

Supplemental Appendix:  
Social Accountability and Minority Status  
Contributions to Public Goods by Hindus and Muslims in Delhi Slums

Melani Cammett  
Harvard University  
[mcammett@g.harvard.edu](mailto:mcammett@g.harvard.edu)

Poulomi Chakrabarti  
Queen's University/ Harvard University  
[pchakrabarti@fas.harvard.edu](mailto:pchakrabarti@fas.harvard.edu)

David Romney  
Brigham Young University  
[david.romney@byu.edu](mailto:david.romney@byu.edu)

December 22, 2022

# Contents

<b>1</b>	<b>Qualitative Fieldwork, Sampling, and Survey Administration</b>	<b>3</b>
1.1	Site Selection . . . . .	3
1.2	Qualitative Fieldwork . . . . .	4
1.2.1	Qualitative Data Protocols . . . . .	6
1.3	Survey Administration . . . . .	10
<b>2</b>	<b>Survey Instrument: Key Measures and Experimental Manipulations</b>	<b>12</b>
2.1	Outcome Measures . . . . .	12
2.2	Measures of Religion and Religiosity . . . . .	13
2.3	Measures of Political Trust . . . . .	16
2.4	Measures of Social Ties . . . . .	17
2.5	Measure of Caste . . . . .	19
2.6	Geocoordinates and Diversity Measure . . . . .	21
2.7	Experimental Manipulations . . . . .	22
<b>3</b>	<b>Main Analysis and Design</b>	<b>25</b>
3.1	Models without Interactions . . . . .	25
3.2	Interactive Model with Separate Treatments and Outcome Measures . . . . .	27
3.3	Full Control Model . . . . .	30
3.4	Pre-Registration . . . . .	33
3.5	Ethics . . . . .	34
<b>4</b>	<b>Robustness Checks</b>	<b>35</b>
4.1	Balance . . . . .	35
4.2	Survey Purpose and Manipulation Checks . . . . .	37
4.3	Timing of Survey Administration and Treatments . . . . .	39

4.4	Factor Analysis with Religious Practice Variables . . . . .	40
4.5	Political Trust as a Mechanism . . . . .	44
4.6	Diversity as an Alternative Hypothesis . . . . .	48
4.7	What is a Neighborhood? . . . . .	51
4.8	Effects by Neighborhood . . . . .	51
4.9	Ceiling/Floor Effects (Tobit Model) . . . . .	54
4.10	Enumerator Effects . . . . .	56
4.11	Effects by Caste Level, for Hindus . . . . .	58
4.12	Religious Ties as Mechanism, Separate Treatments . . . . .	59

# 1 Qualitative Fieldwork, Sampling, and Survey Administration

## 1.1 Site Selection

We selected nine candidate sites in different parts of Delhi based on their ethnic composition. These sites are administratively categorized as one of several types of informal settlements by the government, including *Jhuggi Jhopri* Cluster, Slum Designated Area, Unauthorized Colony, and Resettlement Colonies (see Table 1) (Heller et al., 2015). With the exception of resettlement colonies, these sites are unplanned and its residents do not have tenure rights. Regardless of their legal status, all areas have limited and poor access to public services. The sites are further comprised of twenty *bastis* or smaller organic communities.

Table 1: Types of Informal Settlements in Delhi

Type of Settlement	Population (in millions, 2006)	Percentage of Delhi Population
<i>Jhuggi Jhopri</i> Cluster (JJC)	2.448	14.8%
Slum Designated Area	3.148	19.1%
Unauthorized Colony	0.874	5.3%
Resettlement Colonies	2.099	12.72%

Our initial goal was to identify three types of slums –“heterogeneous” communities with a close to 50/50 Hindu-Muslim split, “homogeneous” areas with an overwhelmingly Muslim majority, and “homogeneous” sites with an overwhelmingly Hindu majority. We relied on publicly available, booth-level electoral lists, collected from the Chief Electoral Officer of Delhi website, to estimate the proportion of the Muslim population in each site. The names in the electoral lists were digitized and matched against a list of common South Asian Muslim names to arrive at an estimate of the Muslim population in each *basti*. While *bastis* display greater variation in the proportion of Muslim population, settlements at the city level tended to be largely Muslim or Hindu. Mixed communities were identified at the *basti* level from the subset of the Hindu or Muslim sites listed below. The electoral lists also provided information on the number of voters in each *basti* and hence allowed us

to generate a rough estimate of the population of each site. Our final survey was administered in sixteen *bastis* in five sites, selected from the initial list (see Table 2). These sites were similar in socioeconomic characteristics, quality of public services, and age (ranging from 30-40 years). Note that the number of participants here differs from that reported in the paper because three treatment arms excluded from the analysis in the paper are included in this table.

Table 2: Number of Participants Surveyed and Response Rate by *basti*

	Site	Basti	Houses	Total Surveyed	Closed House	Refused	Response Rate
1	Site A	Basti A1	1,400	1,035	320	45	74%
2	Site B	Basti B1	1,260	540	520	200	43%
3	Site C	Basti C1	1,550	752	550	248	49%
4	Site C	Basti C2	130	70	30	30	10%
5	Site C	Basti C3	1,080	609	300	171	56%
6	Site C	Basti C4	1,065	534	176	355	50%
7	Site C	Basti C5	125	52	23	50	42%
8	Site C	Basti C6	850	402	340	108	47%
9	Site D	Basti D1	900	380	340	180	42%
10	Site D	Basti D2	400	215	100	85	54%
11	Site D	Basti D3	400	190	110	100	48%
12	Site D	Basti D4	300	140	120	40	47%
13	Site D	Basti D5	2,400	730	1,050	620	30%
14	Site D	Basti D6	140	60	40	40	43%
15	Site E	Basti E1	230	80	105	45	35%
16	Site E	Basti E2	80	55	20	5	69%

Note: Slum names are changed to ensure the anonymity of research subjects.

## 1.2 Qualitative Fieldwork

Before administering the survey, a team of field researchers from a research institute in Delhi, the Centre for Policy Research (CPR), aided us in gathering qualitative data on slum settlements. From March to June 2018, the researchers collected data on twenty *bastis* from nine sites, as discussed earlier. The team generated 26 detailed reports that provided a wealth of information on the demography, quality of infrastructure and public services, political and social life, and local leadership. These reports helped guide our site selection and design of the survey instrument. For example, we were able to choose heterogeneous and homogeneous communities that were largely

similar to each other in other respects based on findings from the fieldwork. The reports also informed our choice of drainage cleaning as an appropriate focus for the survey. The condition of drainage was consistently poor in all the selected sites, and residents across all settlements identified the quality of drainage as one of the main concerns in the area (See Table 3).

Table 3: Condition of Drainage in Selected Sites



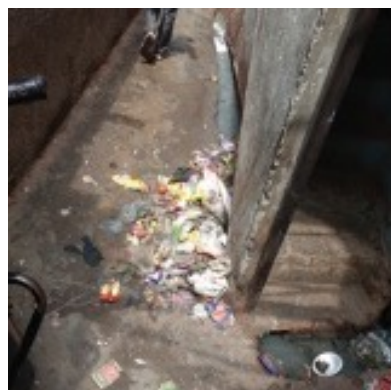
Main drainage clogged with garbage in Site D



Internal drain from *basti* connecting to the main drain in Site C



Sources of water supply right next to internal drains in Site E



Inner drains clogged with garbage and plastic in Site B



Overflowing drains in Site A



Garbage and overflowing drains in Site B

In addition, qualitative fieldwork shaped the design of the experiment. We learned that slum residents did not rate government agencies and non-governmental organizations (NGOs) as their preferred organizations to take charge of drainage. Instead, they favor low-cost private options over both free public services and those offered by NGOs. For example, after a new municipal policy

eliminated fees at community toilet complexes (CTCs) in one of the slums, residents complained that the cleanliness of the toilets declined. Community members were more receptive to paying a small fee to maintain community facilities. Furthermore, residents viewed NGOs as money-making schemes for well-connected locals who want to extract funds from the government and foreign donors, and believed that NGO leaders were not motivated by the welfare of the community. The field reports were also helpful in interpreting the findings from the survey. Fieldwork by the team at CPR was guided by detailed set of protocols for interviews and the recording of observational data, discussed in the following section.

### **1.2.1 Qualitative Data Protocols**

This section provides the protocols and instructions for the qualitative fieldwork carried out by researchers at the Centre for Policy Research in Delhi, India.

\*\*\*

The fieldwork will draw upon a combination of observations, government reports (if available), and information obtained through interviews. *Please note the sources of information in the field notes. For example, for interviews, note who are you interviewing and the selection criteria (i.e., why did you choose that particular person to interview?; Why do you think that they are a reliable source of information?).*

### **Observational Data**

#### **1. Settlement Characteristics**

- (a) Prepare a rough map of the settlement with respect to surrounding roads and landmarks, approach roads/ major flyovers and main drainage networks.
- (b) Focus on the following public services: Drainage, Local Lanes, Garbage Collection, Community Toilets, Water Supply.
- (c) If available, collect information on the settlement (i.e., legal status, demography, etc.) from government reports, such as those issued by the Delhi Development Authority (DDA) and Delhi Urban Shelter Improvement Board (DUSIB).

## 2. Demography

- (a) In addition to the demographic profile of the community, find out any salient cleavages along caste, religious, or other lines. If there are any conflicts on public services or other local issues, is it manifested within or between certain communities?
- (b) Report after the first site visit to decide if the settlement is suitable for the study, based on two criteria: the extent to which these public services are important in the settlement and the extent to which they can be administered by a hypothetical service.

## 3. Drainage and Public Services

- (a) How does the drainage connect to the city-wide network?
- (b) What is the state of the drainage—is it clogged/looks like it's well maintained?
  - i. Who is in charge of maintaining the drainage?
  - ii. Ask local residents about issues with drainage—how often is it clogged? Does the MCD maintain it? If not, do residents attempt to clean it themselves? Are NGOs or other community organizations involved?
  - iii. Under the current system, how many people have to coordinate in order to successfully maintain drainage?
- (c) Which blocks/communities are connected to the drainage network? What is the demographic profile of these communities? *How did you assess the demographic characteristics?*
  - i. Mostly Hindu
  - ii. Mostly Muslim
  - iii. Mixed. If mixed, roughly what proportion of the two?
- (d) Local residents and public service delivery:
  - i. If they are open to a third-party taking care of the public good/utility, then would they be more responsive to an NGO, CBO or for-profit firm?



- ii. Are the residents aware of paid arrangements for a third-party service (either the government or a private company), either in their own settlement or in others, for taking care of the good?
- iii. How open would they be, personally, to having a third party take care of this particular issue, for a reasonable price, and what would a reasonable price be?

## **Interviews with Residents/Leaders**

### **1. Personal Data**

- (a) How long have you been living in this *basti*? Can you tell us how the *basti* has changed over the years? What changes have you seen – could you specify a few examples to help us understand better?
- (b) (For those not born in the *basti*) When did you first come to the *basti*? Where were you living previously (village/another neighborhood/another city or town)? How did you come here - found housing by yourself; through village acquaintance/friend/relative; acquaintance/friend/relative in previous city/neighborhood; through relocation?
- (c) Where do residents in the *basti* hail from?

### **2. Interpersonal Relationships**

- (a) How are interpersonal relationships among the people? Do people go out together for occasions like marriages? Are there major festivals celebrated together in this *basti*? Do you celebrate occasions like marriages and festivals together in this *basti* with other *basti* residents?
- (b) Do you have friends in this *basti*? Are they from the same lane as yours or from other lanes?
- (c) In times of need (illness, family emergency, etc.), would you approach your neighbors in the *basti* for help with borrowing money? Have you done so in the past?

- (d) (For migrants or those who go to village frequently) Are interpersonal relationships stronger in the village or in the city? Where, in your opinion, do people help each other or look out for each other more in the city or in the village?

### 3. Safety

- (a) How safe is the *basti* and neighborhoods around it, particularly for women and newcomers?
- (b) If it is not safe, then do *basti* residents speak about this amongst themselves, and collectively think of solutions?
- (c) (For women informants) By when would you prefer to return home if out for work, or in the marketplace? Do women often go together to the marketplace? Are Community Toilet Complexes (CTCs) safe and accessible after dark?

### 4. Collective Action

- (a) What in your opinion are major issues facing the *basti* that need attention (collectively)?
- (b) Do people organize around collective concerns of the *basti* like water, electricity, roads, safety, etc.? How do residents typically organize – by themselves or are there some who mobilize to lead the process?
- (c) Have there been instances in the past where residents have organized collectively? Can you tell us a bit about them? Were you part of collective action?
- (d) Are there any resident organizations or cooperatives (or a residents' welfare association) in this neighborhood that you are aware of?
- (e) Are you part of any such collectives? How do they operate? Who are the other members of the collective and how often does the collective meet?
- (f) Is there someone who heads or manages these collectives? Are the heads/managers elected or unelected?

