remainder of the paper proceeds as follows. The first section describes the context. The second section details the experimental design. The third section presents results, and the final section concludes.

1. Setting

Pakistan is a setting with a low degree of trust in secular authorities; in the 2012 World Values Survey, 60 percent of Pakistanis reported that they feel little or no confidence in the government (World Values Survey 2012). In national polls carried out during the COVID-19 pandemic, 50 percent agreed that COVID-19 is a foreign conspiracy (Gallup Pakistan 2020).

Religious institutions, on the other hand, enjoy a high degree of trust. The vast majority of the population (96 percent) are Muslim (Pakistan Bureau of Statistics 2017). Religiosity is among the highest in the world, with 94 percent saying religion is "very important" in their lives (Pew Research Center 2018). About 80 percent of respondents in the World Values Survey agree that it is an essential characteristic of democracy for religious authorities to interpret the laws (World Values Survey 2012). Attending Friday prayers in congregation at the mosque is generally considered obligatory for men, and this involves close contact: the Sunnah (example of the Prophet) is to stand shoulder to shoulder during the prayer.

Islam has no central religious authority; any cleric with a certain level of legal qualification (a mufti) can issue a fatwa, or Islamic legal opinion, and many clerics who are not officially muftis do so as well; thus, there may be many contradictory opinions issued by authoritative leaders on any given issue, and these are not considered binding. Thus, local religious leaders in Islam have substantial discretion to influence their congregants' practices.

In March 2020, with COVID-19 cases rising, the government announced a nationwide lockdown, including the suspension of congregational prayers. However, there was limited compliance with this rule. Nationally influential clerics announced opposition to the rule, and in some cases police attempting to enforce it clashed with worshippers outside mosques. Government and a group of influential clerics at the national level met for a series of negotiations and then announced a joint plan. Mosques would stay open but would follow 20 key rules to reduce the spread of COVID. The most clearly defined of these were as follows: (a) prayer mats should be removed and the floor should be washed with chlorinated water, (b)

---

3   Abaluck et al. (2021) also added a focused religious and community leader engagement protocol as part of their package of interventions, part way through the experimental rollout; however, they do not test the impact of engagement with religious leaders separate from other approaches, or investigate the responses of religious leaders.

Table 1. Experimental Design

	Baseline questions	Health information	Government persuasion	Religious persuasion	Mystery shoppers	N individuals who responded to		
						Baseline + treatment call	Mystery shopper	Both
Secular T	Yes	Yes	Yes	No	Yes	283	213	213
Secular + religious T	Yes	Yes	Yes	Yes	Yes	261	185	185
Control	Yes	No	No	No	Yes	285	511	211
Super control	No	No	No	No	Yes	NA	631	NA
Total						829	1,540	609

Note: In the treatment arms, respondents were only attempted for the mystery shopper if they responded to the baseline. For the control and super-control arms, all respondents were attempted for the mystery shopper. Thus, the main analysis compares treatment and control for the sample in column 8 (Sample 1): those who responded to both the baseline and mystery shopper ($N = 211 + 213 + 185 = 609$). Comparisons between control and super control, use column 7 (sample 2): those who responded to the mystery shopper ($N = 511 + 631 = 1, 142$). Figure S1.2 provides more details on the sample composition and reasons for exclusion at each stage of the experiment.

people over 50 years of age and children should not be allowed to attend, (c) six feet of distance during congregational prayers, (d) people should perform ablution at home, and (e) congregants should wear face masks. However, implementation of these rules was limited, with NGOs reporting in May that 80 percent of mosques were not following these rules (Pattan Development Organization 2020); enforcement by state authorities was practically non-existent.

After the first major religious holiday of the year, Eid ul Fitr in May 2020, cases climbed faster (fig. S1.1); many attributed this rise in part to religious and social gatherings on the holiday, including congregational prayers in mosques (Deutsche Welle 2020). In the period leading up to Eid ul Azha, the second major religious holiday of the year, at the end of July, policymakers were concerned that cases could spike again due to large-scale gatherings over the holiday. Self-reported weekly mosque attendance continued to climb after a low during lockdown, reaching 77 percent by the time of Eid ul Azha (Gallup Pakistan 2020).

The randomized control trial (RCT) took place over a three-week period in July 2020, leading up to the Eid ul Azha holiday. The treatments and data collection reference this holiday. Comparing pilot survey data collected in April 2020 to the experiment baseline survey in July 2020 (table S1.1) reveals that by the time of the experiment, respondents were more likely to report asking congregants to wear masks, and less likely to report other steps, such as removing communal prayer mats or encouraging congregants to perform ablution at home. This likely reflects the shift in public-health messaging towards an increased focus on masking given updated knowledge about COVID transmission mechanisms.

2. Experimental Design

The RCT sample is a random draw of religious leaders from community mosques across 19 districts of urban and rural Punjab from a government listing of individuals licensed to register marriages, provided by the Government of Punjab as part of a separate experiment (Field and Vyborny 2020). Approximately two-thirds of individuals in this sample frame are local religious leaders (who typically conduct a marriage and then register it); for this RCT, the randomization sample is restricted to this group. These licenses are issued once, with no official qualifications or training; there is no process for expiry or renewal of these licenses, and the respondents do not receive any government compensation.

Table 1 gives an overview of the experimental design, and fig. S1.2 shows the steps and sample composition in more detail. Respondents were randomized individually into one of four experimental conditions. Randomization was stratified by subdistrict, rural/urban location, and education level. Baseline phone calls were carried out in a randomized order within a fixed time frame available for field work

(fig. S1.2, box 1); thus about half of the religious leaders originally randomized into each treatment arm were actually attempted for the baseline and become part of the main sample.

In the baseline phone call, the enumerator identifies himself as a researcher from the Center for Economic Research in Pakistan (CERP). He emphasizes that he is not calling from the government, but rather as part of a research study. This framing was used both to avoid deception and also to help elicit truthful responses to the survey questions and avoid concerns that the call might lead to any punishment related to enforcement of COVID regulations. Enumerator debriefing revealed that most respondents seemed assured that the caller was not from the government; the handful who seemed unconvinced of this point declined to participate, apparently due to concerns that the call related to enforcement. CERP is well known in policy circles, but is not generally familiar to the public; as a result, respondents were not familiar with CERP, and CERP's own reputation did not play a role. Thus the treatment represents a potentially neutral researcher or non-state actor engaging with the respondent, similar to Acemoglu et al. (2020) and Khan et al. (2021).

The caller then confirms the identity of the respondent and his role as a religious leader. Non-response to this call was high, primarily because the government listing was several years old and had many wrong and outdated numbers. Crucially, there was no difference in protocol or information provided to respondents between treatment arms until the end of this call (fig. S1.2, box 2), so differential non-response by treatment arm is not possible in this stage of the design. Thus non-response at this stage is a form of selection into the RCT sample, rather than attrition. Notably, this selection into the study is not correlated with the respondents' education or urban/rural location (table S1.2, column 1).

For the religious leaders who are reached and agree to be surveyed, the enumerator carries out the baseline survey, collecting information on baseline knowledge about COVID and existing steps taken to prevent spread in the mosque. For the two treatment groups, the persuasion script follows immediately after the survey (fig. S1.2, box 2). Both treatments (secular persuasion; secular + religious persuasion) emphasize the importance of the respondent's leadership role in the community and appeal to him in protecting vulnerable community members from COVID. In addition, they both emphasize the key actionable points in the official protocols for mosques, and ask the respondent to follow them and to spread the word to his community through his sermons and mosque loudspeaker announcements. The script is framed as a persuasive appeal by a non-government organization, with no reference to enforcement measures. It is interactive, involving frequent elicitation of the respondent's reactions and agreement, as well as asking him to commit to action. The interactive nature of the messaging is a key difference from mass-media messages about COVID, which were widely disseminated during this period.

In addition to these elements, the secular + religious persuasion treatment arm includes an appeal to religious authority. This includes (a) the fact that the top religious leaders have endorsed the protocols for mosques, (b) hadith (sayings of the Prophet) about avoiding spread of plague, (c) international Sunni and Shia authorities' pronouncements (fatwas) on the importance of complying with official authorities to prevent spread of COVID, (d) examples of other Muslim countries following strong measures to prevent COVID spread. These components are all designed to convey endorsement from a higher religious authority than the respondent. Because the outcome of interest is compliance with government regulations, the RCT design incorporates a combined arm adding religious framing to the script concerning these regulations, rather than using a separate religious-only script. The scripts of the treatments are detailed in supplementary online appendix S2.

The main followup data collection uses mystery shoppers to obtain a measure of the religious leaders' instructions to members of the mosque congregation free of social desirability bias. This approach has been used extensively in recent development economics research to observe behavior of agents in contexts where self-reports are expected to be affected by social desirability or other forms of bias (Gaddis 2018; Banerjee et al. 2021; Anagol, Cole, and Sarkar 2017; Gine and Mazer 2016; Bertrand et al. 2007; Dizon-Ross, Dupas, and Robinson 2017; Hetzel et al. 2008; Mohanan et al. 2015; Currie, Lin, and Zhang 2011;

Friedman, Woodman, and Chatterji 2015; Fitzpatrick 2021; Bate, Mooney, and Hess 2010; Bate, Jin, and Mathur 2011; Das et al. 2012; Field and Vyborny 2020; Broockman 2013, 2014; McClendon 2016; Kalla and Broockman 2016; Mendez and Grose 2018; Driscoll et al. 2018; Gaikwad and Nellis 2021).

A different enumerator from the original surveyor calls each respondent in the days before Eid, posing as a member of the community saying he and his father want to confirm the timing for Eid prayer services at the mosque. Timings for services are usually set one to two days before Eid and differ for each mosque and each year; thus community members must enquire to confirm the time. These calls were credible to respondents; a number of respondents later called back to let the mystery-shopper callers know the confirmed time of prayers. Response to this call is not correlated with treatment (table S1.3). It is also uncorrelated with respondents' education, although rural respondents are more likely to answer the mystery-shopper call (table S1.2, column 2). Mystery-shopper callers were blinded to the treatment status of the respondent. Enumerators also recorded whether the respondent seemed suspicious of the mystery-shopper call; approximately 10 percent of calls were tagged as suspicious. Treatment does not affect whether respondents seemed suspicious, nor are there heterogeneous effects on suspicion by baseline knowledge of COVID or congregation size (table S1.4). The scripts of the mystery-shopper calls are included in supplementary online appendix S2. Supplementary online appendix S3 discusses ethics aspects of the mystery-shopper call.

After asking about the time of prayer services, the caller asks several questions about how he should prepare for attending mosque given COVID conditions, e.g. by wearing a mask, bringing his own prayer mat, or doing ablution at home. This is credible during the Eid prayers as it is a larger gathering than usual, and congregants who do not frequently attend the mosque are likely to attend. The primary outcome is a simple mean index of the following indicators: (a) whether the respondent advises wearing a mask, (b) says wearing a mask is required when the caller says he would prefer not to wear it, (c) tells the caller to bring a prayer mat (i.e. because the mosque mats would have been removed), (d) indicates the caller should do ablution at home, and (e) asks about the caller's father's age or indicates that the elderly father should not attend. Results tables also report effects on each of the five individual components of the index.

For the primary analysis, enumerators make mystery-shopper calls with respondents who responded to the baseline survey in treatment and control arms (fig. S1.2, box 3, blue and green nodes); this is denoted as Sample 1. Of the 819 respondents who were surveyed successfully in the treatment and control groups and called by mystery shoppers, 629 (75 percent) answered the mystery-shopper calls, of whom 609 were planning to hold Eid prayers at their mosque and were asked questions about prayers;[4] this is not differential by treatment arm (table S1.3, column 1). To quantify the main treatment effect of interest, the impact of persuasion, Specification (1) compares mystery-shopper outcomes between treatment and control arms for these 609 respondents reached at baseline (Sample 1, shown in blue and green in fig. S1.2):

$$Y_i = \beta_0 + \beta_1 \text{ANYTREAT}_i + \epsilon_i. \tag{1}$$

To test whether religious messaging moderates the treatment effects, an alternative specification breaks down the treatment into its two arms:[5]

$$Y_i = \beta_0 + \beta_1 \text{SECULARTREAT}_i + \beta_2 \text{COMBINEDTREAT}_i + \epsilon_i. \tag{2}$$

To distinguish information and persuasion, the next analysis divides the sample by whether the respondent correctly answered two questions about COVID at baseline: whether it can be transmitted by

---

4 Smaller mosques do not always hold Eid prayers.

5 Note that the estimation sample here includes the secular treatment, the combined treatment, and the control group. The latter group is the omitted category in equation (2). These categories are mutually exclusive and exhaustive (i.e. the experiment does not use a factorial design)—see table 1—thus the issue of interaction terms for a factorial design does not arise.

people who show no symptoms, and whether it can be transmitted through coughing or sneezing without touching. Understanding these two facts is crucial, as it provides the basis for the need for preventive measures even among apparently healthy congregants. Tt baseline,both questions were answered correctly by 69 percent of respondents. Equation (3) interacts the main estimates with the respondent's baseline knowledge of how COVID is transmitted:

$$Y_i = \beta_0 + \beta_1 \text{ANYTREAT}_i + \beta_2 \text{BASELINERIGHT}_i$$
$$+ \beta_3 \text{ANYTREAT}_i \times \text{BASELINERIGHT}_i + \epsilon_i, \tag{3}$$

where BASELINERIGHT is an indicator for whether the respondent answered both baseline knowledge questions correctly. If information transmission is the key mechanism, this implies $\beta_3 < 0$ (there is a smaller impact on respondents who already understood COVID transmission at baseline). In contrast, if persuasion is the key mechanism, this implies $\beta_3 \geq 0$ (there is an equal or greater impact on respondents who understood COVID transmission at baseline). To differentiate between respondents who are uncertain versus those who are convinced of misinformation, the sample is divided into three groups based on a Likert scale measure of the respondent's certainty in his answer: those who are certain and correct, uncertain, or certain and wrong. An additional heterogeneity specification adapts equation (3), dividing respondents into these groups, as an alternative approach to distinguishing these mechanisms.

It is possible that simply answering the baseline survey questions plays a role in making COVID salient to religious leaders and increasing their likelihood of recommending preventive measures. To test this, a secondary sample (Sample 2) facilitates a test for the effect of the control group call on behavior. Sample 2 is constructed by calling respondents for the mystery-shopper call from a "super-control" group who were not contacted for baseline calls. To ensure an equivalent sample in both control and super-control groups in Sample 2, enumerators conduct mystery-shopper calls with control-group respondents regardless of baseline response status. Thus this secondary analysis compares mystery-shopper responses between control ($N = 511$) and super-control respondents ($N = 631$) without any sample restriction on baseline response (fig. S1.2, box 3, yellow and green nodes). Response rates to the mystery-shopper calls do not differ between the control and super-control arms (table S1.3, column 2). Equation (4) simply compares the control to the super-control arm:

$$Y_i = \beta_0 + \beta_1 CONTROL_i + \epsilon_i. \tag{4}$$

The scripts of the treatments and mystery-shopper data collection are detailed in supplementary online appendix S2. The study design and outcomes were pre-registered (AEARCTR-0005740, Version 2.0); supplementary online appendix S4 discusses changes in the analysis from the pre-registration. Replication data for this study are available from the Harvard dataverse at https://doi.org/10.7910/DVN/CV1KCJ.

3. Results

3.1. Main Results

Table 2 shows descriptive statistics and balance for the main sample from the baseline survey. Respondents lead community mosques with an average of 40 people attending daily evening prayer before COVID. Most respondents have less than 10th grade (Matric) education and some training in madrasa (religious schooling). The majority report that they are receiving frequent messages about COVID. However, the majority do not believe COVID is present in their communities. Respondents reported steps they have taken to prevent spread of COVID in the mosque, without being prompted with any specific step; almost all respondents mentioned at least one step, but only about 25 percent mention masks. Respondents answered two basic questions about COVID transmission: whether it can be spread by people who show no symptoms, and whether it can be spread through coughing even if two people do not touch. About

Table 2. Descriptive Statistics and Balance

Variable	Combined secular + religious (1)		Secular only (2)		Control (3)		Pairwise t-test (1) − (2)		(1) − (3)		(2) − (3)	
	N	Mean/(SE)	N	Mean/(SE)	N	Mean/(SE)	N	Mean difference	N	Mean difference	N	Mean difference
Ed Matric or above	185	0.465 (0.037)	213	0.493 (0.034)	211	0.431 (0.034)	398	−0.028	396	0.034	424	0.062
Attended madrasa	178	0.831 (0.028)	208	0.841 (0.025)	207	0.758 (0.030)	386	−0.010	385	0.073*	415	0.083**
Typical daily congregation pre-COVID	169	39.592 (3.330)	202	36.327 (3.008)	193	38.995 (4.453)	371	3.265	362	0.597	395	−2.668
Urban	185	0.211 (0.030)	213	0.239 (0.029)	211	0.213 (0.028)	398	−0.029	396	−0.002	424	0.026
Governing party constituency	158	0.430 (0.040)	170	0.376 (0.037)	174	0.362 (0.037)	328	0.054	332	0.068	344	0.014
Believes COVID may be present in community	173	0.116 (0.024)	202	0.084 (0.020)	199	0.186 (0.028)	375	0.031	372	−0.070*	401	−0.102***
BL COVID knowledge: Certain and correct	162	0.580 (0.039)	191	0.576 (0.036)	188	0.580 (0.036)	353	0.004	350	0.000	379	−0.004
BL COVID knowledge: Uncertain	162	0.321 (0.037)	191	0.293 (0.033)	188	0.324 (0.034)	353	0.028	350	−0.003	379	−0.031
BL COVID knowledge: Certain and wrong	162	0.099 (0.024)	191	0.131 (0.024)	188	0.096 (0.022)	353	−0.032	350	0.003	379	0.035
Received few/no COVID messages last week	168	0.333 (0.036)	200	0.350 (0.034)	195	0.318 (0.033)	368	−0.017	363	0.015	395	0.032
Baseline step: Short sermon	131	0.061 (0.021)	181	0.088 (0.021)	187	0.086 (0.021)	312	−0.027	318	−0.024	368	0.003
Baseline step: Clean mosque	131	0.313 (0.041)	181	0.359 (0.036)	187	0.289 (0.033)	312	−0.046	318	0.024	368	0.070
Baseline step: Soap	131	0.328	181	0.260	187	0.246	312	0.069	318	0.082	368	0.014

Table 2. Continued

Variable	Combined secular + religious (1) N	Combined secular + religious (1) Mean/(SE)	Secular only (2) N	Secular only (2) Mean/(SE)	Control (3) N	Control (3) Mean/(SE)	Pairwise t-test (1) − (2) N	Pairwise t-test (1) − (2) Mean difference	Pairwise t-test (1) − (3) N	Pairwise t-test (1) − (3) Mean difference	Pairwise t-test (2) − (3) N	Pairwise t-test (2) − (3) Mean difference
Baseline step: Remove mats	131	0.565 (0.043)	181	0.564 (0.037)	187	0.572 (0.036)	312	0.001	318	−0.007	368	−0.009
Baseline step: Elderly/sick	131	0.099 (0.026)	181	0.155 (0.027)	187	0.166 (0.027)	312	−0.055	318	−0.067*	368	−0.011
Baseline step: Distancing	131	0.634 (0.042)	181	0.586 (0.037)	187	0.588 (0.036)	312	0.048	318	0.045	368	−0.003
Baseline step: Announcements	131	0.115 (0.028)	181	0.133 (0.025)	187	0.128 (0.025)	312	−0.018	318	−0.014	368	0.004
Baseline step: Ablution at home	131	0.191 (0.034)	181	0.227 (0.031)	187	0.225 (0.031)	312	−0.036	318	−0.034	368	0.002
Baseline step: Mask	131	0.221 (0.036)	181	0.276 (0.033)	187	0.219 (0.030)	312	−0.055	318	0.002	368	0.057
Baseline step: Other step	131	0.107 (0.027)	181	0.171 (0.028)	187	0.150 (0.026)	312	−0.064*	318	−0.043	368	0.022
Baseline step: Number of steps	131	2.634 (0.126)	181	2.818 (0.126)	187	2.668 (0.122)	312	−0.184	318	−0.035	368	0.149
Reports no steps to prevent COVID in mosque	131	0.038 (0.017)	181	0.055 (0.017)	187	0.080 (0.020)	312	−0.017	318	−0.042	368	−0.025
Gov't trusted source	168	0.214 (0.032)	196	0.173 (0.027)	194	0.180 (0.028)	364	0.041	362	0.034	390	−0.007
Ulema trusted source	167	0.707 (0.035)	195	0.790 (0.029)	193	0.699 (0.033)	362	−0.083*	360	0.007	388	0.090**

Note: Example for table source text. BL COVID knowledge correct is an indicator for whether the respondent answered both knowledge questions about COVID transmission correctly at baseline (i.e. COVID can "definitely" or "probably" be transmitted without touching, and a person can "definitely" or "probably" have COVID if he/she does not appear visibly ill. BL COVID knowledge correct and certain is an indicator for whether the respondent answered both knowledge questions about COVID transmission correctly at baseline and indicated certainty (i.e. "definitely" responses to both questions). BL COVID knowledge wrong and certain is an indicator for whether the respondent answered either baseline knowledge question about COVID transmission incorrectly with certainty (i.e. "definitely not" answers on either or both of the questions). All other cases are classified as BL COVID knowledge uncertain. The p-values of F-test for joint significance of all variables in predicting assignment to any treatment:.1; secular only arm:.08. Sample sizes for covariates vary because of non-response to baseline questions due to some respondents hanging up before completing the baseline call. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Table 3. Impact of Persuasion Treatment on Religious Leaders' Instructions to Mystery Shopper

| | Panel A: Effect of persuasion treatment (pooled) | | | | | |
| | Index components: | | | | | |
	Index—advice to MS (1)	Bring own prayer mat (2)	Ablution at home (3)	Mask recommended (4)	Mask required (5)	Asks father's age (6)
Any treatment	0.066**	0.040	0.084**	0.109**	0.097**	−0.001
	(0.026)	(0.042)	(0.042)	(0.042)	(0.041)	(0.013)
Observations	609	609	609	609	609	609
Control group mean	0.375	0.531	0.526	0.436	0.360	0.024
	Panel B: Disaggregated by individual persuasion scripts					
	Index components:					
	Index—advice to MS (1)	Bring own prayer mat (2)	Ablution at home (3)	Mask recommended (4)	Mask required (5)	Asks father's age (6)
Secular + religious treatment	0.054*	0.010	0.085*	0.094*	0.072	0.009
	(0.030)	(0.050)	(0.050)	(0.050)	(0.049)	(0.017)
Secular treatment	0.076**	0.065	0.084*	0.123**	0.119**	−0.010
	(0.030)	(0.048)	(0.048)	(0.048)	(0.048)	(0.013)
Observations	609	609	609	609	609	609
Covariates	No	No	No	No	No	No
p-value, treatment effects equal	0.464	0.264	0.992	0.564	0.354	0.233

Note: Sample 1 (treatment and control, respondents who answered both the baseline survey and the mystery-shopper call). "Bring own prayer mat" was included in the instrument but excluded from the pre-analysis plan in error; index results are robust to the exclusion of this variable (table S4.1). Robust standard errors. $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

60 percent of respondents answered both questions correctly in the affirmative and expressed confidence in their answers, a third were unsure, and 10 percent gave a definite no to one or both questions.