- (g) Were people more organized in the past, or are they more organized now? Has a sense of collective solidarity or brotherhood increased or decreased?
- (h) Apart from this, are there cultural or religious associations in this *basti* that you are aware of? Are you part of these associations?

#### 5. Local Leadership: *Pradhans*, Political Party Workers, and Volunteers

- (a) Is there a *pradhan* or are there *pradhans* in this *basti*? Can you name them?
- (b) 2. Since when has this person been a *pradhan*? Is the *pradhan* elected?
- (c) What work does the *pradhan* or *pradhans* do in the *basti*?
- (d) Apart from *pradhans*, are there people from political parties who assist *basti* residents with issues? What kind of assistance do they offer?
- (e) (If speaking to a *pradhan*) Since when have you been a *pradhan* of this neighborhood? What work do you do as a *pradhan*?
- (f) What issues or grievances do people in the *basti* approach you with? What do you do to resolve these issues?

#### 6. Relationship with Elected Representatives

- (a) Have the MLA (Member of Legislative Assembly), Municipal Councillor, and/ or MP (Member of Parliament) visited this *basti*?
- (b) Have you ever approached the MLA, Municipal Councillor, and/ or MP with individual problems or problems of your neighborhood or problems of the *basti*? Were they able to help resolve your grievances?

### 1.3 Survey Administration

Before administering the survey, we piloted the instrument in three *bastis*. This allowed us to make several adjustments to improve the clarity and effectiveness of the instrument. The survey

was administered by a local firm, Across Research and Communications (ARC). Together with an administrator from the survey firm and representatives from CPR, we trained a team of 25 enumerators on the instrument, recruitment strategy, and location characteristics (See Table4).

Table 4: Piloting and Enumerator Training



Piloting the Survey



Training Enumerators

Enumerators worked in teams of 3 to 5 individuals, each of which were assigned to portions of the settlement communities where they carried out the survey. Because of the extreme heat during the survey period and the work schedules of residents of the settlement communities, the enumerators mainly carried out surveys in the morning and early afternoon. They were instructed to approach every third household until all households in an area had been approached, and to interview only one person per household. When approaching a residence, they first introduced their company and the topic of the survey. They then asked to speak to the person in the house responsible for financial and other major household decisions. If this individual was not available, they asked for someone in the house who could make such decisions in the absence of the head of household. If neither was available, then the enumerator attempted to come back a different time. Because of this approach and social norms and practices in these communities, our sample is overwhelmingly (80%) male.

## 2 Survey Instrument: Key Measures and Experimental Manipulations

Our instrument draws upon a number of existing surveys, including the World Values Survey (Inglehart et al., 2012), surveys conducted for the The State and Citizens: A Study of Delhi's Slum by the the Center for the Study of Developing Societies (CSDS)(Lokniti, 2012), National Election Survey (NES) by CSDS (Lokniti, 2018), State of Democracy in South Asia conducted by CSDS (Lokniti, 2008), and instruments developed by the Program on Governance and Local Development (GLD) (Lust et al., 2020). Specifically, questions on political trust and political participation were adapted from the World Values Survey, National Election Survey, and State of Democracy in South Asia. Questions on religious participation, caste, financial assets were based on the National Election Survey templates. Question of social ties and networks relied heavily on GLD questionnaires.

### 2.1 Outcome Measures

We use five questions to measure outcomes from our experiment, all of which are answered on a 1–4 scale (with 4 indicating more interest in the drainage cleaning program):

- Would this program be beneficial for your neighborhood?
- How interested would you be in the program?
- How likely would you be to pay the monthly fee?
- Would you be willing to sign a six-month contract for this service?
- How likely would you be to try to get your neighbors to sign up for the program?

We combine these measures into an index by averaging a respondent's answers across the five questions. For those who do not answer all of the questions, we average the questions they did respond to. Figure 5 in our robustness checks shows the distribution of this index, which is trimodal

because of small floor and ceiling effects (which we address in our robustness checks). These five questions have a Cronbach's alpha of 0.89, indicating a high internal consistency for this index.

## 2.2 Measures of Religion and Religiosity

Eleven questions address religion and religiosity in our survey. (See Figure 1 for a graphical representation of these questions.) The first two are simple questions about religious affiliation:

- What is your religion?
- *If respondent is a Muslim:* To which sect do you belong, Shia or Sunni?

For the first question, respondents are asked to pick from the following list: Hindu, Muslim, Christian, Buddhist/Neo-Buddhist, Jain, or no religion, and can specify their religion if it is not on the list. Because of our choice of neighborhoods, the vast majority of respondents approached was either Hindu or Muslim. Those from other religions were dropped from the survey. For the second question, respondents were asked to pick Shia or Sunni (almost all indicated they were Sunni).

After this, respondents were asked five questions about religious activities and practices:

- Please tell me, how often do you practice these religious activities?
  - Prayer/puja/namaz
  - Visiting temple/mosque/church/gurudwara
- And what about these, how often do you practice them?
  - Participating in kathas, sangats, bhajan-kirtans, jalsas, church services, etc.
  - Giving donations for religious activities
  - Keeping fast, rozas, etc.

For the first two of these questions, respondents were asked to respond with one of the following: daily, 1-2 times a week, on festivals, during periods of distress, or never. For the last three, they

were asked to respond with one of the following: Whenever I get a chance, sometimes, rarely, during periods of distress, or never.

Elsewhere in the survey, respondents were asked a series of questions about local social ties, including the following two questions measuring religious ties:

- We would like to understand how people interact in your area. How obligated are you to help the following people, even if it costs you a day's wages?
  - A member of your religion
  - Someone associated with the day-to-day functioning of their temple/mosque/gurudwara

Respondents stated their level of obligation on a three-point scale: Very obligated, somewhat obligated, or not obligated at all.

Lastly, two questions asked respondents about their involvement in local organizations and with local elites:

- Now I will ask you about some important people. Please indicate if you are related to or a close friend of the following.
  - Religious leader in your area
- Have you or a member of your household participated in any of the following voluntary organizations?
  - Religious organization

For the first question, respondents indicated that they were the relative of, close friend of, both relative and close friend of, or had no relationship with a local religious leader. For the second, they responded with a simple yes or no.

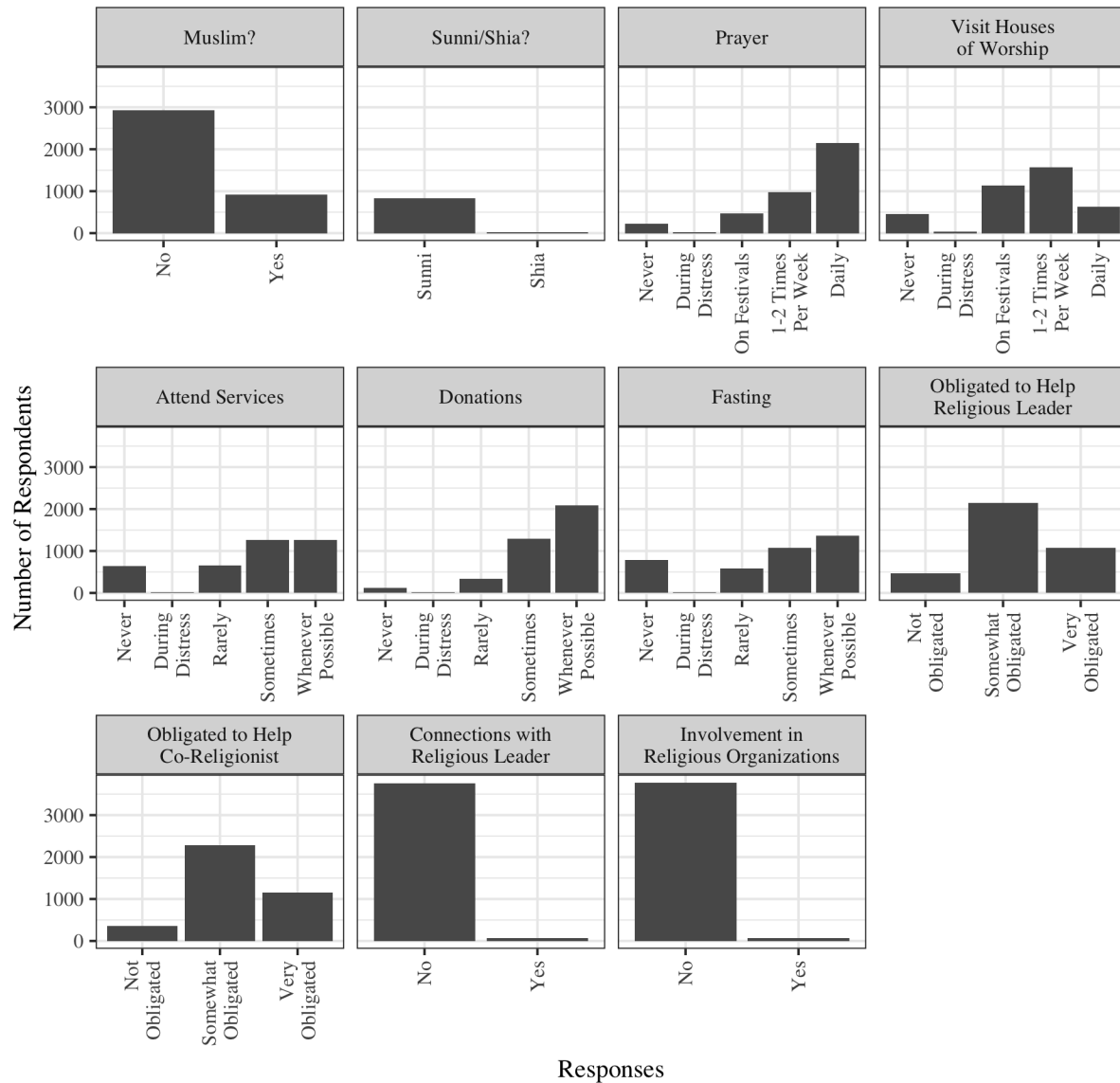


Figure 1: Measures of Religion and Religiosity (low total count in the “Sunni/Shia?” panel is due to the fact that only Muslims responded to that question).



## 2.3 Measures of Political Trust

The survey includes six measures of political trust. All of these questions are in the same section of the survey and follow a common introduction:

- I'm going to name a number of institutions. For each one, please tell me how well you think they perform their responsibilities?
  - Prime Minister and Central Government
  - Chief Minister and Delhi Government
  - The MCD
  - Political Parties
  - The Police
  - Your Local *Pradhan*

Respondents rated these institutions on a 1-4 scale: Not well at all, somewhat well, mostly well, and extremely well.

In our control models, we include an index of all six of these measures. However, the index does not have very good internal consistency (Cronbach's  $\alpha = 0.57$ ). We also explored the creation of factors from these items; however, the number of factors suggested by parallel analysis (two) produced factors that were not substantively useful. In exploring mechanisms, we therefore look at measures individually, focusing on local institutions that are of importance to drain cleaning and maintenance: evaluations of the MCD and the local Pradhan. However, the Pradhan question exhibits a high degree of missingness. Other questions only have between 25 (the MCD) and 266 (political parties) missing responses, but the Pradhan question has 1,518 (~40%) missing responses because many respondents did not know enough about their Pradhan to evaluate them.

## 2.4 Measures of Social Ties

In the survey, 16 questions relate to social ties with others in their neighborhood. The first is a standalone question that is not incorporated in an index:

- In your neighborhood, would you say:

With a three-point scale of possible responses: Few people know each other, some people know each other, or most people know each other.

The second social ties question is an independent question about the frequency with which respondents talk to their neighbors:

- How often do you talk to your neighbors?

With a five-point scale of possible responses: Never, only when necessary, rarely, sometimes, or daily.

Respondents are then asked to answer a set of five questions that form an index measure:

- When asked for help, do your neighbors help you with the following issues?
  - Childcare
  - Finances (help with small loans, etc)
  - Obtaining government services or documents
  - Solving disputes among family members or neighbors
  - Keeping the area clean

Respondents could answer one of the following on a four-point scale: Never, rarely, sometimes, or always. This index had a high internal consistency, with a Cronbach's alpha of 0.86.

Lastly, we included one additional nine-question index of social ties, this one measuring individuals' feeling of obligation to others in their area:

- We would like to understand how people interact in your area. How obligated are you to help with the following people, even if it costs you a day's wages?
  - A member of your immediate family (father, mother, children)
  - A member of your extended family (uncle, aunt, grandparent)
  - A neighbor
  - A distant relative
  - A member of your caste
  - A member of your religion
  - People who come from the same town/village
  - A fellow work mate/co-worker
  - Someone associated with the day-to-day functioning of their temple/mosque/gurudwara

Respondents could answer one of the following on a three-point scale: Not obligated at all, somewhat obligated, and very obligated. This index had a high internal consistency, with a Cronbach's alpha of 0.93.

Unlike some other variables, these four measures – the two standalone questions as well as the two indices – varied little by the religious identity of respondents, as can be seen in Table 5.