Overall, the randomization is well balanced. Of 72 tests, 8 are significant at the 10 percent level or greater; the *p*-value on the joint *F*-test of the significance of all variables in predicting any treatment is 0.1, and on predicting the secular-only treatment 0.08. Respondents in the control group appear to be slightly more likely to believe that COVID is present in their community, and more likely to report discouraging the elderly or sick from attending mosque during the pandemic. Both of these apparent imbalances should not drive the treatment effects of interest; if anything, they should bias the estimates towards zero.

Table 3 shows the main results of the experiment. In the control group, respondents recommended on average 38 percent of the counter-COVID measures to callers; about half recommended the caller bring his own prayer mat and do ablution at home, 44 percent recommended a mask, and only 36 percent said a mask was required. Only 2 percent of respondents asked the caller (unprompted) about his father's age. Respondents may not have imagined the callers' fathers to be elderly (most enumerators are in their 20s and 30s). Perhaps more importantly, the religious importance of congregational prayer for men may be sufficient that religious leaders disregard this component of the official rules about COVID; in contrast, wearing a mask or doing ablution before attending still allows members of the congregation to completely fulfill their religious obligations.

Panel A shows the main treatment effects, comparing treatment and control groups (equation (1) estimated on Sample 1). Overall, the treatment increased the index of COVID compliance instructions by 18 percent (7 percentage points). The effects are driven by an increase in recommendations to do ablution at home and to wear a mask; the proportion who tell callers they are required to wear a mask increases by

25 percent over the control group mean (10 percentage points over a mean of 0.36). Overall, the treatment had a large and significant impact on respondents' observed behavior.

Panel B uses the same sample to break down the results by treatment arm, to investigate whether religious persuasion has any additional effect with religious leaders. The two treatment effects on the index are similar in size ($\hat{\beta}_1 = 0.054$; $\hat{\beta}_2 = 0.076$ and statistically indistinguishable, $p = 0.46$). The treatment is equally effective with or without religious content.

3.2. Information as a Mechanism

Table 4 estimates equation (3) to investigate whether the effect of the treatment works through a mechanism of new information about COVID transmission, rather than persuasion. The sample is divided by whether the respondent correctly answered two questions about COVID at baseline: whether it can be transmitted by people who show no symptoms, and whether it can be transmitted through coughing or sneezing without touching. In panel A, respondents who gave the correct answers but were uncertain are classified as having correct knowledge at baseline. These groups are interacted with the treatment indicator.

The effects are completely driven by respondents who gave the correct answers at baseline; treatment effects on the two groups are statistically significantly different at the 5 percent level. To investigate the role of the treatment in resolving uncertainty, panel B splits respondents into three groups: those who answered the knowledge questions correctly at baseline, those who were uncertain, and those who gave the wrong answer but said they were certain about it. The effects are driven completely by those who are correct and certain in their beliefs at baseline.[6]

As in any heterogeneous treatment effects analysis, it is possible that other factors, such as education, could be correlated with respondents' knowledge about COVID and could themselves mediate responsiveness to treatment. Table S1.5 tests for robustness of the results in table 4 to this possibility. Panel A shows the correlation between respondents' characteristics and correct answers to the baseline COVID questions. More-educated respondents are more likely to answer these questions correctly; respondents who report receiving fewer messages about COVID are less likely to answer correctly. Panel B shows the results of equation (3), in which the dependent variable is the index of advice to the mystery shopper; column 1 repeats the original specification, while columns 2–4 also include these respondent baseline characteristics and their interaction with treatment as additional control variables. If the heterogeneity results in table 4 were driven by differences in these other characteristics, the coefficient on treatment would be expected to change. Instead, it is similar in size across specifications, suggesting that these differences do not drive the heterogeneity results.

3.3. Testing for Survey or Reminder Effects

Table S1.6 uses Sample 2 to investigate whether simply receiving the baseline survey, which has no informational content but makes COVID salient to respondents or may remind them of existing knowledge about government rules, drives the results (equation (4)). This specification compares the responses between the control (baseline survey + mystery shopper) and super-control group (mystery shopper only). Administering a survey alone has no detectable effect on the index of instructions to the mystery shopper.

As an additional test for the possibility of reminder effects as an alternative mechanism, fig. S1.3 tests whether treatment effects decay in the short run. The experiment was carried out over a period of several

---

6   Treatment group respondents were also asked as a part of the treatment script whether they were already aware of the 20-point plan for mosque COVID compliance measures (see supplementary online appendix S2.1); 97 percent indicated they were aware of the plan. This supports the results of the formal heterogeneity analysis, suggesting that information about the COVID compliance rules themselves is an unlikely channel for treatment effects; however, this question could be subject to social desirability bias effects, and did not test respondents' knowledge, as it was designed primarily to encourage engagement with the treatment script.

Table 4. Heterogeneous Effects by Baseline Knowledge about COVID Transmission

	Index—advice to MS	Bring own prayer mat	Ablution at home	Mask recommended	Mask required	Asks father's age
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A			
Any treatment = 1	−0.022	−0.056	0.015	0.014	−0.041	−0.044
	(0.050)	(0.080)	(0.079)	(0.080)	(0.078)	(0.034)
BL COVID knowledge correct = 1	−0.044	−0.040	−0.048	−0.040	−0.040	−0.056*
	(0.049)	(0.077)	(0.077)	(0.077)	(0.075)	(0.032)
Any treatment = 1 × BL COVID knowledge correct = 1	0.131**	0.149	0.106	0.139	0.198**	0.064*
	(0.060)	(0.097)	(0.096)	(0.097)	(0.094)	(0.036)
Observations	544	544	544	544	544	544
Control mean Y | baseline correct	0.370	0.520	0.520	0.440	0.360	0.010
Control mean Y | baseline wrong	0.410	0.560	0.570	0.480	0.400	0.060
Proportion of sample correct	0.690	0.690	0.690	0.690	0.690	0.690
p-value: Total effect on subgroup = 0	0.001	0.085	0.025	0.005	0.003	0.129

Table 4. Continued

	Panel B					
	Index—advice to MS (1)	Bring own prayer mat (2)	Ablution at home (3)	Mask recommended (4)	Mask required (5)	Asks father's age (6)
Any treatment = 1	0.012	0.033	−0.007	0.050	0.003	−0.021
	(0.050)	(0.080)	(0.079)	(0.080)	(0.078)	(0.032)
BL COVID knowledge correct and certain	−0.033	−0.027	−0.086	−0.000	−0.010	−0.040
	(0.050)	(0.080)	(0.080)	(0.080)	(0.078)	(0.029)
BL COVID knowledge wrong and certain	−0.003	0.070	−0.035	−0.070	0.012	0.006
	(0.089)	(0.132)	(0.134)	(0.132)	(0.131)	(0.061)
Any treatment = 1 × BL COVID knowledge correct and certain	0.105*	0.051	0.149	0.104	0.179*	0.042
	(0.061)	(0.100)	(0.099)	(0.100)	(0.097)	(0.035)
Any treatment = 1 × BL COVID knowledge wrong and certain	−0.026	−0.156	0.110	0.024	−0.074	−0.034
	(0.105)	(0.161)	(0.160)	(0.161)	(0.157)	(0.063)
Observations	541	541	541	541	541	541
Control mean Y | certain and correct	0.371	0.514	0.505	0.459	0.367	0.009
Control mean Y | uncertain	0.403	0.541	0.590	0.459	0.377	0.049
Control mean Y | certain and wrong	0.400	0.611	0.556	0.389	0.389	0.056
Proportion of sample certain and correct	0.579	0.579	0.579	0.579	0.579	0.579
Proportion of sample uncertain	0.312	0.312	0.312	0.312	0.312	0.312
Proportion of sample certain and wrong	0.109	0.109	0.109	0.109	0.109	0.109
p-value: Total effect on subgroup BL correct and certain = 0	0.001	0.155	0.016	0.009	0.002	0.179
p-value: Total effect on subgroup certain and wrong = 0	0.875	0.378	0.460	0.594	0.600	0.307

Note: Sample 1 (treatment and control arms, respondents who answered baseline survey and mystery-shopper call). Sample size varies from table 3 because of non-response to baseline questions due to some respondents hanging up before completing the baseline call. BL COVID knowledge correct is an indicator for whether the respondent answered both knowledge questions about COVID transmission correctly at baseline (i.e. COVID can "definitely" or "probably" be transmitted without touching, and a person can "definitely" or "probably" have COVID if he/she does not appear visibly ill. BL COVID knowledge correct and certain is an indicator for whether the respondent answered both knowledge questions about COVID transmission correctly at baseline and indicated certainty (i.e. "definitely" responses to both questions). BL COVID knowledge wrong and certain is an indicator for whether the respondent answered either baseline knowledge question about COVID transmission incorrectly with certainty (i.e. "definitely not" answers on either or both of the questions). Robust standard errors in parentheses. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

weeks, and call order was randomized; this allows us to test for a decay in effects over time. The estimates are imprecise, but do not suggest a decay in treatment effects.

Simply providing basic information about COVID transmission or reminding respondents about government regulations does not appear to be the mechanism for the results. Rather, the one-on-one persuasion, which made salient the importance of the mosque as a venue for transmission and emphasized the respondents' key leadership role in protecting the vulnerable in their community, mobilizes respondents who already believe that there is a risk of COVID spread to respond.

3.4. Heterogeneity by Relationship with Secular and Religious Authorities

Table S1.7 uses baseline survey information as well as geographic variation in the sample to test whether there is heterogeneity in responses by the respondent's relationship with secular and religious authorities. The treatment dummy is interacted with a series of indicators of these relationships:

$$Y_i = \beta_0 + \beta_1 \text{TREAT}_i + \beta_2 X_i + \beta_3 \text{TREAT}_i \times X_i + \epsilon_i.$$

There is no pattern of stronger response among respondents with a closer relationship with secular or religious authorities, whether measured by respondents' survey responses on sources of information they trust, the political alignment of the constituencies where they live, or madrasa training. This result contrasts with recent evidence from the United States and Brazil on the politicization of responses to COVID-19 public-health advice (Grossman et al. 2020; Alcott et al. 2020; Bursztyn et al. 2020; Painter and Qiu 2020; Milosh et al. 2021; Gadarian, Goodman, and Pepinsky 2021). This could be because these patterns are particular to the high degree of polarization in those countries (Barari et al. (2020) find no relationship between reported COVID compliance and trust in government across Italian respondents). Alternatively, it may be because personal interaction is more effective than mass communications in crossing party lines, as consistently shown in the political science literature on voter mobilization (Gerber and Green (2019) review this literature in detail).

4. Discussion

In this study, a randomized controlled trial establishes that one-on-one interactive communication can be effective in influencing community religious leaders to instruct congregants to follow government-mandated public-health measures in their mosques. The evidence is consistent with a *persuasion* channel, rather than an *information* or education channel, for these effects. Perhaps surprisingly, there is no evidence that an extensive script linking the persuasive argument to Islamic scripture and the advice of internationally respected religious authorities has any additional effect compared to the basic persuasion script designed to motivate respondents based on their identity as community leaders.

Given the limitations on the effectiveness of generic mass messaging and the infeasibility of direct citizen engagement at a population scale, engaging community and religious leaders may be necessary for many low-capacity states to effectively carry out policy. Although the intervention tested in this RCT requires time input at an individual level, its scripted nature and phone-based delivery means it is still low cost. The call center costs (call center agents, managers, and connection costs) amount to approximately $0.75 (in 2022 USD) per religious leader targeted in the list. The average of 230 congregants at Friday prayer reported by respondents in the pilot data suggests there are approximately 200,000 such community mosques nationwide. Thus, carrying out such a campaign at a national scale would involve outreach to 200,000 religious leaders for a cost of approximately $150,000 USD. This is equivalent to the cost estimate from a Pakistan-based advertising agency for a recent mass-media campaign on polio, another high-priority public-health issue, for 150 one-minute advertising spots each week for three weeks. Thus the cost of expanding this intervention at a national level would be comparable to a typical mass-media campaign approach.