Table 5: Social Ties by Respondent Religion

	Hindus		Muslims	
	$\bar{X}$	SD	$\bar{X}$	SD
Neighborhood Knows Each Other	0.86	0.27	0.90	0.23
Talk to Neighbors	0.90	0.19	0.93	0.16
Neighbors Help Me	3.11	0.79	3.12	0.78
Obligation to Others	2.38	0.46	2.35	0.42

## 2.5 Measure of Caste

We used an open-ended survey question to gather data on respondent caste identification, based on *jati* or *zat*. A closed-ended question would have been too limited for the variety of castes. (Our enumerators entered over 800 unique text strings to record responses to this question!) We use the responses to categorize Hindus into three broad caste groups – upper castes (*Brahmins*, *Kshtriya*, and *Vaishya*), Other Backward Classes (OBC) (*Shudra*), and Lower Castes (Scheduled Castes and Scheduled Tribes). Muslims are categorized into two groups – upper castes (*Syed*, *Sheikh*, *Pathan*, and *Mughal*) and lower castes. The category of OBC Muslims, though administratively relevant in some states, is substantively less meaningful. Unlike Hindu society where caste reflects ideas of purity and pollution, caste or *zat* among Muslims is largely based on lineage. Scholars have categorized Muslims who claim foreign origin or are converts from upper caste Hindu groups as upper caste, while lower caste Muslims are generally converts from low-caste Hindus (Ali, 2002; Gayer and Jaffrelot, 2012).

After administering the survey, we identified the caste of all the unique responses based on the caste code protocol developed by the National Election Survey (NES), carried out by the Center for the Study of Developing Societies (CSDS; see Table 6). This resulted in 1,368 respondents identified as belonging to a low caste and 2,366 identified as belonging to a high caste. The proportion of caste membership based on this dichotomous categorization did not significantly vary by respondents' religious identification; approximately 37% of Hindus and 36% of Muslims were categorized as low caste (See Table 6). For the main models, we categorize individuals into two main groups – upper caste (upper caste Hindus and Muslims), and lower caste (OBC and lower caste Hindus and lower caste Muslims).

Table 6: Coding Protocol for Caste

	Groups	Sub-castes/ <i>Jati</i>
Hindu	Upper	Forward Caste and Brahmin; Bhumihar; Lohana; Rajput/Thakur; Kayasthas; Vaishya/Bania; Thondai, Mandala, Saiva, Vellala; Jain; Punjabi Khatri; Sindhi; Karana; Other Upper Castes
	OBC	Jat; Reddy; Kamma; Nair; Marathas; Patel (Karwa, Leuva, Patidar); Raju; Velama; Kapu; Bunt; Naidu; Bhuyan; Rai Bahadur; Lingayats; Thigala; Gujjar; Thevar; Ahir; Yadav; Kurmi; Mudaliars; Gowda; Lodh; Vanniyars; Munnuru Kapu; Koeri; Mutharayars; Mudiraj, Mutraju, Tenugollu; Vokkaliga; Kalinga; Lingayat; Thurpu Kapu; Gaderia; Koppulu Velama; Kunbi; Maratha-Kunbi; Koli; Charan; Rabari; Bharwad; Kshtriya; Chaudhary; Senai; Nadars; Ahom; Koch; Dhangar; Vanjari; Leva-Patil; Gowari; Agri; Powar; Teli; Kashyap; Kamboj; Mali/Saini; Bunkar; Gaderia; Darzee; Thatihar; Lakhera; Badhai; Kumhar; Lohar; Sunar; Medara, Mahendra; Kumar; Namasudra, Dami; Kosti-Sali-Padmashali; Kewat; Dhobi; Nai; Teli; Jogi; Trader OBCs (Jaiswal, Suri, Kalwar, Kanu, Halwai, Modi, Kasera, Tamoli, Tati); Toddy Tappers; Landless Labourers; Entertaining Castes; Veluthedathu Nair; Vathi; Chutiya, Koch; Vaishya; Banik/Bania/Bjujel; Dewan/Khatri; Gurung; Jogi; Kirat Rai; Mangar; Sunwar; Thami; Bahun; Chettri; Newar; Sanyasi; Jhimar; Deh, Bhat, Bharbhuj, Pinje; Other OBCs
	Lower	Chamar/Jatav; Satnami; Balmiki/Bhangi; Pasi; Pano; Devendrakula Velars; Rajbanshis; Mala; Namashudras; Mahar; Boyar/Mang; Dhobi, Julaha, Kewat; Dom; Nomadic Castes; Shilpkar; Musahar; Mayavanshi; Vankar; Dhanuk; Gond; Khatik; Kori; Thiruvalluvar; Adi Karnataka; Adi Dravida; Banjara; Bhambi; Bhovi; Chalavadi; Holaya; Korama; Mala Hannai; Pulaya; Paraya; Kuruva; Vettuva; Velan; Thandan; Mali; Koibarta; Das; Hadi; Damai/Musahar; Bhuian; Manjhi/Nag; Matang/Sarki; Basith; Megh; Bagdi Duley; Bauri; Jalia Kaibartta; Jhalo Malo Malo; Kaora; Lohar; Mal; Pod Poundra; Sunri; Other Scheduled Caste; Mina; Bhil; Gond/Rajgond; Oraon; Kamars; Santhall; Munda; Andh/Kondh; Naikda; Baiga; Bhattra; Binjhwar; Halba; Karku; Kavar; Korwa; Nagesia; Sawar; Ho; Kharia; Bhumij; Chero; Kharwar; Lohra; Mahli; Mal Pahariya; Mahali; Patelia; Baiga; Bhilala; Mahadev-Koli; Korku; Barela; Bhotia; Buksa; Jamnsari; Tharu; Malayali; Todas; Brus; Chakma; Kuki Tribes; Angami; Ao; Chakhesang; Chang; Khi-amniungan; Konyak; Lotha; Phom; Rengma; Sumi; Yimchungru; Zeliang; Bhutia; Lepcha; Garo; Jamatias; Mog; Other Scheduled Tribes
Muslim	Upper	Ashrafs Muslims (Sayyad, Shaikh); Mughal Muslims; Pathan; Muslim Rajputs; Other Upper Caste Muslims
	Lower	Pasmanda Muslims; Ansari; Qureshi; Qassab; Other lower caste Muslims

Table 7: Caste, by Participant Religion

	Lower Caste	Other Backward Class (OBC)	Upper Caste	Unable to Categorize
Hindu	1,290 (44%)	931 (32%)	672 (23%)	31 (1%)
Muslim	352 (36%)	-	555 (60%)	12 (1%)

Percentages calculated by row and don't add to 100% due to rounding.

## 2.6 Geocoordinates and Diversity Measure

Because we were interested in how local diversity affected the treatments in our survey, we used two methods to collect geolocation data from participants as the survey was administered. First, the tablets used to administer the survey automatically recorded the enumerator's geolocation at the beginning of the survey. Second, we instructed enumerators to use Google Maps to find their current location and record their coordinates in the notes section at the end of the survey. In piloting the survey, we found that using both approaches simultaneously yielded the most reliable geolocation data. In assigning geolocations to respondents, we default to the manually-entered Google Maps coordinates and use the automatically-recorded coordinates as a fallback.

However, cellular service can be spotty in Delhi. This sometimes resulted in inaccurate geolocation coordinates, with the enumerator either unable to access the Google Maps application to record their location or the coordinates providing the location of the closest cell phone tower or the center of town rather than the enumerator's precise location. Approximately 7% of our surveys had evidence that this occurred. For these cases, there were often coordinates available from the previous or next respondent interviewed by the same enumerator (or the same device) on the same day. We averaged these to provide approximate coordinates for most of the 7% of respondents who did not have reliable coordinates through the primary data collection. We deemed it better to provide an approximate estimate rather than eliminate these data. For five respondents, none of these methods provided reliable coordinates. Results that depend on geolocation data do not significantly vary if we exclude manually-edited coordinates.

## 2.7 Experimental Manipulations

The text of the experimental manipulations was as follows:

### **1: Control** *Our basis for comparison with the treatments.*

This service was offered in and around the neighborhood of a resident of Delhi. He shares his experience:

“Everything they told us about the service is true. Because of them, the drains are in a much better condition and there are no blockages due to garbage. The company is quick to respond in case people have any maintenance needs. But the most difficult aspect of this service is to get enough people to contribute. There were many residents in my neighborhood—among them, Sunil and Abdul—who did not contribute, and because of this the service could not be implemented in my neighborhood. But in the adjoining neighborhood, many residents did contribute, such as Salim and Mahesh, and the service was implemented as a result.”

If this service were implemented in your neighborhood, all information about you, including your name, address, and whether or not you contribute would remain anonymous and will not reported to the other residents in your area.

### **2: Horizontal Accountability** *Same as control until the final paragraph.*

... Once this service is started, community members would discuss amongst themselves who contributed and who did not pay the monthly fee. For example, in one of the neighborhoods that recently adopted this service, community members singled out and shamed the following non-contributors:

Javed, Vikas, Anwar, Gaurav, Quayum, and Sachin

### **3: Vertical Accountability** *Same as control until the final paragraph.*

... Once this service is started, the local leader would inform local residents of the benefits of hiring the company and would follow up personally with some of those who did not contribute money to the startup fee. For example, in one of the neighborhoods that recently adopted this service, the local leader personally visited the following non-contributors and shamed them.

Javed, Vikas, Anwar, Gaurav, Quayum, and Sachin

**4: Ingroup Underperformance (Hindu/Muslim)** *Differs from control in the italicized portions of the following paragraphs. Portions of the treatment vary based on the religious identification of the participant; In the paragraphs below, religion-dependent portions are displayed in the format Hindu/Muslim.*

This service was offered in and around the neighborhood of a resident of Delhi. He shares his experience:

“Everything they told us about the service is true. Because of them, the drains are in a much better condition and there are no blockages due to garbage. The company is quick to respond in case people have any maintenance needs. But the most difficult aspect of this service is to get enough people to contribute. There were many residents in my neighborhood *of Saraswatina-gar/Muhammadpur*—among them, *Sunil/Salim* and *Mahesh/Abdul*—who did not contribute, and because of this the service could not be implemented in my neighborhood. But in the adjoining neighborhood *of Muhammadpur/Saraswatinagar*, many residents did contribute, such as *Salim/Sunil* and *Abdul/Mahesh*, and the service was implemented as a result.”

If this service were implemented in your neighborhood, all information about you, including your name, address, and whether or not you contribute would remain anonymous and will not reported to the other residents in your area.

**5: Horizontal Accountability and Ingroup Underperformance (Hindu/Muslim)** *Same as the ingroup underperformance treatment until the final paragraph.*

... Once this service is started, community members would discuss amongst themselves who contributed and who did not pay the monthly fee. For example, in one of the neighborhoods that recently adopted this service, community members singled out and shamed the following non-contributors:

Prakash/Qasam, Gaurav/Jamal, Sachin/Quayum, Manish/Javed, Vikas/Iqbal, Arun/Anwar



**6:Vertical Accountability and Ingroup Underperformance (Hindu/Muslim)** *Same as the ingroup underperformance treatment until the final paragraph.*

... Once this service is started, the local leader would inform local residents of the benefits of hiring the company and would follow up personally with some of those who did not contribute money to the startup fee. For example, in one of the neighborhoods that recently adopted this service, the local leader personally visited the following non-contributors and shamed them.

Prakash/Qasam, Gaurav/Jamal, Sachin/Quayum, Manish/Javed, Vikas/Iqbal, Arun/Anwar

### **3 Main Analysis and Design**

#### **3.1 Models without Interactions**

Table 8 shows the results of the experiment on the outcome index and each of the separate outcomes, with no interaction effects. For this table, the treatments have been combined as they were for all analyses in the paper. Table 9, on the other hand, shows the results with each separate treatment.

For the sample as a whole, these tables show uniformly null effects of the combined treatments as well as for the separate treatments.

Table 8: Model with Combined Treatments and No Interactions

	Dependent Variable:					
	Benefit	Interest	Fee	Contract	Influence	Index
Constant	3.02 (0.02)	2.95 (0.02)	2.74 (0.02)	2.46 (0.02)	2.51 (0.02)	2.73 (0.02)
Treatments (Comb.)	0.02 (0.03)	-0.02 (0.03)	0.04 (0.03)	0.04 (0.03)	0.04 (0.03)	0.03 (0.03)
Observations	3,801	3,818	3,804	3,769	3,718	3,832
R <sup>2</sup>	0.0002	0.0001	0.0003	0.0004	0.0004	0.0003
Adjusted R <sup>2</sup>	-0.0001	-0.0001	0.0001	0.0001	0.0001	0.0000
Residual Std. Error	1.34 (df = 3799)	1.27 (df = 3816)	1.41 (df = 3802)	1.43 (df = 3767)	1.47 (df = 3716)	1.17 (df = 3830)
F Statistic	0.64 (df = 1; 3799)	0.48 (df = 1; 3816)	1.21 (df = 1; 3802)	1.42 (df = 1; 3767)	1.31 (df = 1; 3716)	1.11 (df = 1; 3830)

Table 9: Model with Separate Treatments and No Interactions

	Dependent Variable:					
	Benefit	Interest	Fee	Contract	Influence	Index
Constant	3.02 (0.04)	2.95 (0.03)	2.74 (0.04)	2.46 (0.04)	2.51 (0.04)	2.73 (0.03)
Horizontal Acct.	0.04 (0.05)	-0.01 (0.05)	0.05 (0.06)	0.04 (0.06)	0.07 (0.06)	0.04 (0.05)
Vertical Acct.	0.03 (0.05)	-0.01 (0.05)	0.05 (0.05)	0.06 (0.06)	0.06 (0.06)	0.04 (0.04)
Ing. Und.	0.03 (0.05)	0.001 (0.05)	0.04 (0.06)	0.04 (0.06)	0.04 (0.06)	0.03 (0.05)
Ing. Und. + Hor.	0.01 (0.05)	-0.05 (0.05)	-0.02 (0.06)	0.001 (0.06)	-0.03 (0.06)	-0.01 (0.05)
Ing. Und. + Vert.	0.02 (0.05)	-0.02 (0.05)	0.06 (0.05)	0.06 (0.06)	0.05 (0.06)	0.04 (0.04)
Observations	3,801	3,818	3,804	3,769	3,718	3,832
R <sup>2</sup>	0.0002	0.0004	0.001	0.001	0.001	0.001
Adjusted R <sup>2</sup>	-0.001	-0.001	-0.0004	-0.001	-0.0001	-0.001
Residual Std. Error	0.94 (df = 3795)	0.90 (df = 3812)	0.99 (df = 3798)	1.00 (df = 3763)	1.02 (df = 3712)	0.81 (df = 3826)
F Statistic	0.15 (df = 5; 3795)	0.31 (df = 5; 3812)	0.68 (df = 5; 3798)	0.44 (df = 5; 3763)	0.95 (df = 5; 3712)	0.55 (df = 5; 3826)

### **3.2 Interactive Model with Separate Treatments and Outcome Measures**

In the main analysis in the paper, we opt to combine our five treatment groups. In the heterogeneous effects model, which we present in Table 10 and Figure 2, all treatments tend to have the same effect magnitude and direction, with a few exceptions. In particular, the Ingroup Underperformance + Vertical Accountability treatment has a particularly large effect on individuals' willingness to try to influence others to join the program. However, we think it is more fruitful to examine the overall pattern rather than one particular effect. If hypothesis testing, statistical significance would vary across the treatments and outcomes in this analysis, but point estimates for Muslims are always positive while for Hindus there is a clear null effect.

Table 10: Interaction Model with Separate Treatments

	Dependent Variable:					
	Benefit	Interest	Fee	Contract	Influence	Index
Constant	3.06 (0.04)	2.98 (0.04)	2.79 (0.04)	2.47 (0.05)	2.55 (0.05)	2.76 (0.04)
Muslim	-0.16 (0.08)	-0.12 (0.08)	-0.20 (0.09)	-0.06 (0.09)	-0.16 (0.09)	-0.13 (0.07)
Hor. Acct.	-0.02 (0.06)	-0.05 (0.06)	-0.01 (0.06)	-0.02 (0.06)	0.03 (0.07)	-0.01 (0.05)
Hor. Acct. x Mus.	0.25 (0.12)	0.12 (0.12)	0.23 (0.13)	0.23 (0.13)	0.16 (0.13)	0.19 (0.10)
Vert. Acct.	-0.03 (0.06)	-0.07 (0.06)	-0.04 (0.06)	0.01 (0.06)	-0.001 (0.07)	-0.02 (0.05)
Vert. Acct. x Mus.	0.20 (0.12)	0.24 (0.12)	0.34 (0.13)	0.21 (0.13)	0.24 (0.13)	0.23 (0.10)
Ing. Und.	-0.04 (0.06)	-0.05 (0.06)	-0.02 (0.06)	0.02 (0.07)	-0.03 (0.07)	-0.02 (0.05)
Ing. Und. x Mus.	0.27 (0.12)	0.19 (0.12)	0.23 (0.13)	0.04 (0.13)	0.26 (0.13)	0.20 (0.11)
I.U. + H.A.	-0.03 (0.06)	-0.10 (0.06)	-0.07 (0.06)	-0.01 (0.07)	-0.07 (0.07)	-0.05 (0.05)
(I.U. + H.A.) x Mus.	0.17 (0.12)	0.19 (0.11)	0.19 (0.13)	0.06 (0.13)	0.14 (0.13)	0.14 (0.10)
I.U. + V.A.	-0.03 (0.06)	-0.05 (0.06)	0.02 (0.06)	0.04 (0.06)	-0.04 (0.07)	0.003 (0.05)
(I.U. + V.A.) x Mus.	0.18 (0.12)	0.12 (0.12)	0.13 (0.13)	0.06 (0.13)	0.38 (0.13)	0.15 (0.11)
Observations	3,801	3,818	3,804	3,769	3,718	3,832
R <sup>2</sup>	0.002	0.002	0.003	0.002	0.004	0.002
Adjusted R <sup>2</sup>	-0.001	-0.001	0.0002	-0.001	0.001	-0.0004
Residual Std. Error	0.94 (df = 3789)	0.90 (df = 3806)	0.99 (df = 3792)	1.00 (df = 3757)	1.02 (df = 3706)	0.81 (df = 3820)
F Statistic	0.68 (df = 11; 3789)	0.66 (df = 11; 3806)	1.06 (df = 11; 3792)	0.78 (df = 11; 3757)	1.31 (df = 11; 3706)	0.85 (df = 11; 3820)

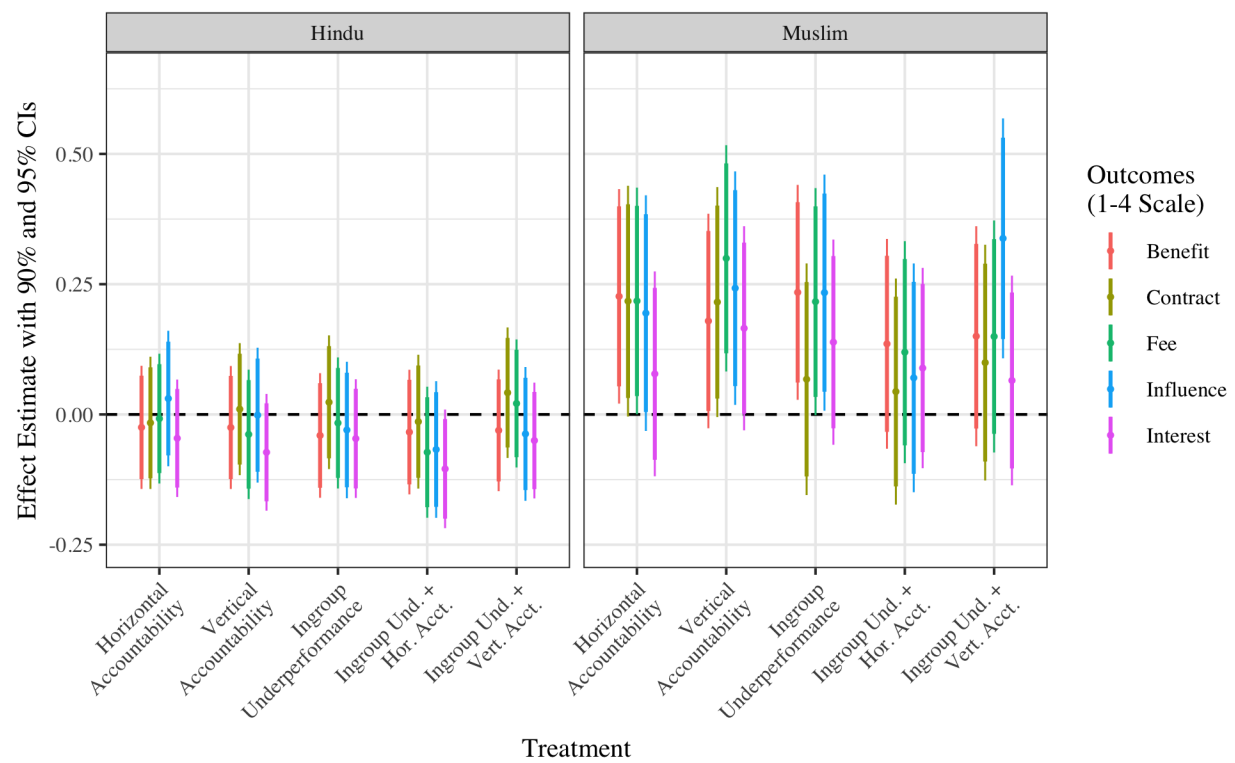


Figure 2: Interaction Model with Separate Treatments

### 3.3 Full Control Model

In the paper, we provide a condensed version of our model with controls. In Table 11 below, we provide the full model. Note that the coefficients on the combined treatment and the interaction terms vary only marginally between these specifications and are statistically significant at the 0.05 level in all models.

Readers will note that we do not include neighborhood or enumerator fixed effects in any of these models. If we were running a non-interactive model, including these fixed effects as control variables is not problematic. However, our main model interacts the treatments with religious identification, which makes these fixed effects problematic for our coefficient estimates of interest. As noted in the paper and elsewhere in this appendix, areas differ greatly in their religious composition, such that religious identification is highly correlated with neighborhoods. Enumerator fixed effects are problematic for a related reason: enumerators were not randomly assigned to areas, so the issue of high correlation with specific neighborhoods also affects enumerator fixed effects as a control. Evidence that this is problematic includes wild swings in coefficient estimates and significance when fixed effects are introduced, as well as a very large (>50) variance inflation factor. Given this limitation, we feel that the best way to account for important neighborhood and enumerator characteristics is to focus on the characteristics that are important to our analysis. For this reason, our robustness checks include an examination of how our main model estimates vary based on the religious composition of neighborhoods as well as an examination of how enumerator religious identification affects our analysis.

Table 11: Main Model with Controls

	Dependent Variable: Index of Favorability toward Drainage Program				
	(1)	(2)	(3)	(4)	(5)
Constant	2.76 (0.02)	1.00 (0.16)	1.01 (0.16)	0.95 (0.16)	0.96 (0.16)
Treatments (Combined)	-0.02 (0.03)	-0.00 (0.03)	0.00 (0.03)	0.00 (0.03)	0.00 (0.03)

Table 11 (Continued)

	Dependent Variable: Index of Favorability toward Drainage Program				
	(1)	(2)	(3)	(4)	(5)
Muslim	−0.13 (0.04)	−0.07 (0.04)	−0.06 (0.04)	−0.11 (0.04)	−0.10 (0.04)
Treatments (Combined) x Muslim	0.18 (0.06)	0.17 (0.06)	0.17 (0.06)	0.15 (0.06)	0.15 (0.06)
Gender (1 = Female, 0 = Male)		−0.01 (0.04)	0.00 (0.04)	0.00 (0.04)	−0.01 (0.04)
Married (1 = Married, 0 = Other)		0.02 (0.05)	0.00 (0.05)	0.01 (0.05)	0.01 (0.05)
Age		0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Years Residing in Settlement		−0.00 (0.00)	−0.00 (0.00)	−0.00 (0.00)	−0.00 (0.00)
Ownership (1 = Own Residence)		−0.00 (0.04)	−0.01 (0.04)	0.00 (0.04)	0.00 (0.04)
Number of Rooms in Residence		0.04 (0.01)	0.04 (0.01)	0.03 (0.01)	0.03 (0.01)
Employment Status (1 = Employed)		0.18 (0.04)	0.19 (0.04)	0.19 (0.04)	0.19 (0.04)
Fin. Hardship (1 = More Hardship)		−0.03 (0.06)	−0.05 (0.06)	−0.07 (0.06)	−0.08 (0.06)
Assets Index (1-3, 3 = Most Assets)		0.09 (0.03)	0.09 (0.03)	0.09 (0.03)	0.09 (0.03)
Education Level (1-10 Scale)		0.02 (0.01)	0.02 (0.01)	0.02 (0.01)	0.02 (0.01)
Practice of Prayer (1-5 Scale)		0.06 (0.01)	0.06 (0.01)	0.06 (0.01)	0.06 (0.01)
Religious Index (1-5 Scale)		0.05 (0.02)	0.05 (0.02)	0.04 (0.02)	0.03 (0.02)
Political Efficacy Index (1-4 Scale)		0.04 (0.02)	0.04 (0.02)	0.05 (0.02)	0.05 (0.02)
Relation to Local Figures (1 = Any)		−0.04 (0.04)	−0.04 (0.04)	−0.06 (0.04)	−0.05 (0.04)
Involve in Local Org. (1 = Any)		−0.02 (0.05)	−0.02 (0.05)	−0.03 (0.05)	−0.02 (0.05)
General Social Ties (0-1 Scale)		0.25 (0.08)	0.24 (0.08)	0.22 (0.08)	0.24 (0.08)
Helpfulness Index (1-4 Scale)		0.02 (0.02)	0.01 (0.02)	0.01 (0.02)	0.00 (0.02)
Forgo Wages, Family (1-3 Scale)		0.11 (0.04)	0.11 (0.04)	0.12 (0.04)	0.16 (0.05)
Forgo Wages, Others (1-3 Scale)		0.20 (0.03)	0.20 (0.03)	0.21 (0.03)	0.18 (0.03)