Beyond this specific, standardized intervention, governments could consider establishing a mechanism for community-level engagement with local community leaders, such as district-level outreach teams. In this study setting, Pakistan, the government engages with prominent religious leaders at the national level. It is this engagement that led to the 20-point plan for mosques. Yet the lack of compliance with this plan at the community level illustrates the need for engagement at the community level. In Pakistan, there is no strategy or government body that has direct communication or engagement with religious or informal leaders. The results of this study demonstrate that such engagement can be effective.

Such approaches may be promising to explore not only for the case of public-health campaigns (such as encouraging trust in vaccination), but also for a much broader set of policies where establishing public trust is key—ranging from encouraging households to educate daughters to discouraging participation in violent conflict.

Anecdotally, the caller's identity as part of a non-government institution may have been important not only to elicit truthful information in the baseline survey, but also potentially to engage effectively in persuasion given that respondents might be concerned about speaking to a government caller about regulations with which they might not be compliant. This study cannot test this directly; future work could investigate this issue systematically by varying the framing of the identity of the caller.

Data Availability Statement

Replication data for this study are available at the https://doi.org/10.7910/DVN/CV1KCJ, the Harvard dataverse.

REFERENCES

Abaluck, J., L. H. Kwong, A. Styczynski, A. Haque, A. Kabir, E. Bates-Jeffries, E. Crawford *et al.*, 2022. "The Impact of Community Masking on COVID-19: A Cluster-Randomized Trial in Bangladesh." *Science* 375: 6577.

Acemoglu, D., A. Cheema, A. I. Khwaja, and J. A. Robinson, 2020. "Trust in State and Nonstate Actors: Evidence from Dispute Resolution in Pakistan." *Journal of Political Economy* 128(8): 3090–147.

Acemoglu, D., T. Reed, and J. A. Robinson, 2014. "Chiefs: Economic Development and Elite Control of Civil Society in Sierra Leone." *Journal of Political Economy* 122(2): 319–68.

Acemoglu, D., and J. Robinson, 2023. "Weak, Despotic, or Inclusive? How State Type Emerges from State versus Civil Society Competition." *American Political Science Review* 117(2): 407–420.

Acemoglu, D., D. Ticchi, and A. Vindigni, 2011. "Emergency and Persistence of Inefficient States." *Journal of the European Economic Association* 9(2): 177–208..

Alatas, V., A. Banerjee, R. Hanna, B. A. Olken, R. Purnamasari, and M. Wai-Poi, 2019. "Does Elite Capture Matter? Local Elites and Targeted Welfare Programs in Indonesia." *AEA Papers and Proceedings* 109: 334–39.

Alatas, V., A. Banerjee, R. Hanna, B. A. Olken, and J. Tobias, 2012. "Targeting the Poor: Evidence from a Field Experiment in Indonesia." *American Economic Review* 102(4): 1206–40.

Alcott, H., L. Boxell, J. Conway, M. Gentzkow, M. Thaler, and D. Yang, 2020. "Polarization and Public Health: Partisan Differences in Social Distancing during the Coronavirus Pandemic." *Journal of Public Economics* 191: 104254.

Alsan, M., and S. Eichmeyer, 2023. "Experimental Evidence on the Effectiveness of Non-experts for Improving Vaccine Demand." *American Economic Journal: Economic Policy* 16 (1): 394–414.

Anagol, S., S. Cole, and S. Sarkar, 2017. "Understanding the Advice of Commissions-Motivated Agents: Evidence from the Indian Life Insurance Market." *Review of Economics and Statistics* 99(1): 1–15.

Ashraf, N., O. Bandiera, and B. K. Jack, 2014. "No Margin, No Mission? A Field Experiment on Incentives for Public Service Delivery." *Journal of Public Economics* 120: 1–17.

Auriol, E., J. Lassebie, A. Panin, E. Raiber, and P. Seabright, 2020. "God Insures Those Who Pay? Formal Insurance and Religious Offerings in Ghana." *Quarterly Journal of Economics* 135(4): 1799–1848.

Balán, P., A. Bergeron, G. Tourek, and J. L. Weigel, 2022. "Local Elites as State Capacity: How City Chiefs Use Local Information to Increase Tax Compliance in the Democratic Republic of the Congo." *American Economic Review* 112(3): 762–97.

Baldwin, K., 2013. "Why Vote with the Chief? Political Connections and Public Goods Provision in Zambia." *American Journal of Political Science* 57(4): 794–809.

———. 2014. "When Politicans Cede Control of Resources: Land, Chiefs and Coalition-Building in Africa." *Comparative Politics*: 46(3): 253–271.

———. 2019. "Elected MPs, Traditional Chiefs, and Local Public Goods: Evidence on the Role of Leaders in Co-production from Rural Zambia." *Comparative Political Studies* 52(12): 1925–1956.

Banerjee, A., M. Alsan, E. Breza, A. Chandrasekhar, A. Chowdhury, E. Duflo, P. Goldsmith-Pinkham *et al.*, 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 27496.

Banerjee, A., R. Chattopadhyay, E. Duflo, D. Keniston, and N. Singh, 2021. "Improving Police Performance in Rajasthan, India: Experimental Evidence on Incentives, Managerial Autonomy, and Training." *American Economic Journal: Economic Policy* 13(1): 36–66.

Banerjee, A., and L. Iyer, 2005. "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." *American Economic Review* 95(4): 1190–213.

Banuri, S., S. Dercon, and V. Gauri, 2019. "Biased Policy Professionals." *World Bank Economic Review* 33(2): 310–27.

Barari, S., S. Caria, A. Davola, P. Falco, T. Fetzer, S. Fiorin, L. Hensel *et al.*, 2020. "Evaluating COVID-19 Public Health Messaging in Italy: Self-Reported Compliance and Growing Mental Health Concerns." medRxiv: 1–19.

Bassi, V., and I. Rasul, 2017. "Persuasion: A Case Study of Papal Influences on Fertility-Related Beliefs and Behavior." *American Economic Journal: Applied Economics* 9(4): 250–302.

Basurto, M. P., P. Dupas, and J. Robinson, 2020. "Decentralization and Efficiency of Subsidy Targeting: Evidence from Chiefs in Rural Malawi." *Journal of Public Economics* 185: 104047.

Bate, R., G. Z. Jin, and A. Mathur, 2011. "Does Price Reveal Poor-Quality Drugs? Evidence from 17 Countries." *Journal of Health Economics* 30(6): 1150–63.

Bate, R., L. Mooney, and K. Hess, 2010. "Medicine Registration and Medicine Quality: A Preliminary Analysis of Key Cities in Emerging Markets." *Research and Reports in Tropical Medicine*: 1: 89–93.

Ben-Nun, Bloom P., G. Arikan, and M. Courtemanche, 2015. "Religious Social Identity, Religious Belief, and Anti-immigration Sentiment." *American Political Science Review* 109(2): 203–21.

Bennett, D., A. Naqvi, and W.-P. Schmidt, 2018. "Learning, Hygiene and Traditional Medicine." *Economic Journal* 128(612): 545–74.

Bertrand, M., S. Djankov, R. Hanna, and S. Mullainathan, 2007. "Obtaining a Driver's License in India: An Experimental Approach to Studying Corruption." *Quarterly Journal of Economics* 122(4): 1639–1676.

Besley, T. J., R. Burgess, A. Khan, and G. Xu, 2022. "Bureaucracy and Development." *Annual Review of Economics* 14: 397–424.

Besley, T., and T. Persson, 2009. "The Origins of State Capacity: Property Rights, Taxation, and Politics." *American Economic Review* 99(4): 1218–1244.

Blair, G., R. Littman, E. R. Nugent, R. Wolfe, M. Bukar, B. Crisman, A. Etim *et al.*, 2021. "Trusted Authorities Can Change Minds and Shift Norms during Conflict." *Proceedings of the National Academy of Sciences of the United States of America* 118(42): e2105570118.

Blair, R. A., T. Curtice, D. Dow, and G. Grossman, 2022. "Public Trust, Policing, and the COVID-19 Pandemic: Evidence from an Electoral Authoritarian Regime." 305: 115045. Social Science & Medicine.

Blattman, C., A. Hartman, and R. Blair, 2014. "How to Promote Order and Property Rights under Weak Rule of Law? An Experiment in Changing Dispute Resolution Behavior." *American Political Science Review* 108(1): 100–120.

Boyer, C., E. Levy Paluck, J. Annan, T. Nevatia, J. Cooper, J. Namubiru, L. Heise *et al.*, 2022. "Religious Leaders Can Motivate Men to Cede Power and Reduce Intimate Partner Violence: Experimental Evidence from Uganda." *Proceedings of the National Academy of Sciences* 119 (31): e2200262119.

Broockman, D. E., 2013. "Black Politicians Are More Intrinsically Motivated to Advance Blacks' Interests: A Field Experiment Manipulating Political Incentives." *American Journal of Political Science* 57(3): 521–36.

———. 2014. "Distorted Communication, Unequal Representation: Constituents Communicate Less to Representatives Not of Their Race." *American Journal of Political Science* 58(2): 307–21.

Bryan, G., J. Choi, and D. Karlan, 2020. "Randomizing Religion: The Impact of Protestant Evangelism on Economic Outcomes." *Quarterly Journal of Economics* 136(1): 293–380.

Bursztyn, L., A. Rao, C. Roth, and D. Yanagizawa-Drott, 2020. "Misinformation during a Pandemic." NBER Working Paper 27417.

Callen, M., S. Gulzar, A. Hasanain, M. Y. Khan, and A. Rezaee, 2020. "Data and Policy Decisions: Experimental Evidence from Pakistan." *Journal of Development Economics* 146: 102523.

Casey, K., R. Glennerster, and E. Miguel, 2012. "Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan." *Quarterly Journal of Economics* 127(4): 1755–812.

Casey, K., R. Glennerster, E. Miguel, and M. Voors, 2023. "Skill versus Voice in Local Development." *Review of Economics and Statistics* 105(2): 311–326.

Cavgias, A., R. Bruce, and L. Meloni, 2023. "Policy Enforcement in the Presence of Organized Crime: Evidence from Rio de Janeiro." *Journal of Development Economics* 162: 103071.

Chhibber, P., and J. S. Sekhon, 2014. "The Asymmetric Role of Religious Appeals in India." Mimeo, Berkeley University.

Clingingsmith, D., A. Khwaja, and M. Kremer, 2009. "Estimating the Impact of the Hajj: Religion and Tolerance in Islam's Global Gathering." *Quarterly Journal of Economics* 124(3): 1133–70.

Condra, L. N., M. Isaqzadeh, and S. Linardi, 2019. "Clerics and Scriptures: Experimentally Disentangling the Influence of Religious Authority in Afghanistan." *British Journal of Political Science* 49(2): 401–19.

Conning, J., and M. Kevane, 2002. "Community Based Targeting Mechanisms for Social Safety Nets: A Critical Review." *World Development* 30(3): 375–394.

Currie, J., W. Lin, and W. Zhang, 2011. "Patient Knowledge and Antibiotic Abuse: Evidence from an Audit Study in China." *Journal of Health Economics* 30(5): 933–49.

Dai, H., S. Saccardo, M. A. Han, L. Roh, N. Raja, S. Vangala, H. Modi *et al.*, 2021. "Behavioral Nudges Increase COVID-19 Vaccinations." *Nature* 597: 404–409.

Das, J., A. Holla, V. Das, M. Mohanan, D. Tabak, and B. Chan, 2012. "In Urban and Rural India, A Standardized Patient Study Showed Low Levels of Provider Training and Huge Quality Gaps." *Health Affairs* 31(12): 2774–84.

De Kadt, D., and H. A. Larreguy, 2018. "Agents of the Regime? Traditional Leaders and Electoral Politics in South Africa." *Journal of Politics* 80(2): 382–99.