Table 11 (Continued)

	Dependent Variable: Index of Favorability toward Drainage Program				
	(1)	(2)	(3)	(4)	(5)
% Political Activities Engaged In		0.15 (0.05)	0.17 (0.05)	0.19 (0.05)	0.19 (0.05)
Quality of Drainage (1-5)		-0.10 (0.01)	-0.11 (0.01)	-0.10 (0.01)	-0.10 (0.01)
Drainage Problem Requiring Help		-0.12 (0.03)	-0.12 (0.03)	-0.12 (0.03)	-0.10 (0.03)
Caste Level (1 = Upper, 0 = Lower)			0.03 (0.03)	0.02 (0.03)	0.02 (0.03)
Diversity				0.24 (0.05)	0.23 (0.05)
Enumerator Religion					-0.30 (0.07)
Observations	3,832	3,684	3,577	3,555	3,549
R <sup>2</sup>	0.00	0.10	0.10	0.11	0.11
Adjusted R <sup>2</sup>	0.00	0.09	0.09	0.10	0.10
Residual Std. Error	1.16 (df = 3828)	1.08 (df = 3658)	1.08 (df = 3615)	1.08 (df = 3592)	1.07 (df = 3585)
F Statistic	3.90 (df = 3; 3828)	15.69 (df = 25; 3658)	14.90 (df = 26; 3615)	15.45 (df = 27; 3592)	15.57 (df = 28; 3585)

### 3.4 Pre-Registration

We pre-registered the design of our experiment with EGAP, now migrated to Open Science Foundation. Crucially, we pre-registered that we would look at heterogeneous effects for Muslims and Hindus because we anticipated that Muslims, as a minority, might respond differently to our intervention. However, we found it necessary to make several adjustments from our pre-registered design in the main analysis reported in our paper.

*First:* Instead of using outcome measures separately, we formed an index that we used as our primary outcome measure. We used an index because it allowed us to simplify the number of tests we were conducting, though we still include results with separate outcomes in the main paper.

*Second:* We combined our five treatment groups—horizontal accountability, vertical accountability, ingroup underperformance, and the two combination treatments—into a single treatment. We did this for a few reasons. As seen in Section 3.2 of the Appendix, we saw broadly similar results across the different treatments. Additionally, combining treatments made it easier to explore mechanisms. With five treatments and numerous mechanisms to test, the number of tests would increase quickly if we had not combined the treatments. And, with fewer Muslims in the sample than Hindus, looking at mechanisms underlying the differing responses of Muslims and Hindus quickly becomes untenable because of the number of participants.

*Third:* We report some p-values with corrections for multiple comparisons in the paper, while in the pre-registration we did not specify corrections that we would use.

*Fourth:* We do not include all treatment arms of the experiment in the analysis in this paper. We include the relevant controls and treatment to the current analysis only.

*Fifth:* Our pre-analysis plan focused primarily on diversity, not on minority status, as the factor that we would use to explore heterogeneous effects. However, we focus on minority status for a few reasons. First, it quickly became clear to us that diversity and minority status overlap by virtue of the relative sizes of groups. This is in line with recent research that has come out since the planning stages of our project (Kustov and Pardelli, 2018). Additionally, we found more evidence for minority status having an effect as opposed to diversity (see Section 4.6).

### 3.5 Ethics

Our research project adheres to all APSA Principles and Guidelines for Human Subjects Research and was approved through Harvard University Institutional Review Board (IRB18-0258). Aspects of this research project caused us to consider four principles in particular: Power, Consent, Deception, and Confidentiality. We discuss each of these below.

*Power:* Because the participants in our study are from a low-power community, we carefully considered how to structure our survey to respect participants' autonomy and not cause them harm. For instance, we (1) designed our survey to be as short as possible, (2) carefully constructed the religious cues in our treatments to be subtle and not exacerbate community tensions, and (3) clearly communicated the hypothetical nature of the drainage cleaning initiative multiple times during the survey. We believe that these steps produced a survey that respected the valuable time of our participants, their desire to maintain peace and stability in their communities, and their right to accurate information about the nature of the proposed drain cleaning program.

*Consent:* Consent was informed and voluntary. No monetary compensation was offered, which might have made some participants feel pressure to participate. Because our surveys were administered via tablet, consent was given verbally to the enumerator rather than via signature.

*Deception:* Our survey manipulations outline a drain cleaning program that does not exist; however, we did *not* use deception in this aspect of the survey. It was clearly communicated to participants, multiple times, that the drain cleaning program was strictly hypothetical.

*Confidentiality:* Our survey collected geocoordinates, which is the only data we collected of a sensitive nature for identifying participants. In our replication data, we include only the diversity measure calculated from these coordinates, not the original geocoordinates.

## 4 Robustness Checks

### 4.1 Balance

In the paper, we combine treatment groups for our analyses. The lower panel in Figure 3 shows that, in this combined analysis, our sample is well-balanced between the control group and the group of the combined treatments. Standardized differences for the pre-treatment covariates all fall within a range of -0.75 to 0.75. An omnibus balance test (Hansen and Bowers, 2008), reported in the final row of Table 12, also indicates that we cannot reject the null of a balanced sample for the combined treatments.

We also address balance by examining the treatment groups separately rather than combined. The upper panel in Figure 3 shows the standardized differences for pre-treatment covariates between the control and each of the treatment groups. Only a few variables show minor imbalances: Gender, Number of Rooms in Residence, the Religious Behavior Index, Religious ID, and the Forgo Wages Index. However, the omnibus balance tests in rows 1-5 of Table 10 indicate that these minor imbalances are not unexpected for our randomization and we cannot reject the null of a balanced sample.

Table 12: Omnibus Balance Tests

Treatment	$\chi^2$	DF	P-value
Horizontal Accountability	32.14	31	0.41
Vertical Accountability	23.91	32	0.85
Ingroup Underperformance	26.95	31	0.67
Horizontal Accountability + Ingroup Underperformance	15.67	31	0.99
Vertical Accountability + Ingroup Underperformance	39.47	31	0.14
Treatments Combined	30.57	34	0.64



Figure 3: Comparison of Baseline Characteristics between Control and Treatments

## 4.2 Survey Purpose and Manipulation Checks

In order to make sure that participants received and understood the information they were presented about the proposed drainage service, they were asked two questions immediately after the administration of the survey. First, participants were asked: “How many people in your neighborhood need to agree to pay the fee before the company will begin the new service?” They were read three options to choose from: (a)  $1/4$  or 25%, (b)  $1/2$  or 50%, and (c)  $2/3$  or 67%. These choices mirror the way these proportions were presented in the experimental manipulation itself, where participants were told that a condition for the drainage cleaning program was that “ $2/3$  or 67% of the residents in your neighborhood indicate their agreement with this contract.” Both fractions and percentages were used because pretesting revealed that both were common ways of expressing proportions. A total of 3,416 respondents, or 88.9% of the sample, passed the manipulation check. Those who did not were told the correct answer after their response to make sure that they completed the rest of the survey with the correct conditions in mind.

Second, participants were asked: “What is the monthly subscription fee required by this service?” They were read three options as answers: (a) Rs. 50, (b) Rs. 100, and (c) Rs. 150. A total of 3,651 respondents, or 95.0% of the sample, passed this manipulation check. Those who did not were told the correct answer after their response.

We present results in the paper *without* excluding those who failed the manipulation check. However, results are largely the same if we exclude these respondents, as shown in Table 13. If anything, the results excluding those who failed the manipulation checks are stronger, with a coefficient estimate that is 0.035 larger than the estimate in our main analysis with everyone included. This may indicate that the effects are stronger for those who absorbed the most information from the treatment administration.

Table 13: Main Analysis, Excluding Those Who Failed Manipulation Checks

	Dependent Variable: Index of Favorability toward Drainage Program
Constant	2.84 (0.02)
Treatments (Combined)	−0.04 (0.03)
Muslim	−0.17 (0.04)
Treatments (Combined) x Muslim	0.22 (0.06)
Observations	3,387
R <sup>2</sup>	0.005
Adjusted R <sup>2</sup>	0.004
Residual Std. Error	1.15 (df = 3383)
F Statistic	5.22 (df = 3; 3383)

### 4.3 Timing of Survey Administration and Treatments

Because our survey was administered using tablets, we have detailed information on the amount of time it took to administer each portion of the survey. It took, on average, 13.5 minutes to administer the survey, although this average varied from a low of 9 minutes to a high of 16.5 minutes depending on the enumerator.<sup>1</sup>

Out of the 13.5 minutes on average spent in the survey, less than half a minute was spent on the experimental manipulation. Enumerators spent just under 17 seconds on average explaining the drainage cleaning program in general, which did not vary between treatments, and 7 seconds on average explaining the portions of the program that varied between treatments. This second portion varied only slightly, by less than 0.75 seconds, between the different treatment texts (see Table 14). This indicates that, in terms of time spent completing the experimental manipulation, there is little difference between the treatments.

Table 14: Length of Treatment Assignment

Treatment	Average Time Administering Treatment (Seconds)
Control	6.95
Horizontal Accountability	6.77
Vertical Accountability	7.32
Ingroup Underperformance	7.38
Hor. Acct. + Ingroup Underperformance	7.36
Vert. Acct. + Ingroup Underperformance	6.67

<sup>1</sup>This calculation includes the twenty enumerators who administered approximately 93% of the surveys; it excludes the five enumerators who each administered 30 surveys or fewer, within which there was more variation (minimum of 9 minutes, maximum of 32 minutes).



## 4.4 Factor Analysis with Religious Practice Variables

In the paper, we use a set of indices to measure treatment heterogeneity in terms of religious private practice, public practice, and social ties. However, we considered an alternative approach using factor analysis. In this section, we provide a description of that alternative approach, its similarities and differences to the approach we present in the paper, and how an approach using factor analysis also provides strong evidence for the mechanism outlined in the paper.

As a reminder, our survey included nine questions focused on religion. Five of these focused on religious practice and activities, while the other four focused on religious social ties and connections to religious institutions. For a more in-depth discussion of these questions and descriptive statistics, refer to Section 2.2 in this appendix. Importantly, from a theoretical perspective, we anticipate that our experimental results will vary based on some aspects of religion but not others. Specifically, we expected those with strong ties to their religious community to drive the treatment effect we saw among Muslims, since members of the minority with strong ties to their group would be the most likely to respond to our treatments' primes of the prospect of social shaming.

However, we wanted the factor analysis to be data-driven to see if our theoretical expectations lined up with what appeared in the data. For this reason, when choosing the number of factors, we used parallel analysis as implemented in the `fa.parallel` function from the `psych` package in R. Parallel analysis compares the “scree” of factors of the data with that of a random data matrix of the same size as the original. One complication to this process was the variation in the types of questions we wanted to include in our factor analysis. Most of these variables are Likert-style items, and two of them – those regarding connections with religious leaders and involvement with religious organizations – are dichotomous. With such a set of questions, the regularly used Pearson correlation is biased from the latent correlation used to estimate factors. For this reason, we used a correlation matrix with mixed correlations consisting of polychoric correlations for the Likert-style items and tetrachoric correlations for the dichotomous items.

The scree plot that summarizes the results of this process can be found in Figure 4. This plot shows the eigenvalues (i.e., the variance explained) for each factor that could be extracted, both for the actual data as well as for simulated data. The figure clearly indicates that the factor analysis eigenvalues for the actual data are higher than those for the simulated data until we reach factor 5, at which point they are smaller, thus

## Parallel Analysis Scree Plots

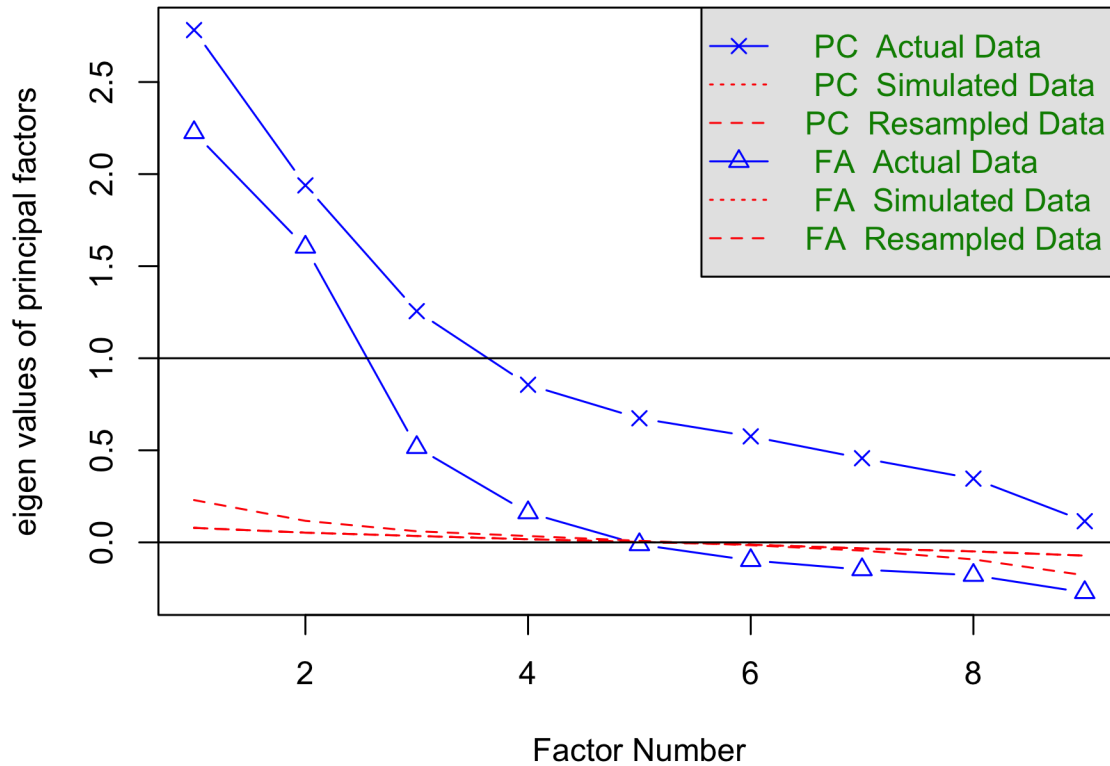


Figure 4: Scree Plot from Parallel Analysis of Questions on Religion

leading to the conclusion that a four-factor model is appropriate for these data.

We then run factor analysis with four factors, using varimax rotation, resulting in the loadings shown in Table 15. These factors clearly separate between the different sets of questions we asked. Factor 1, which we label as “Institutional Connectedness,” loads primarily on the questions regarding respondents’ connections with religious leaders and their involvement in religious organizations. This factor does not have a corollary in our paper, since few respondents actually had the connections with religious leaders or involvement with religious organizations when questioned, and we were not comfortable using questions with such low variation. Factor 2, which we label as “Obligation to Co-Religionists,” loads primarily on questions regarding respondents’ feelings of obligation to help religious leaders or other co-religionists. It also to a lesser extent loads on a measure of how often the respondent donates to their religion. This factor corresponds to the “Religious Social Ties” index used in the paper. Factor 3, which we label as “Participation

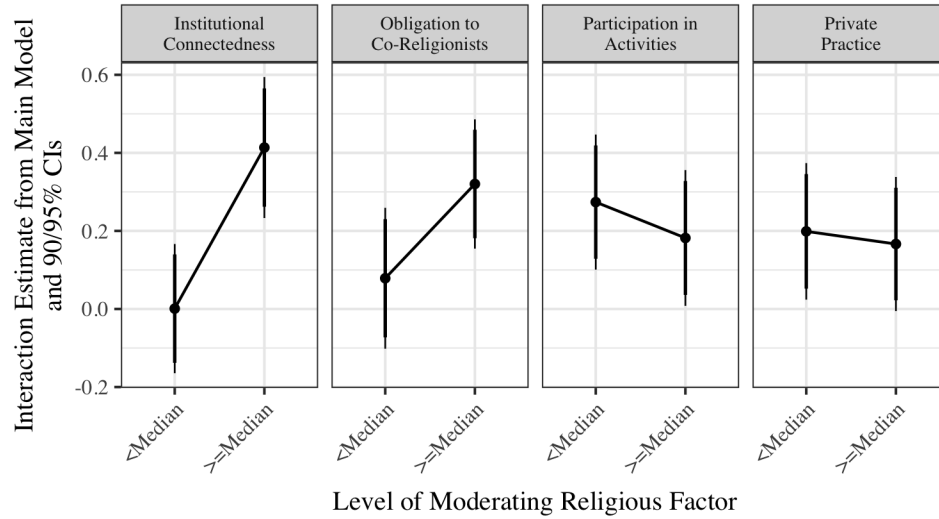


Figure 5: Estimate of Interaction Coefficient from Main Analysis, by Religiosity Factors

in Activities,” loads primarily on questions about attending houses of worship and services, donating, and fasting, with smaller loading on the question about frequency of prayer. This corresponds to the “Public Participation” index used in the paper. Factor 4, which we label “Private Participation,” loads primarily on just the question about prayer. In the paper, we use prayer by itself as another measure of religious practice.

With these factors, we can then conduct the same procedure used in the paper to evaluate a mechanism of defensive cooperation. This involves estimating our main model – the effect of the combined treatments, interacted with religious identification, on an index of our five outcome measures – separately among those with high and low values for each of these factors. If one of the factors explains the difference between Hindus and Muslims in our main model, then we would expect the interaction variable to vary greatly when estimated for those with different values of these factors. Using this method, we find results that are not only similar to those in the paper but stronger. We find that the Institutional Connectedness and Obligation to Co-Religionist factors differentiate for the interaction, as shown in Figure 5.

Table 15: Factor Loadings for Questions on Religiosity

	Factor 1	Factor 2	Factor 3	Factor 4
	Institutional Connectedness	Obligation to Co-Religionists	Participation in Activities	Private Practice
Prayer	-0.002	0.02	0.29	0.95
Visit Houses of Worship	0.13	0.01	0.48	0.28
Attend Services	0.18	-0.03	0.83	0.05
Donations	-0.02	0.25	0.57	0.14
Fasting	0.07	-0.04	0.73	0.10
Obligated to Help Religious Leader	0.13	0.99	0.08	0.02
Obligated to Help Co-Religionist	0.13	0.87	0.01	0.003
Connections with Religious Leader	0.39	0.07	0.10	0.0004
Involvement with Religious Organization	0.99	0.11	0.04	0.03

## 4.5 Political Trust as a Mechanism

When considering alternative explanations besides defensive cooperation that could explain the difference between Hindus and Muslims in their responses to the treatments, one explanation we particularly considered was differing levels of trust in political institutions. Given the role of the state in discrimination toward Muslims, it is reasonable to expect that Muslims would have lower levels of trust in political institutions. In turn, lower trust in the political institutions that deliver drainage service – local neighborhood leaders, i.e. *Pradhans*, as well as the Municipal Corporation of Delhi (MCD) – could also cause individuals to respond more to the treatments that we administered. Specifically, individuals with lower trust in these institutions may respond more to appeals based on horizontal or vertical pressure in particular because they are more likely to be open to collective action and market-based solutions. Note that, in contrast to defensive cooperation as a mechanism, we are not positing that political trust functions differently for Hindus and Muslims. Instead, this mechanism depends on differing levels of political trust between the groups. If this mechanism is at play, we would expect two empirical implications: First, lower levels of trust among Muslims in local political institutions involved in drain cleaning and maintenance; second, an interaction effect between the treatments and trust in these local institutions *that is in the opposite direction of the interaction in our main model*. We look at each of these individually and, for the most part, rule out this mechanism. It is possible that a lack of political trust in *Pradhans* drives the result we see for some of our participants. However, important factors point against this mechanism – numerous respondents were unfamiliar with the identity of their *Pradhan* and, in addition, there is little difference between Muslims and Hindus when it comes to trust in the MCD.

We address the first empirical implication – lower levels of trust among Muslims in local political institutions involved in drain cleaning and maintenance – in Table 16, where we examine differences in levels of trust between Hindus and Muslims for six salient political institutions and leaders: local neighborhood leaders called *Pradhans*, the Municipal Corporation of Delhi (MCD), the Prime Minister (at the time Narendra Modi), political parties, the police, and the Chief Minister of Delhi (at the time Arvind Kejriwal). Participants were asked to rate each of these institutions and leaders in terms of how effectively they carry out their responsibilities, on a 1-4 scale (higher numbers indicating more effectiveness/trust). Those familiar with the tense relationship between Muslims and the Indian state would find many of the results in Table 16 unsurprising. The biggest divergence in trust is not in institutions related to drain cleaning but, rather, in

Table 16: Levels of Trust in Political Figures and Institutions, by Religion

	Pradhan		MCD		Prime Minister		Parties		Police		Chief Minister	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Hindu	2.13	1.06	2.14	0.98	2.93	1.01	1.96	1.01	2.63	1.06	1.62	0.49
Muslim	1.83	0.99	2.12	0.97	2.07	1.01	1.65	0.89	2.19	1.02	1.62	0.49

the prime minister (a difference of 0.86, almost a third of our scale), the police (a difference of 0.44), and political parties (a difference of 0.31). Evaluations of the *Pradhan* also differ, with Muslims on average rating their *Pradhan* 0.3 lower than Hindus, a difference comparable to that for political parties. However, while *Pradhans* are often involved in requesting services from the state, such as drain cleaning, the main local political institution responsible for maintaining drains is the MCD, for which the difference between Muslims and Hindus is basically nonexistent and equivalent to the lack of difference in evaluations of the Chief Minister.

We address the second empirical implication of political trust as a mechanism – negative interaction effects between trust in local institutions and the treatments – by estimating a new interactive model for each of the six political trust measures. Figure 6 shows the marginal effect of the treatments for those below and above the median value for each measure. The pattern in these data is clear: The two local leaders/institutions directly involved in drain cleaning moderate the effect of the treatments, while other local, regional, or national political institutions and leaders do not. Although the lack of a difference between Muslims and Hindus in evaluations of the MCD effectively rules out that measure as a mechanism, it does independently moderate the effect of the treatments such that those with lower trust in the MCD exhibit a positive treatment effect while those with higher trust have a negative effect. Similarly, those with low trust in their *Pradhan* have a positive treatment effect, while those with high trust have a negative one. This indicates that the lower trust in *Pradhan*'s among Muslims, relative to Hindus, could drive the interactive effect of religion in our main model.

However, other factors regarding the relationship between *Pradhans* and their communities in this context indicate that we should question whether a lack of trust in *Pradhans* drives the higher treatment effect for Muslims. Though *Pradhans* have been shown to be important actors in local level collective action in Indian slums (Auerbach, 2016; Auerbach and Thachil, 2018; Auerbach, 2017; Chidambaram, 2020; Jha et al., 2007),

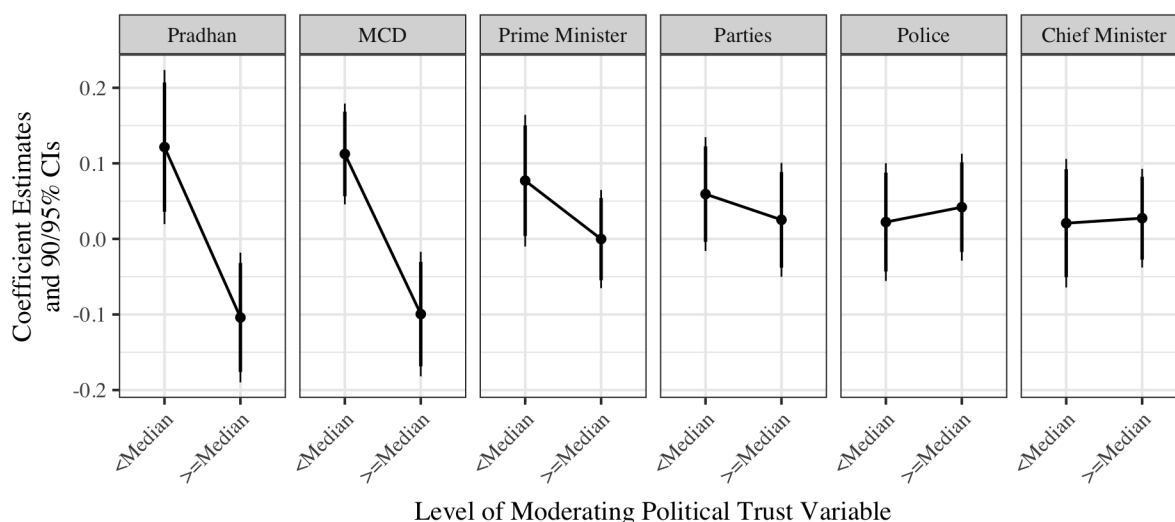


Figure 6: Marginal Effect of Combined Treatments, by Levels of Political Trust.

results from our survey show that the relationship between *Pradhans* and the community in our locations is more complicated. Overall, just under 40% of people did not even respond to our question about the political effectiveness of *Pradhans*; the rates for Hindus (35%) and Muslims (53%) separately indicate that this non-response is particularly bad for Muslims. None of the other political effectiveness questions even approached this level of non-response. The closest two are the questions about political parties (6.9% non-response) and the police (3.6%).