De Walque, D., 2007. "How Does the Impact of an HIV/AIDS Information Campaign Vary with Educational Attainment? Evidence from Rural Uganda." *Journal of Development Economics* 84(2): 686–714.

Dell, M., N. Lane, and P. Querubin, 2018. "The Historical State, Local Collective Action, and Economic Development in Vietnam." *Econometrica* 86(6): 2083–121.

Dellavigna, S., and M. Gentzkow, 2010. "Persuasion: Empirical Evidence." *Annual Review of Economics* 2: 643–69.

Deserranno, E., 2018. "Financial Incentives as Signals: Experimental Evidence from the Recruitment of Village Promoters in Uganda." *American Economic Journal: Applied Economics* 11(1): 277–317.

Desposato, S., 2015. *Ethics and Experiments Problems and Solutions for Social Scientists and Policy Professionals*, Routledge. New York NY.

Deutsche Welle, 2020. "Eid Festivities Raise Coronavirus Surge Fears in South Asia."Online news report.

Dizon-Ross, R., P. Dupas, and J. Robinson, 2017. "Governance and the Effectiveness of Public Health Subsidies: Evidence from Ghana, Kenya and Uganda." *Journal of Public Economics* 156: 150–69.

Driscoll, A., G. Cepaluni, F. D. S. Guimarães, and P. Spada, 2018. "Prejudice, Strategic Discrimination, and the Electoral Connection: Evidence from a Pair of Field Experiments in Brazil." *American Journal of Political Science* 62(4): 781–95.

Druckman, J., 2021. "A Framework for the Study of Persuasion." *Annual Review of Political Science* 25: 65–88.

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

Dupas, P., 2011. "Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya." *American Economic Journal: Applied Economics* 3(1): 1–34.

Field, E., and K. Vyborny, 2020. "Information Gaps and de Jure Legal Rights: Evidence from Pakistan." *EDI Working Paper Series*.

Finan, F., B. A. Olken, and R. Pande, 2017. "The Personnel Economics of the Developing State." *Handbook of Economic Field Experiments* 2: 467–514.

Fitzpatrick, A., 2022. "The Impact of Public Health Sector Stockouts on Private Sector Prices and Access to Healthcare: Evidence from the Anti-malarial Drug Market." *Journal of Health Economics* 81(102544): 1–41.

Freedman, M., 2020. "Vote with Your Rabbi: The Electoral Effects of Religious Institutions in Israel." *Electoral Studies* 68: 102241.

Friedman, W., B. Woodman, and M. Chatterji, 2015. "Can Mobile Phone Messages to Drug Sellers Improve Treatment of Childhood Diarrhoea? - A Randomized Controlled Trial in Ghana." *Health Policy and Planning* 30(Supplemental 1): i82–i92.

Gadarian, S. K., S. W. Goodman, and T. Pepinsky, 2021. "Partisanship, Health Behavior, and Policy Attitudes in the Early Stages of the COVID-19 Pandemic." *PLOS One* 16(4): e0249596.

Gaddis, S., 2018. "An Introduction to Audit Studies in the Social Sciences." In *Audit Studies: Behind the Scenes with Theory, Method, and Nuance. Methodos Series*, edited by S. Gaddis. Springer. New York, NY.

Gaikwad, N., and G. Nellis, 2021. "Do Politicians Discriminate against Internal Migrants? Evidence from Nationwide Field Experiments in India." *American Journal of Political Science* 65(4): 790–806.

Gallup Pakistan, 2020. "Attitude Tracker Survey Pakistan Wave 8 Results." Online report.

Gerber, D., and A. Green, 2019. *Get Out the Vote: How to Increase Voter Turnout*. Brookings Institution Press. Washington, DC.

Gine, X., and R. K. Mazer, 2022. "Financial (Dis-)Information: Evidence from a Multi-Country Audit Study." *Journal of Public Economics* 208: 104618.

Grépin, K., V. Mueller, and N. Wu, 2023. "Unmasking the Truth: Experimental Evidence of Facemask Compliance in Bangladesh, Kenya, and Nigeria during the COVID-19 Pandemic." *PLOS Global Public Health* 3(3): e0001086.

Grose, C. R., 2014. "Field Experimental Work on Political Institutions." *Annual Review of Political Science* 17: 355–70.

Grossman, G., S. Kim, J. Rexer, and H. Thirumurthy, 2020. "Political Partisanship Influences Behavioral Responses to Governors – Recommendations for COVID-19 Prevention in the United States." *Proceedings of the National Academy of Sciences*.117 (39): 24144–24153.

Gruber, J. H., 2005. "Religious Market Structure, Religious Participation, and Outcomes: Is Religion Good for You?" *Advances in Economic Analysis & Policy* 5(1).

Grundler, K., and N. Potrafke, 2022. "Policy Advice of Economic Experts: Facts or Preferences?" Mimeo, ifo Institute Munich.

Guiteras, R. P., D. I. Levine, S. P. Luby, T. H. Polley, K. Khatun-E-Jannat, and L. Unicomb, 2016. "Disgust, Shame, and Soapy Water: Tests of Novel Interventions to Promote Safe Water and Hygiene." *Journal of the Association of Environmental and Resource Economists* 3(2): 321–59.

Hetzel, M. W., A. Dillip, C. Lengeler, B. Obrist, J. J. Msechu, A. M. Makemba, C. Mshana *et al.*, 2008. "Malaria Treatment in the Retail Sector: Knowledge and Practices of Drug Sellers in Rural Tanzania." *BMC Public Health* 8: 1–11.

Hjort, J., D. Moreira, G. Rao, and J. F. Santini, 2021. "How Research Affects Policy: Experimental Evidence from 2,150 Brazilian Municipalities." *American Economic Review* 111(5).

Hong, J. Y., and C. Paik, 2021. "Hate Thy Communist Neighbor: Protestants and Politics in South Korea." *Journal of Economic Behavior and Organization* 186: 707–23.

Hsiung, B. O., and P. A. Djupe, 2019. "Religion and the Extension of Trust." *Political Behavior* 41(3): 609–31.

Kalla, J. L., and D. E. Broockman, 2016. "Campaign Contributions Facilitate Access to Congressional Officials: A Randomized Field Experiment." *American Journal of Political Science* 60(3): 545–58.

Karing, A., 2024. "Social Signaling and Childhood Immunization: A Field Experiment in Sierra Leone." Quarterly Journal of Economics.

Kerwin, J. T., 2020. "Scared Straight or Scared to Death? Fatalism in Response to Disease Risks." Mimeo, University of Minnesota.

Khan, A., S. Nasim, M. Shaukat, and A. Stegmann, 2021. "Building Trust in the State with Information: Evidence from Urban Punjab." *Journal of Public Economics* 202: 104494.

Khan, M. Y., 2020. "Mission Motivation and Public Sector Performance: Experimental Evidence from Pakistan." Mimeo, University of Pittsburgh.

Kilic, T., E. Whitney, and P. Winters, 2015. "Decentralised Beneficiary Targeting in Large-Scale Development Programmes: Insights from the Malawi Farm Input Subsidy Programme." *Journal of African Economies* 24(1): 26–56.

Kitschelt, H., and S. Wilkinson, 2007. *Patrons, Clients and Policies: Patterns of Democratic Accountability and Political Competition* Cambridge University Press, Cambridge, United Kingdom.

Lau, R. R., 2020. "Classic Models of Persuasion." *The Oxford Handbook of Electoral Persuasion*: 27–50. Oxford University Press, Oxford, United Kingdom.

Lee, D. S., 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies* 76(3): 1071–102.

Lee, N., 2022. "Do Policy Makers Listen to Experts? Evidence from a National Survey of Local and State Policy Makers." *American Political Science Review* 116(2): 677–88.

Mansuri, G., and V. Rao, 2004. "Community-Based and -Driven Development: A Critical Review." *World Bank Research Observer* 19(1): 1–39.

Martinez-Bravo, M., and A. Stegmann, 2022. "In Vaccines We Trust? The Effect of the CIA's Vaccination Ruse on Immunization in Pakistan." *Journal of the European Economic Association* 20(1): 150186.

McCauley, J. F., 2014. "Pentecostalism as an Informal Political Institution: Experimental Evidence from Ghana." *Politics and Religion* 7(4): 761–87.

McClendon, G. H., 2012. "Ethics of Using Public Officials as Field Experiment Subjects." *Newsletter of the APSA Experimental Section* 3(1): 13–19.

———. 2016. "Race and Responsiveness: An Experiment with South African Politicians." *Journal of Experimental Political Science* 3(1): 60–74.

———. 2019. "Religious Communication and the Effects of Priming." *Oxford Research Encyclopedia of Politics*. Oxford University Press, Oxford, United Kingdom.

McClendon, G., and R. Riedl, 2015. "Religion as a Stimulant of Political Participation: Experimental Evidence from Nairobi, Kenya." *Journal of Politics* 77(4): 1045–57.

———. 2016. "Individualism and Empowerment in Pentecostal Sermons: New Evidence from Nairobi, Kenya." *African Affairs* 115(458): 119–44.

———. 2019. *From Pews to Politics: Religious Sermons and Political Participation in Africa*. Cambridge University Press, Cambridge, United Kingdom.

———. 2021. "Using Sermons to Study Religions' Influence on Political Behavior." *Comparative Political Studies* 54(5): 779–822.

Mehmood, S., S. Naseer, and D. Chen, 2024. AI Education as State Capacity: Experimental Evidence from Pakistan *Mimeo, Oxford*.

Mehmood, S., and A. Seror, 2023. "Religious Leaders and Rule of Law." *Journal of Development Economics* 160: 102974.

Mendez, M. S., and C. R. Grose, 2018. "Doubling Down: Inequality in Responsiveness and the Policy Preferences of Elected Officials." *Legislative Studies Quarterly* 43(3): 457–91.

Milosh, M., M. Painter, D. Van Dijcke, and A. L. Wright, 2021. "Unmasking Partisanship: How Polarization Influences Public Responses to Collective Risk." *Journal of Public Economics* 204.

Mohanan, M., M. Vera-Hernández, V. Das, S. Giardili, J. D. Goldhaber-Fiebert, T. L. Rabin, S. S. Raj *et al.*, 2015. "The Know-Do Gap in Quality of Health Care for Childhood Diarrhea and Pneumonia in Rural India." *JAMA Pediatrics* 169(4): 349–57.

Moreno-Medina, J., 2021. "Sinning in the Rain: Weather Shocks, Church Attendance and Crime." *Review of Economics and Statistics* 105 (1): 54–69.

Murphy, D. M. A., V. Nourani, and D. R. Lee, 2020. "Chatting at Church: Information Diffusion through Religious Networks." *Review of Economics and Statistics* 04(3): 449–464.

Nakajima, N., 2021. "Evidence-Based Decisions and Education Policymakers." Mimeo, Harvard University.

Naurin, E., and P. Öhberg, 2021. "Ethics in Elite Experiments: A Perspective of Officials and Voters." *British Journal of Political Science* 51(2): 890–98.

Nichter, S., 2018. *Votes for Survival: Relational Clientelism in Latin America*. Cambridge University Press, Cambridge, United Kingdom.

Nichter, S., and M. Peress, 2017. "Request Fulfilling: When Citizens Demand Clientelist Benefits." *Comparative Political Studies* 50(8): 1086–117.

Nyhan, B., J. Reifler, S. Richey, and G. L. Freed, 2014. "Effective messages in Vaccine Promotion: A Randomized Trial." *Pediatrics* 133(4): 835–842.

Painter, M., and T. Qiu, 2021. "Political Beliefs Affect Compliance with COVID-19 Social Distancing Orders." *Journal of Economic Behavior and Organization* 185(May): 688–701.

Pakistan Bureau of Statistics, 2017. "Pakistan Census 2017 - Population by Sex, Religion and Rural/Urban."Online report.

Pattan Development Organization, 2020. "Overwhelming Pandemic, Overwhelmed by Fatalist Mindset in Pakistan."Online report.

Pew Research Center, 2018. "The Age Gap in Religion around the World."Online report.

Rahman, K. W., 2022. "International migration and the religious schooling of children in the home country: evidence from Bangladesh." *Journal of Population Economics* 36: 1963–2005.

Rhee, M., M. Sissoko, S. Perry, W. McFarland, J. Parsonnet, and O. Doumbo, 2005. "Use of Insecticide-Treated Nets (ITNs) following a Malaria Education Intervention in Piron, Mali: A Control Trial with Systematic Allocation of Households." *Malaria Journal* 4(35): 1–9.

Rink, A., 2018. "Do Protestant Missionaries Undermine Political Authority? Evidence from Peru." *Comparative Political Studies* 51(4): 477–513.

Rogger, D. O., and R. Somani, 2023. "Hierarchy and Information." *Journal of Public Economics* 219: 1–69.

Sandefur, J., and B. Siddiqi, 2015. "Delivering Justice to the Poor: Theory and Experimental Evidence from Liberia." 1–58. Mimeo, Stanford University.

Sharma, K., 2018. "Clerical Persuasion and Religious Extremism: An Experiment among Sunni and Shia Muslims in Northern India." 1–49. Mimeo, ETH Zurich.

Smith, A. E., 2016. "When Clergy Are Threatened: Catholic and Protestant Leaders and Political Activism in Peru." 9(3): 431–455. Politics and Religion.

Sperber, E., G. McClendon, and O. Kaaba, 2021. "Estimating the Effect of Christian Messages on Civic Engagement: Evidence from a Community-Collaborative Study in Zambia." University of Notre Dame Kellogg Institute Working Paper 442.

Stokes, S. C., T. Dunning, and V. Brusco, 2012. *Brokers, Voters, and Clientelism*, Cambridge University Press, Cambridge, United Kingdom.

Tauchmann, H., 2014. "Lee (2009) Treatment-Effect Bounds for Nonrandom Sample Selection." *Stata Journal* 14(4): 884–94.

Torgler, B., 2006. "The Importance of Faith: Tax Morale and Religiosity." *Journal of Economic Behavior and Organization* 61(1): 81–109.

Vivalt, E., and A. Coville, 2023. "How Do Policy-Makers Update Their Beliefs?" *Journal of Development Economics* 165.

Voors, M., T. Turley, E. Bulte, A. Kontoleon, and J. A. List, 2018. "Chief for a Day: Elite Capture and Management Performance in a Field Experiment in Sierra Leone." *Management Science* 64(12): 5855–76.

Wade, R., 1990. *Governing the Market: Economic Theory and the Role of Government in East Asian Industrialization*. Princeton University Press.

Wilke, A. M., 2024. "How the State Discourages Vigilantism – Evidence from a Field Experiment in South Africa." *American Journal of Political Science*.

Inglehart, R. C. Haerpfer A. Moreno C. Welzel K. Kizilova J. Diez-Medrano M. Lagos P. Norris E. Ponarin, & B. Puranen *et al.*, (eds.), 2018. World Values Survey: Round Six - Country-Pooled Datafile. doi.org/10.14281/18241.8.

Supplementary Online Appendix
Religious Leaders' Compliance with State Authority: Experimental Evidence from COVID-19 in Pakistan

Kate Vyborny ⓘD

S1. Additional Figures and Tables

Table S1.1. Self-Reported Steps to Control COVID during April 2020 Pilot Survey versus July 2020 Experiment

Variable	Pilot (1) Mean/(SE)	Baseline (2) Mean/(SE)	Pairwise t-test (1) − (2) Mean difference
Keeping sermon short	0.289	0.065	0.224***
	(0.022)	(0.009)	
Cleaning mosque	0.781	0.290	0.490***
	(0.020)	(0.016)	
Providing soap	0.520	0.251	0.270***
	(0.025)	(0.015)	
Remove prayer mats	0.639	0.477	0.161***
	(0.024)	(0.017)	
Encourage own prayer mat	0.263	0.124	0.139***
	(0.022)	(0.011)	
Discourage elderly/sick from attending	0.388	0.130	0.258***
	(0.024)	(0.012)	
Distancing	0.542	0.510	0.033
	(0.024)	(0.017)	
No handshakes/hugs	0.369	0.189	0.180***
	(0.024)	(0.014)	
Health announcements in sermon	0.205	0.086	0.119***
	(0.020)	(0.010)	
Health announcements on loudspeaker	0.169	0.047	0.122***
	(0.018)	(0.007)	
Ablution at home	0.277	0.182	0.095***
	(0.022)	(0.013)	
Requesting/requiring masking	0.104	0.240	−0.136***
	(0.015)	(0.015)	
Number of observations	415	830	1,245

Table S1.2. Correlates of Response to Baseline Survey and Mystery-Shopper Call

	Responded to baseline survey (1)	Mystery-shopper call (2)
Rural	−0.013	0.050**
	(0.022)	(0.023)
Primary	−0.016	0.024
	(0.057)	(0.066)
Middle	0.010	0.015
	(0.057)	(0.065)
Matric	−0.033	−0.022
	(0.056)	(0.065)
Intermediate	−0.039	−0.055
	(0.063)	(0.072)
Graduate	−0.012	0.012
	(0.063)	(0.071)
Postgraduate/PhD	0.016	0.070
	(0.067)	(0.074)
Constant	0.330***	0.600***
	(0.057)	(0.064)
Observations	2,700	2,489

Note: Column 1: Sample consists of respondents attempted for baseline in treatment and control arms (excludes super control, who were not called in the baseline). Column 2: Sample consists of respondents attempted for mystery shopper. See fig. S1.2 for more details on sample selection for each instrument. Robust standard errors in parentheses. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Table S1.3. Balance in Response to Mystery-Shopper Calls

|	Responded to mystery shopper	
	(1)	(2)
Secular + religious treatment	−0.025	–
	(0.038)	
Secular treatment	0.010	–
	(0.036)	
Control group (assigned to baseline survey)	–	−0.017
		(0.022)
Constant	0.754***	0.591***
	(0.026)	(0.015)
Observations	813	1,956
p-value, treatment effects equal	0.350	–
p-value, all treatments = 0	0.640	–
Sample	Sample 1	Sample 2

Note: Column 1: Treatment and control arms, respondents who answered baseline survey. Column 2: Control and super-control arms, respondents for whom mystery shopper was attempted. Robust standard errors in parentheses. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Table S1.4. Treatment Effects on Suspicion of Mystery-Shopper Caller

	Enumerator noted respondent was suspicious			
	(1)	(2)	(3)	(4)
Any treatment	−0.010	–	–	–
	(0.024)			
Secular + religious treatment	–	0.007	–	–
		(0.029)		
Secular treatment	–	−0.024	–	–
		(0.026)		
Any treatment = 1	–	–	0.018	−0.009
			(0.047)	(0.038)
BL COVID knowledge correct = 1	–	–	−0.008	–
			(0.044)	
Any treatment = 1 × BL COVID knowledge correct = 1	–	–	−0.025	–
			(0.056)	
Large congregation = 1	–	–	–	−0.025
				(0.042)
Any treatment = 1 × large congregation = 1	–	–	–	−0.006
				(0.050)
Observations	629	629	563	583
Control group mean	0.092	–	–	–

Note: BL COVID knowledge correct is an indicator for whether the respondent answered both knowledge questions about COVID transmission correctly at baseline. Robust standard errors in parentheses. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Figure S1.1. Timeline.
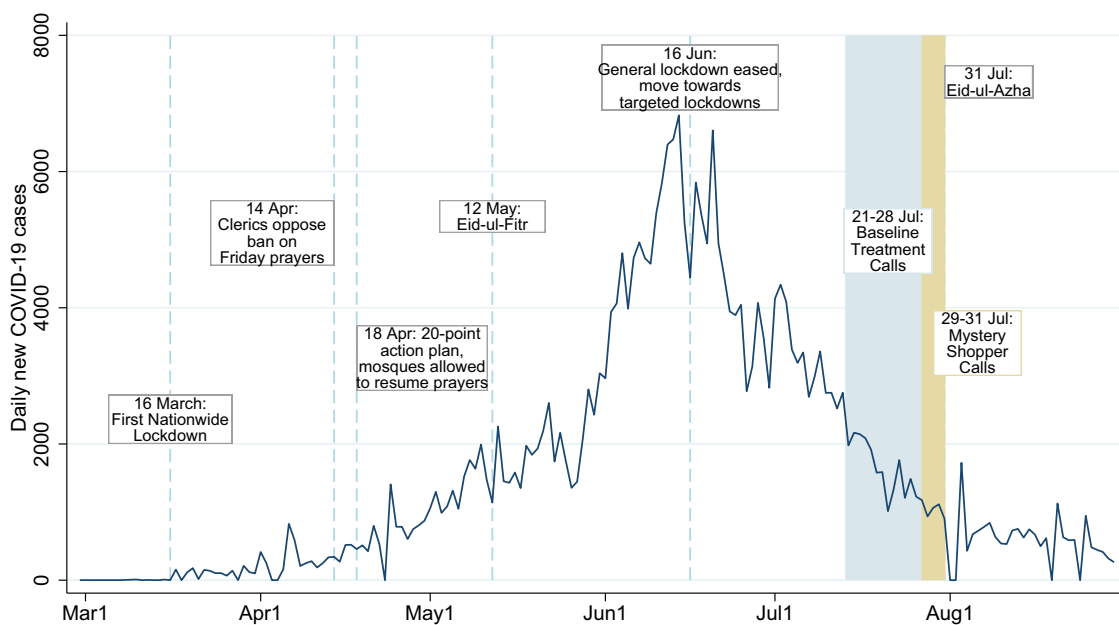
Source: COVID case data obtained from ourworldindata.org.

Table S1.5. Alternative Mediators of Heterogeneous Treatment Effects by Baseline COVID Knowledge

| | Panel A | | |
| | Baseline knowledge correct | | |
	(1)	(2)	(3)
Ed 6–8 years	–	0.021	–	–
		(0.056)		
Ed 9–10 years	–	0.081	–	–
		(0.056)		
Ed 12+ years	–	0.131**	–	–
		(0.058)		
Urban = 1	–	–	−0.041	–
			(0.049)	
Received few/no COVID messages last week = 1	–	–	–	−0.145***
				(0.043)
Observations	–	544	544	544

| | Panel B | | | |
| | Index—Advice to Mystery Shopper | | | |
	(1)	(2)	(3)	(4)
Any treatment = 1	−0.022	−0.091	−0.032	−0.006
	(0.050)	(0.065)	(0.054)	(0.057)
BL COVID knowledge correct = 1	−0.044	−0.053	−0.045	−0.047
	(0.049)	(0.048)	(0.049)	(0.050)
Any treatment = 1 × BL COVID knowledge correct = 1	0.131**	0.132**	0.132**	0.128**
	(0.060)	(0.060)	(0.061)	(0.061)
Ed 6–8 years	–	−0.010	–	–
		(0.060)		
Ed 9–10 years	–	0.038	–	–
		(0.062)		
Ed 12+ years	–	0.066	–	–
		(0.069)		
Any treatment = 1 × Ed 6–8 years	–	0.096	–	–
		(0.075)		
Any treatment = 1 × Ed 9–10 years	–	0.072	–	–
		(0.075)		
Any treatment = 1 × Ed 12+ years	–	0.088	–	–
		(0.084)		
Urban = 1	–	–	−0.001	–
			(0.051)	
Any treatment = 1 × Urban = 1	–	–	0.045	–
			(0.063)	
Received few/no COVID messages last week = 1	–	–	–	−0.015
				(0.050)
Any treatment = 1 × Received few/no COVID messages last week = 1	–	–	–	−0.039
				(0.061)
Observations	544	544	544	544