What is the cause of these extreme levels of non-response rates for this particular question? To answer that, we can look at another survey question, where we asked individuals to rate how influential the *Pradhan* in their neighborhood is. Participants were asked to rate the *Pradhan's* influence on a 1-4 scale, but enumerators were also instructed to record the reason for cases of non-response, with options for “do not know,” “refuse to answer,” and “does not apply (no *pradhan*).” A summary of responses to this question, broken down by religion, can be found in Table 17. The results indicate that only three participants refused to answer the question, while most participants either did not think their neighborhood had a *Pradhan* (Hindus 28%, Muslims 39%) or did not feel like they knew enough information about their *Pradhan* to rate how influential they were (Hindus 17%, Muslims 28%). This is strong evidence that a large proportion of our sample, and a majority of Muslim participants, either are not aware of a *Pradhan* serving in their neighborhood or know little to nothing about the *Pradhan*. This finding is also reflected in our qualitative fieldwork from these locations. Our qualitative fieldworkers recorded, for instance, that residents in *Settlement D* complained about

Table 17: Pradhan Influence, by Participant Religion

	Not at All Influential	A Little Influential	Somewhat Influential	Extremely Influential	Do Not Know	Refuse to Answer	No Pradhan
Hindu	738 (25%)	284 (10%)	398 (14%)	181 (6%)	509 (17%)	3 (0%)	811 (28%)
Muslim	150 (16%)	67 (7%)	58 (6%)	31 (4%)	255 (28%)	0 (0%)	358 (39%)

having not had a *Pradhan* for more than a decade. In another area, *Settlement A*, these fieldworkers found that even residents who were politically aware – for example, they knew the name of their state legislator – could not name their *Pradhan*.

These findings lead us to conclude that, in spite of lower trust among Muslims in their *Pradhans*, and the presence of a negative interaction effect between trust in the Pradhan and the treatments, this mechanism is unlikely to account for the interaction effect from our main model for most participants.



## 4.6 Diversity as an Alternative Hypothesis

The interaction in our main analysis is with religious identification – whether the participant was Muslim or Hindu. However, in our pre-analysis plan, we also pre-registered an analysis examining the interaction of the treatments with religious diversity. Because of a robust literature tying diversity to poor public goods provision, we anticipated finding that the treatments would be less effective in diverse locations. We thought that participants in more diverse areas might not respond as much to our treatment mechanisms because they would not care as much about accountability to peers or to local elites.

For our measure of diversity, we use the polarization index from Montalvo and Reynal-Querol (2005). This measure, labeled as RQ below, is represented by the equation

$$RQ = 1 - \sum_{i=1}^N \left( \frac{\frac{1}{2} - \pi_i}{\frac{1}{2}} \right)^2 \pi_i,$$

where  $\pi_i$  represents the proportion of the population that belongs to group  $i$ . This measure captures “how far the distribution of the... groups is from the  $(1/2, 0, 0, \dots, 0, 1/2)$  distribution (bipolar), which represents the highest level of polarization” (pp. 798). However, because our context only involves two groups, this measure simplifies to  $RQ = 4\pi_1\pi_2$ , where  $\pi_1$  and  $\pi_2$  are the group proportions, and it is almost identical to the ethnic fractionalization index for two groups ( $2\pi_1\pi_2$ ) that is also often used to measure diversity. Using the geocoordinates collected as a part of the survey, we calculate the polarization index for each participant based on the religious identification of other participants within 100 meters. This provides us with a near-continuous measure of diversity at the local level.

Using this measure, we find that there is no interaction with religious diversity. Results for our analysis with a diversity interaction effect can be found in Table 18 and Figure 7. The coefficient on the interaction between the combined treatments and the diversity index is 0.1 ( $SE = 0.09$ ), statistically insignificant and in the opposite direction of our hypothesis. Because diversity is a continuous variable, we also explore the interaction effect graphically in Figure 7. This figure shows the estimate for the linear interaction effect from Table 18, an estimate of interaction effects based on binning observations into tertiles from the diversity index, and a histogram at the bottom of the plot showing the distribution of the diversity measure (pink colored bars indicate observations in the combined treatment groups; gray bars indicate those in the control group). This figure confirms what we find in the regression model – that is, no significant interaction effect.

As noted in the paper, we think the most likely explanation for this finding is the correlation between being a minority (Muslim) and living in a “diverse” location, which exhibit a medium-sized correlation in our sample ( $r = 0.36$ ). By this account, respondents’ religious identification explains our finding that diversity has little interaction effect, and that whatever effects exists is in the opposite direction of our prior hypotheses and what the literature would expect.

Table 18: Model with Diversity Interaction

	Dependent Variable: Index of Favorability toward Drainage Program
Constant	2.71 (0.03)
Treatments (Combined)	−0.02 (0.04)
Diversity (100m)	0.04 (0.06)
Treatments (Combined) x Diversity (100m)	0.10 (0.09)
Observations	3,810
R <sup>2</sup>	0.002
Adjusted R <sup>2</sup>	0.001
Residual Std. Error	1.17 (df = 3806)
F Statistic	2.19 (df = 3; 3806)

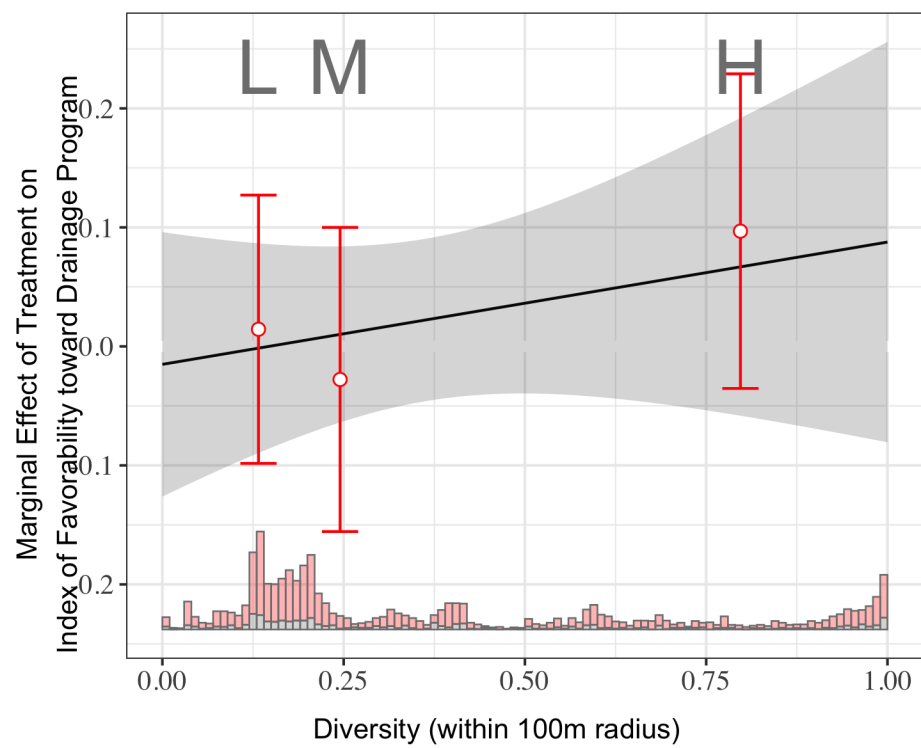


Figure 7: Graphical Representation of Interaction with Diversity

## 4.7 What is a Neighborhood?

Some might be concerned about using *bastis* to group people for the administration of our treatment. What if participants do not consider the residents of their *basti* to be their neighbors? To address this question, we included a question in the survey that asked participants what they considered to be their neighborhood. Participants were given three options: (i) the 1-2 streets closest to my place of residence, (ii) the 3-4 streets closest to my place of residence, and (iii) the entire settlement in which the person was surveyed. Approximately 67%, or two-thirds, of the participants chose the final option. This indicates that, although there is significant variation in how individuals define their neighborhood, most respondents consider their entire *basti* as their neighborhood.

## 4.8 Effects by Neighborhood

We explore the effect estimates of the main model and how they vary between different types of neighborhoods. This is an alternative way of exploring the interaction effect of diversity because we can see how the estimates change depending on the religious composition of each neighborhood. By design, the neighborhoods included in this study vary widely in terms of their religious composition. As noted in the section of this appendix on site selection, we used data from electoral rolls to select three types of locations: majority Hindu, mixed, and majority Muslim. We were successful in obtaining appropriate variation on this variable by neighborhood. As a reminder, we surveyed participants in sixteen neighborhoods, and these neighborhoods themselves can be divided into five areas: Settlement A, Settlement B, Settlement C, Settlement D, and Settlement E. By the estimates in our sample, the six neighborhoods in the Settlement C area, as well as the neighborhood of Settlement A, all contained a small percentage of Muslim residents (below 10%). Five of the neighborhoods – Settlement B, three of the neighborhoods in the Settlement D area, and one of the Settlement E neighborhoods – we classify as mixed because they contain more than 10% but less than 50% Muslim residents. Lastly, four neighborhoods – one from Settlement D and three from Settlement E – have at least 50% Muslim residents in our sample, with the three Settlement E neighborhoods having especially high percentages of Muslim residents (ranging between 84% and 97% Muslim).

When we estimate effects separately in these three types of locations, we find that effects for both Muslims and Hindus vary in interesting ways. Table 19 shows the estimates from each of these models, the left column

containing results for the large Hindu majority locations, the middle column containing results for the mixed locations, and the right column containing results for the majority Muslim locations. Three aspects of these estimates deserve mentioning. First, the main model's finding that Muslims respond to the treatments while Hindus do not seem to be primarily driven by residents in neighborhoods with either strong Hindu majorities (<10% Muslim) or Muslim majorities (>50% Muslim). The interaction effect for a respondent being Muslim in these neighborhoods is higher than the main model; at 0.26 and 0.25, respectively, the interaction estimate is approximately 1.5 times the estimate in the main model, although the standard error for these estimates is also larger. In comparison, the interaction estimate for respondents in the mixed neighborhoods (>10% but <50% Muslim) is effectively zero. Second, we also note that the main model's estimate of a lower baseline for Muslim participants in the control condition (the lower-level coefficient on Muslim in the main model is estimated at -0.13, ) seems to be primarily driven by respondents in the Muslim majority neighborhoods. In these neighborhoods, Hindu participants have a very high baseline of interest (2.95,  $SE = 0.09$ , compared to 2.76,  $SE = 0.02$ , in the main model) in the drainage cleaning program. Third, we note that Hindus in mixed neighborhoods do have a marginally significant effect estimate for the combined treatments (0.12,  $SE = 0.07$ , compared to an effect of -0.02,  $SE = 0.03$  in the main model).

It is difficult to draw strong conclusions from these findings; we did not hypothesize about effects varying by these neighborhood groups prior to carrying out this project, and the differences we find here are all of marginal significance compared to the findings in the main model of the paper. However, if we are to draw conclusions from these varying effects by neighborhood, it would be that Muslims are least likely to be affected by our treatments in non-polarized neighborhoods.

Table 19: Main Model by Percent Muslim in Neighborhoods

	Dependent Variable: Index of Favorability toward Drainage Program		
	Low (<10%)	Medium (>10%, <50%)	High (>50%)
Constant	2.77 (0.02)	2.62 (0.05)	2.95 (0.09)
Treatments (Combined)	-0.04 (0.03)	0.12 (0.07)	-0.11 (0.13)
Muslim	-0.13 (0.10)	-0.04 (0.09)	-0.30 (0.10)
Treatments (Combined) x Muslim	0.26 (0.14)	0.08 (0.13)	0.25 (0.15)
# Hindu Respondents	2157	579	188
# Muslim Respondents	114	211	594
Observations	2,262	789	781
R <sup>2</sup>	0.002	0.01	0.01
Adjusted R <sup>2</sup>	0.0004	0.004	0.01
Residual Std. Error	1.09 (df = 2258)	1.20 (df = 785)	1.33 (df = 777)
F Statistic	1.30 (df = 3; 2258)	2.00 (df = 3; 785)	3.31 (df = 3; 777)

## 4.9 Ceiling/Floor Effects (Tobit Model)

Our main outcome variable in this study is an index composed of five questions about participant interest in the drainage cleaning program. These questions are designed to elicit a range of responses on the 1-4 scale used to measure the outcomes. For the first two questions – would this program be beneficial to your neighborhood and how interested would you be in the program – respondents were much more likely to register high values, with averages of 3.02 and 2.94, respectively, for those in the control condition. In response to the third question – how likely would you be to pay the monthly fee – respondents reported slightly lower values, with an average of 2.74 for control condition respondents. Lastly, in response to the fourth and fifth questions – would you be willing to sign a six-month contract for this service and how likely would you be to try to get your neighbors to sign up for the program – respondents in control reported the lowest average responses of 2.46 and 2.51, respectively.