Note: Sample 1 (treatment and control arms, respondents who answered baseline survey and mystery-shopper call). Sample size varies from table 3 because of non-response to baseline questions due to some respondents hanging up before completing the baseline call. Robust standard errors in parentheses. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Figure S1.2. Sampling and Response by Treatment Arm.
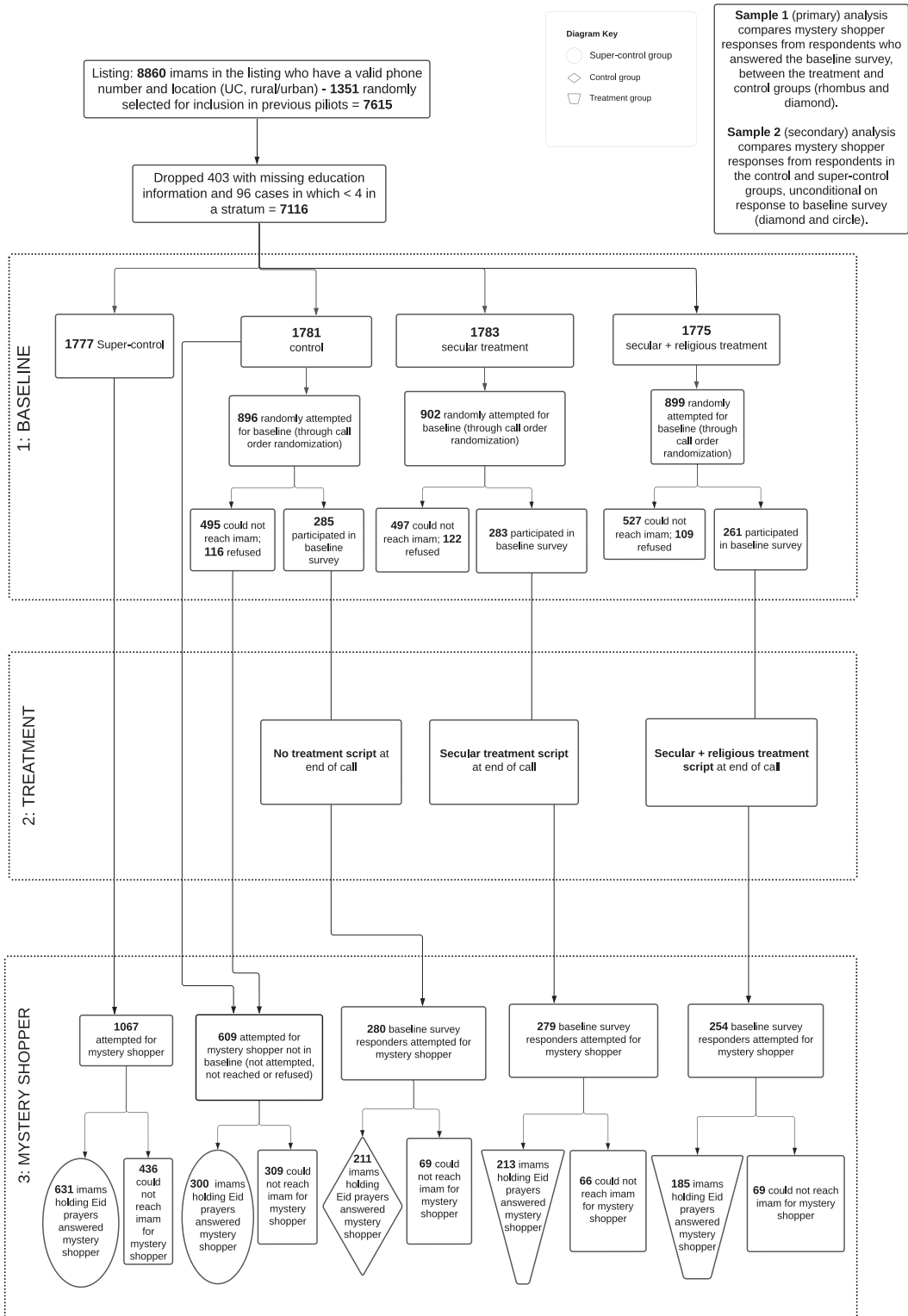
Table S1.6. Effect of Receiving Baseline Survey on Religious Leaders' Instructions to Mystery Shoppers: Control versus Super Control

	Index—Advice to MS
	(1)
Control group (assigned to baseline survey)	0.019
	(0.018)
Observations	1,142

Note: Comparison of control versus super control in Sample 2 (respondents who answered the mystery-shopper call). Robust standard errors in parentheses. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Figure S1.3. Heterogeneous Treatment Effects by Gap between Treatment and Mystery Shopper.
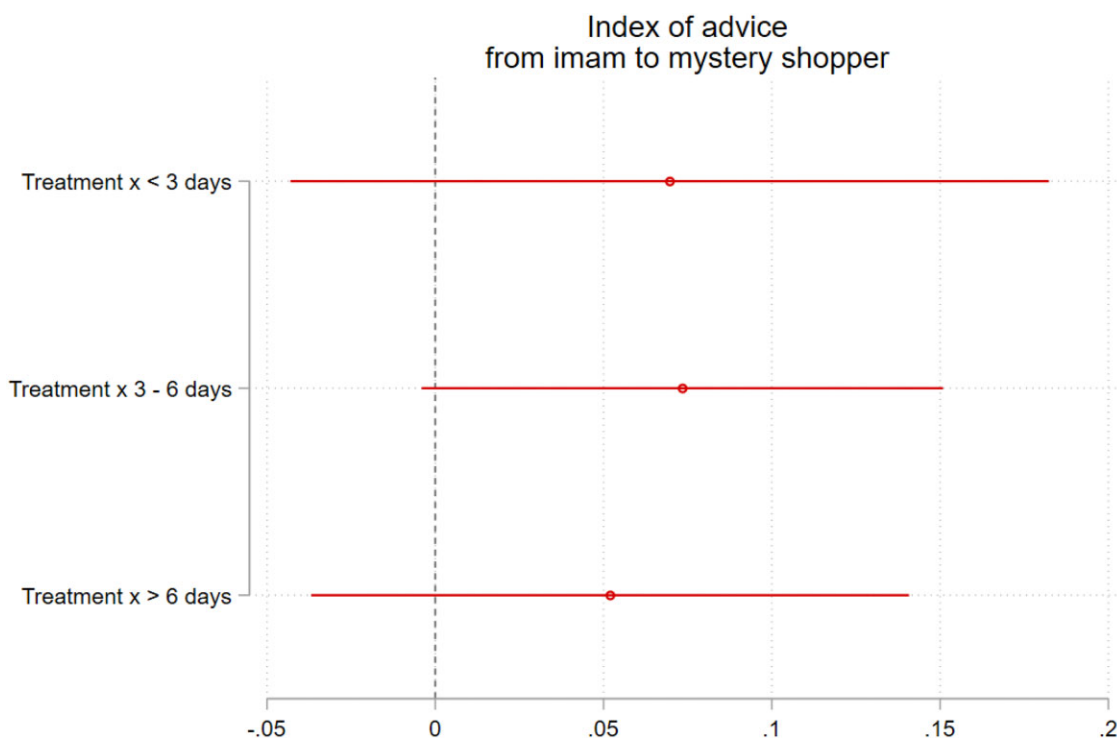
Note: Figure shows coefficients β_1 to β_3 from regressing the mystery-shopper index $INDEX = \beta_0 + \beta_1 TREAT \times SHORTGAP + \beta_2 TREAT \times MEDIUMGAP + \beta_3 TREAT \times LONGGAP + \beta_4 MEDIUMGAP + \beta_5 LONGGAP + \epsilon$. Order of phone calls was randomized. Robust standard errors; 95 percent confidence intervals shown.

Table S1.7. Heterogeneous Effects by Relationship with Secular and Religious Authorities

	Panel A: Pooled treatments Index—Advice to mystery shopper			
	(1)	(2)	(3)	(4)
Any treatment = 1	0.078**	0.062*	0.050	0.102*
	(0.030)	(0.036)	(0.059)	(0.053)
Subgroup = 1	0.103*	−0.028	0.001	0.033
	(0.059)	(0.049)	(0.052)	(0.049)
Any treatment = 1 × subgroup = 1	−0.052	0.009	0.022	−0.037
	(0.072)	(0.060)	(0.066)	(0.062)
Observations	558	502	593	555
Subgroup	Gov't trusted source	Governing party constituency	Attended madrasa	Ulema trusted source
p-value: Total effect on subgroup = 0	0.692	0.135	0.016	0.046

	Panel B: Individual treatment arms Index—Advice to mystery shopper			
	(1)	(2)	(3)	(4)
Secular treatment = 1	0.081**	0.077*	0.076	0.136**
	(0.035)	(0.042)	(0.069)	(0.066)
Subgroup = 1	0.103*	−0.028	0.001	0.033
	(0.060)	(0.049)	(0.052)	(0.049)
Secular treatment = 1 × subgroup = 1	−0.009	−0.023	0.004	−0.067
	(0.081)	(0.070)	(0.077)	(0.076)
Secular + religious treatment = 1	0.074**	0.045	0.021	0.073
	(0.035)	(0.042)	(0.072)	(0.059)
Secular + religious treatment = 1 × subgroup = 1	−0.092	0.043	0.041	−0.014
	(0.082)	(0.068)	(0.080)	(0.070)
Observations	558	502	593	555
Subgroup	Gov't trusted source	Governing party constituency	Attended madrasa	Ulema trusted source
p-value: Total effect of secular T on subgroup = 0	0.324	0.335	0.020	0.060
p-value: Total effect of secular + religious T on subgroup = 0	0.804	0.103	0.072	0.121
Control mean Y | subgroup = 0	0.360	0.393	0.372	0.352
Control mean Y | subgroup = 1	0.463	0.365	0.373	0.385
Proportion of sample in subgroup	0.188	0.388	0.809	0.733

Note: Sample 1 (treatment and control arms, respondents who answered baseline survey and mystery-shopper call). Sample size varies from table 3 and across columns because of non-response to baseline questions due to some respondents hanging up before completing the baseline call (columns 1, 3, and 4), and insufficient geographical information to map some respondents to their political constituency (column 2). Robust standard errors in parentheses. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

S2. Additional Information on Study Design

S2.1. Treatment Scripts

The full SurveyCTO version of baseline and treatment call including Urdu is available at https://www.dropbox.com/s/moatri1ymc437at/Imam%20Calls%20COVID-19_scale_up_first_round.xlsx?dl=0.

The full SurveyCTO version of the Mystery Shopper call including Urdu is available at https://www.dropbox.com/s/h2y9bk8fmhfnae4/Imam%20Calls%20COVID-19_mystery_shopper_scaleup_first_round.xlsx?dl=0.

S2.1.1. Secular Information: All Treatment Arms

- According to leading health organizations like the World Health Organization and Pakistan Medical Association, when someone coughs, sneezes, or speaks they spray small liquid droplets from their nose or mouth which may contain virus. This makes coronavirus much more contagious than the flu since you can catch the disease from an infected person even if you are standing a few steps away from them.
- Additionally, approximately half the people with Coronavirus don't show any sign/symptom, so they can spread it to people around them without knowing it.
- Elderly people who catch the Coronavirus have a much higher chance of getting seriously ill and dying as compared to normal flu.
- Virus can also be spread when someone touches something—so someone who is sick and doesn't show it, can pass on the virus by touching the door, the wuzu water tap, and other places in the building.
- Because the disease is difficult to detect since many infected people do not show any signs, and there are many elderly and vulnerable people around us who are at high risk of catching it, doctors recommend that one must protect oneself and each other by wearing a mask in public, keeping hands clean, and avoiding close and unnecessary physical contact (such as shaking hands).
- Did you know the Government of Pakistan [COMBINED T ONLY: AND leading ulema] have jointly endorsed a statement that to avoid spread of COVID:

    - Elderly must stay at home
    - People must stay spread apart 6 feet
    - No wuzu (religious ablution) in the mosque
    - People must wear masks
    - Remove the jah namaz/floor mat
    - Wash floors between prayers

- Do you feel that these points are relevant for your decisions about planning for your mosque activities in the coming weeks? Why or why not?
- I would like to send you this information via WhatsApp or SMS. Which medium would you prefer?
- As you are an important leader and an example for your community, people look to you as an example and guide on how to face this issue.
- I would like to request you to consider helping protect your community by taking the following four steps:

    - Wear a mask as an example to others.
    - Announce in the khutbah each day that people should wear masks when they come to prayer.
    - Announce the importance of wearing masks over the mosque loudspeaker to the community.
    - Ask people to do wuzu at home and bring their own jah namaz to mosque.

- Would you consider doing this?

S2.1.2. Additional Text Included in Religious Arm Treatment Script

- Islamic scholars have referenced the following hadiths about avoiding the spread of disease, when discussing how the ummah should respond to Coronavirus threat.
- Abu Salamah reported: The Messenger of Allah, peace and blessings be upon him, said, "There is no infection (without the decree of Allah). Do not mix the sick with the healthy." Source: Sahih Muslim 2221 b
- Because so many people have the illness without showing any sign, the recommendation is to wear a mask to protect others.
- Do you feel this hadith is relevant to the situation with Coronavirus?

- The Islamic affairs minister in Saudi Arabia has issued strict guidelines about social distancing, wearing masks and performing ablution at home for masjids.
- Do you think these guidelines are helpful for you in planning for congregational prayers at mosque?
- Leading muftis at Al Azhar University in Egypt have provided the following guidelines on what mosques should do to prevent the spread of coronavirus: "The real danger lies in the rapid spread of the virus, and that gatherings are a direct cause of infection. The Islamic Shariah is keen on taking care of the interests of people both in this world and in the Hereafter, and on warding any harm off them. The major objective of the Shariah is to preserve people's lives and protect them from all dangers. The pandemic that has hit the whole of humanity constitutes a necessity that leads to using some juristic concessions in the performance of some Islamic obligations, in order to avoid the danger of the further spread of the virus outbreak."
- Do you think these guidelines are helpful for you in planning for congregational prayers at mosque?
- There are similar fatwas from Ayatollah Sistani and Ayatollah Khameini, do you want to hear it?[7] Grand Ayatollah Sayyid Ali Sistani in Iraq has issued a fatwa stating that it is the duty of the believers to follow the health guidelines. He also stated that when health guidelines are ignored where there is fear of becoming infected with the virus and a serious possibility of death or significant harm if the person were to catch the disease, then not adhering to guidelines would be sinful.
- Do you think these guidelines are helpful for you in planning for congregational prayers at mosque?
- Similarly Ayatollah Khameini in Iran announced: "Surely, everything that leads to society's health and everything that helps prevent the disease from becoming rampant is a good deed and on the contrary, everything that helps the disease become widespread is a sin. Allah the Exalted has made us responsible towards our own health and that of others and the health of the people. So, the first word of advice is that we should consider it our responsibility to completely observe the instructions specified by officials and to act on them
- Do you think these guidelines are helpful for you in planning for congregational prayers at mosque?

S2.1.3. Mystery-Shopper Script
- Assalam-o-alaikum. My father and I wanted to attend the Eid prayers? Could you please tell me when are they being held?
- Do we have to bring our own jah namaz?
- Do we have to do wuzu at home?
- Do we have to wear a mask?
- ENUMERATOR: Did the respondent mention it is COMPULSORY to wear a mask in order to come?
- Thanks so much for your time, goodbye.

S3. Ethics
Duke University Campus IRB approved this study under protocol number 2020-0432.[8] In this appendix, we discuss ethical issues related to the experiment, building on discussions of ethics in field experiments with elites by McClendon (2012), Desposato (2015), Naurin and Öhberg (2021).

Participants in the treatment and control groups were consented at the beginning of the baseline survey and were informed of the duration of the survey, as well as the fact that they would not be compensated for their participation.

The protocol included the use of mystery shoppers (also known as an "audit study" approach), an approach used extensively in recent development economics research to observe behavior of agents in

7   Shia religious authorities; Shias are a religious minority in Pakistan, thus respondents are prompted to check their interest in this fatwa.
8   At the time of the inception of this study, the Center for Economic Research in Pakistan did not have a local IRB, and to the researchers' knowledge other institutions in Pakistan did not accept external applications for IRB review.

contexts where self-reports are expected to be affected by social desirability or other forms of bias, particularly in the fields of public services and the behavior of public officials (Gaddis 2018; Banerjee et al. 2021; Anagol, Cole, and Sarkar 2017; Gine and Mazer 2016; Bertrand et al. 2007; Dizon-Ross, Dupas, and Robinson 2017; Hetzel et al. 2008; Mohanan et al. 2015; Currie, Lin, and Zhang 2011; Friedman, Woodman, and Chatterji 2015; Fitzpatrick 2021; Bate, Mooney, and Hess 2010; Bate, Jin, and Mathur 2011; Das et al. 2012; Field and Vyborny 2020; Broockman 2013, 2014; McClendon 2016; Kalla and Broockman 2016; Mendez and Grose 2018; Driscoll et al. 2018; Gaikwad and Nellis 2021). Some deception is inherent in this methodological approach. In this scenario, it is necessary to use deception in order to observe their advice to congregants in a real-life situation. Respondents will not be punished or reported for their answers. The mystery-shopper data collected is kept confidential and only used to analyze the impact of the intervention on how they advise the community.

Key considerations in the decisions on compensation and deception included the following:

- The limited amount of time required from respondents. The mystery-shopper call asks basic, simple information that other congregants would commonly ask an imam, so this call will be well within the normal scope of activities for respondents. The total time for each call is approximately 2 minutes, so this instrument does not represent an undue burden on their time. The total time taken for respondents who participated in the baseline survey and were called in the mystery-shopper exercise was approximately 10 minutes.
- The non-sensitive nature of the questions asked. Mosque activities during communal prayers (prayers, use of masks, social distancing, etc.) are public and could be observed by any casual observer in the area by stepping into the mosque. Hence the mystery shoppers are not collecting any sensitive or private information.
- The influential role of respondents as community leaders, such that they would not be considered marginalized or vulnerable.
- The ex ante most likely anticipated impact of treatment (enhancing masking and other protective measures in the mosque) would have potential benefits to congregants, while the control group would simply continue with the status quo, and neither respondents nor their congregants would experience any harm as a result of the experiment. The mystery-shopper script involves the caller further asking whether a mask is really required even when the imam initially says he should wear it; however, we anticipated no substantive risk of discouraging mask use because of this single query, given widespread non-compliance which was already documented as the norm (Pattan Development Organization 2020).

For the super-control group only, participants received only the 2 minute mystery-shopper call and no other instruments. Therefore, they would not be briefed on the study, making an active consent process impossible for this subgroup only. This was deemed to be acceptable based on the following criteria:

- Criterion 1: The research involves no more than minimal risk to the subjects. There is no intervention with the control group participants; the only interaction is the collection of a few simple variables on their current practices in the mosque through the mystery-shopper instrument. We do not anticipate any risk to these subjects.
- Criterion 2: The waiver or alteration will not adversely affect the rights and welfare of the subjects. Mosque activities during communal prayers (prayers, use of masks, social distancing, etc.) are public and could be observed by any casual observer in the area by stepping into the mosque. Hence we are not collecting any sensitive or private information. The mystery-shopper call is the only research activity with these participants and presents an extremely minimal burden (a 2 minute call with only simple information frequently asked from imams and also readily available to any observer in the area). Thus inclusion in the study without briefing would not lead to any adverse effects or negatively affect the welfare of the subjects.

Table S4.1. Robustness to Alternative Construction of Index

	Dependent variables: Alternative constructions of Mystery Shopper index		
	(1)	(2)	(3)
Any treatment	0.0659**	0.0629**	0.0623**
	(0.0263)	(0.0269)	(0.0244)
Observations	609	621	617
Variables excluded from index	None	Prayer mat	Prayer mat; mask required

Note: Robust standard errors in parentheses. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

- Criterion 3: The research could not practically be carried out without the waiver or alteration. The goal of the super-control group is to detect the effect of being surveyed. Thus briefing participants on the study would make it impossible to carry out this component effectively.

S4. Variations from Study Pre-registry

The study was pre-registered (AEARCTR-0005740, Version 2.0).

Registry Version 1.0 was registered during early-stage piloting. Piloting was complete, the design was revised, and then Version 2.0 was registered before the main study fieldwork. Observations collected during piloting, before Version 2.0 was registered, are not included in the analysis presented in the paper.

This appendix discusses variations between the study pre-registry and the analysis presented in the paper.

Variable Included in Instrument but Omitted from Pre-registration in Error

Due to an error in coordination between post-pilot questionnaire revisions and registration, the outcome "asks the respondent to bring his own prayer mat" was left out of the registry entry inadvertently.

The full script of the instrument (supplementary online appendix S2) demonstrates that all the mystery-shopper outcome variables collected are presented in the paper.

Our results are not dependent on the inclusion of this variable. The index results are robust to the exclusion of this variable (table S4.1).

Stated intent questions

At the end of the baseline call, after the persuasion script was complete for the treatment group, we also asked respondents in all treatment arms to report their planned steps to reduce the spread of COVID in the mosque. We listed these stated intent variables in the study registry. We consider these variables weaker evidence than the mystery-shopper data, because they are self-reported and therefore subject to response bias. In addition, because of variation in the length of the call between treatment and control, response rates to this module differ between treatment arms, causing potential sample selection bias in the estimates. Table S4.2 shows the results for the pre-specified variables from this module with and without Lee (2009) bounds; the results are consistent with the main results but imprecisely estimated after bounding.

Table S4.2. Treatment Effects on Stated Intent Questions Asked at End of Call 1

Panel A: Basic estimates

	Number of steps (1)	Short sermon (2)	Clean mosque (3)	Soap (4)	Remove mats (5)	Elderly/sick (6)	Distancing (7)	Announcements (8)	Ablution at home (9)	Mask (10)	Other step (11)	Answered intent Qs (12)
Any treatment	0.1823	0.0238	0.0021	−0.0108	0.0535	0.0051	−0.0422	0.0421	0.0750***	0.0568*	−0.0285	−0.1212***
	(0.1162)	(0.0207)	(0.0312)	(0.0302)	(0.0351)	(0.0263)	(0.0372)	(0.0258)	(0.0289)	(0.0305)	(0.0345)	(0.0261)
Observations	670	670	670	670	670	670	670	670	670	670	670	829
Control group mean Y	2.0949	0.0751	0.2292	0.2016	0.2846	0.1423	0.3874	0.1304	0.1542	0.1660	0.3241	

Panel B: Lee bounds

	Number of steps (1)	Short sermon (2)	Clean mosque (3)	Soap (4)	Remove mats (5)	Elderly/sick (6)	Distancing (7)	Announcements (8)	Ablution at home (9)	Mask (10)	Other step (11)
Any treatment											
Lower	−0.1325	0.0090	−0.0257	−0.0392	0.0086	−0.0281	−0.1033**	0.0048	0.0493	0.0356	−0.0780*
	(0.1439)	(0.0242)	(0.0380)	(0.0358)	(0.0415)	(0.0309)	(0.0448)	(0.0307)	(0.0339)	(0.0346)	(0.0425)
Upper	0.6301***	0.0959***	0.1324***	0.1189***	0.1666***	0.1300***	0.0548	0.1559***	0.2074***	0.1936***	0.0801*
	(0.1458)	(0.0144)	(0.0466)	(0.0457)	(0.0480)	(0.0438)	(0.0480)	(0.0178)	(0.0456)	(0.0457)	(0.0477)
Observations	829	829	829	829	829	829	829	829	829	829	829

Note: Panel A: Treatment and control arms, respondents who answered baseline survey. Robust standard errors in parentheses. Panel B shows Lee (2009) treatment effect bounds estimated using the Leebounds package developed by Tauchmann (2014). $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.