Table 20: Tobit Model

Dependent Variable: Index of Favorability toward Drainage Program	
Constant	2.77 (0.02)
Treatments (Combined)	−0.02 (0.03)
Muslim	−0.13 (0.04)
Treatments (Combined) x Muslim	0.21 (0.05)
Observations	3,832
Log Likelihood	−10,601.50
Wald Test	20.98 (df = 3)

However, average responses were higher than the midpoint of the scale for all our measures, which may make some worried about ceiling effects for Hindus, who saw no effect of the treatments but who also had a higher baseline than Muslims for each of the index items. Additionally, these questions evaluate attitudes on a somewhat polarizing topic for communities, and we might be worried about a floor for respondents' attitudes as well because of strong opinions from those who oppose the program. Indeed, when we look at a histogram of the index in Figure 8, we see that that respondents pool at the lower and upper ends of our scale,

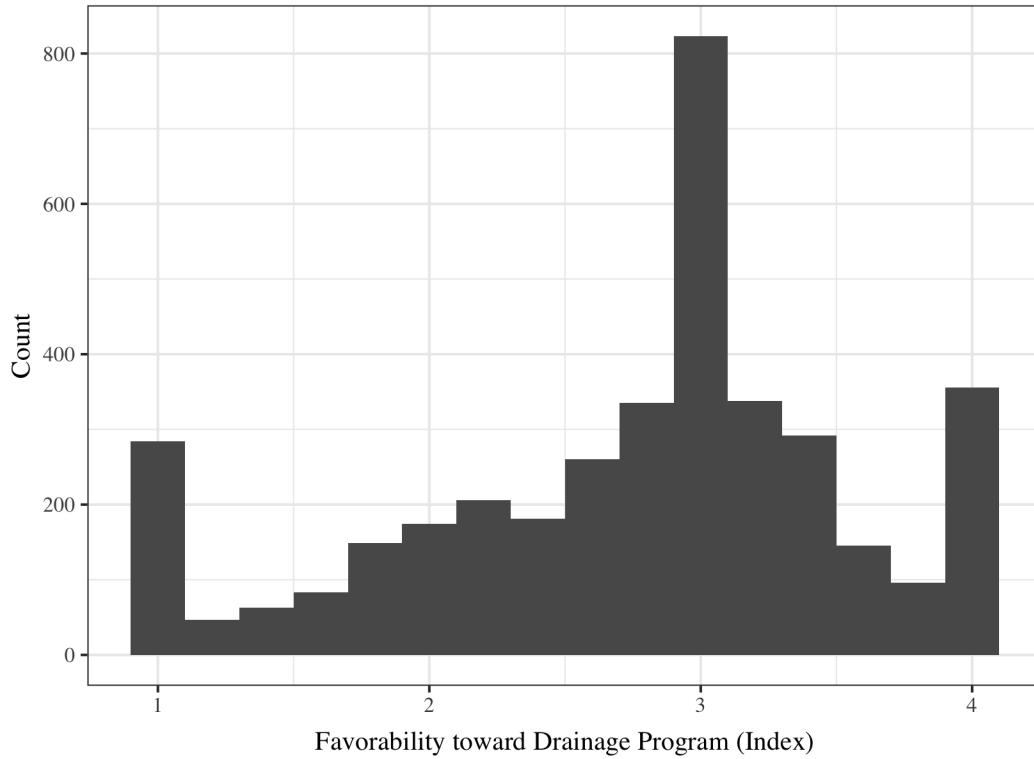


Figure 8: Histogram of Outcome Index, All Respondents

with 7.4% of respondents recording a 1 for all five outcome questions and 9.3% of respondents recording a 4 for all of them. To make sure that the results we find in our main model are not driven by this pooling in any way, we estimate a tobit model with a floor of 1 and a ceiling of 4, the results of which can be found in Table 20. The estimates from this model are almost the same as those in the main analysis; if anything, the coefficients from the tobit model would indicate a slightly higher interaction effect for Muslims.



## 4.10 Enumerator Effects

Because our survey hinges on religious identification, we may be concerned about how the religion of the survey enumerator affects our results. In Table 21 we calculate our main analysis separately for those who had a Hindu enumerator and those who had a Muslim one. Before discussing these results, we want to note that this analysis should be viewed as highly speculative. We did not pre-register an analysis of this sort. Additionally, out of the twenty-five total enumerators who administered our survey, only two of them were Muslim. There are only 141 respondents who took the survey from a Muslim enumerator. Lastly, enumerators were not randomly assigned to settlements, so our Muslim enumerator results come primarily from just six of our sixteen locations.

With these caveats in mind, the analysis in Table 21 indicates that Muslim enumerators may elicit very different effects from participants in a number of ways, especially in terms of the “baseline” support for the program in the control group. For instance, the baseline in the control group for Hindu respondents is much lower when they have a Muslim enumerator (2.15,  $SE = 0.11$ ) than when they have a Hindu enumerator (2.78,  $SE = 0.02$ ). On the other hand, the opposite is true for Muslim respondents; the baseline for Muslims is much higher with a Muslim enumerator (3.17,  $SE = 0.16$ ) than it is with a Hindu enumerator (2.61,  $SE = 0.04$ ). Lastly, from this higher baseline, the treatment effect for Muslim respondents with Muslim enumerators is now estimated to be negative (-0.54,  $SE = 0.31$ ), as opposed to the positive effect with Hindu enumerators and in the main analysis.

As already noted, we do not think we can draw much from these models. The estimates in the right column of Table 21 are entirely dependent on two Muslim enumerators who only administered the survey to 141 respondents, who are not representative of our sample. However, examining these results is useful for considering the scope and generalizability of this research project. Because the vast majority of our enumerators are Hindu, our findings are primarily applicable to situations where programs like ours would be instituted by members of the majority religious/ethnic group.

Table 21: Main Analysis, by Religion of Enumerator

	Dependent Variable: Index of Favorability toward Drainage Program	
	Hindu Enumerators	Muslim Enumerators
Constant	2.78 (0.02)	2.15 (0.11)
Treatments (Combined)	-0.01 (0.03)	-0.06 (0.15)
Muslim	-0.17 (0.04)	1.02 (0.20)
Treatments (Combined) x Muslim	0.20 (0.06)	-0.54 (0.31)
Observations	3,685	141
R <sup>2</sup>	0.01	0.21
Adjusted R <sup>2</sup>	0.004	0.19
Residual Std. Error	1.16 (df = 3681)	1.06 (df = 137)
F Statistic	6.30 (df = 3; 3681)	12.13 (df = 3; 137)

### 4.11 Effects by Caste Level, for Hindus

We used Table 22 to examine how the effect of the treatment differs by caste level for Hindus, for whom the caste system is more rigid and it is therefore more reasonable to expect an interaction effect. Upper caste Hindus start at a higher baseline than lower-caste Hindus, indicating that they are more likely to support the drainage program in the control. However, we find that there is no interaction—the coefficient on the interaction term is insignificant and not large.

Table 22: Effect of Treatments, by Caste Level (Lower v. Upper)

Dependent Variable: Index of Favorability toward Drainage Program	
Constant	2.73 (0.02)
Treatments (Combined)	0.002 (0.03)
Upper Caste	0.11 (0.05)
Treatments (Combined) x Upper	−0.09 (0.07)
Observations	2,886
R <sup>2</sup>	0.002
Adjusted R <sup>2</sup>	0.001
Residual Std. Error	1.12 (df = 2882)
F Statistic	1.79 (df = 3; 2882)

## 4.12 Religious Ties as Mechanism, Separate Treatments

We examine religious ties as a mechanism for each of our separate treatments, rather than combined. We find a similar (though less significant) effect for all treatments, except for vertical ties alone, which is null.

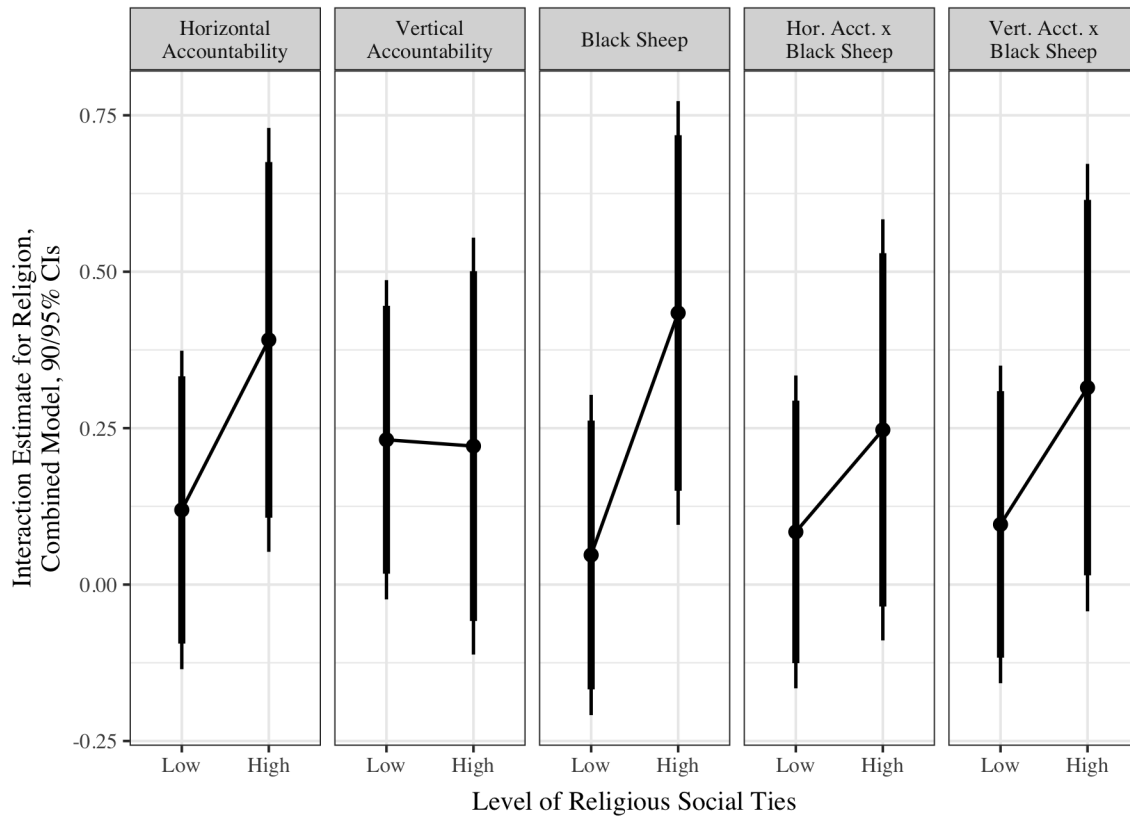


Figure 9: Religious Ties as Mechanism, Separate Treatments

## References

- Ali, S. (2002). Collective and Elective Ethnicity: Caste among Urban Muslims in India. *Sociological Forum*, 17(4):593–620.
- Auerbach, A. M. (2016). Clients and Communities: The Political Economy of Party Network Organization and Development in India's Urban Slums. *World Politics*, 68(1):111–148.
- Auerbach, A. M. (2017). Neighborhood Associations and the Urban Poor: India's Slum Development Committees. *World Development*, 96:119–135.
- Auerbach, A. M. and Thachil, T. (2018). How Clients Select Brokers: Competition and Choice in India's Slums. *American Political Science Review*, 112(4):775–791.
- Chidambaram, S. (2020). How Do Institutions and Infrastructure Affect Mobilization around Public Toilets vs. Piped Water? Examining Intra-slum Patterns of Collective Action in Delhi, India. *World Development*, 132:1–14.
- Gayer, L. and Jaffrelot, C., editors (2012). *Muslims in Indian Cities: Trajectories of Marginalisation*. Columbia University Press, New York, NY.
- Hansen, B. B. and Bowers, J. (2008). Covariate balance in simple, stratified, and clustered comparative studies. *Statistical Science*, 23(2):219–236.
- Heller, P., Mukhopadhyay, P., Banda, S., and Sheikh, S. (2015). Exclusion, Informality, and Predation in the Cities of Delhi An Overview of the Cities of Delhi Project. Working paper.
- Inglehart, R., Haerpfer, C., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., and Puranen, B. (2012). World Values Survey: Round Six - Country-Pooled Datafile. Technical report, JD Systems Institute & WWSA Secretariat, Madrid, Spain & Vienna, Austria.
- Jha, S., Rao, V., and Woolcock, M. (2007). Governance in the Gullies: Democratic Responsiveness and Leadership in Delhi's Slums. *World Development*, 35(2):230–246.
- Kustov, A. and Pardelli, G. (2018). Ethnoracial Homogeneity and Public Outcomes: The (Non)effects of Diversity. *American Political Science Review*, pages 1–8.
- Lokniti (2008). *State of Democracy in South Asia*. Oxford University Press.
- Lokniti (2012). Slums and Citizens: Study of Delhi Slums. Technical report, Center for the Study of Developing Societies, New Delhi.
- Lokniti (2018). National Election Studies. Technical report, Center for the Study of Developing Societies, New Delhi.
- Lust, E., Kristen, L., Nyasente, D., Harris, P., Harris, A., Dulani, B., Metheney, E., Nickel, S., Carlitz, R., Gakii Gatua, J., Jöst, P., Mechkova, V., Mujenja, M. F., Tengeganga, J., Grimes, M., Ahsan Jansson, C., Alfonso, W., Ben Brahim, N., Jordan, J., Bauhr, M., Boräng, F., Ferree, K., Hartmann, F., and Lueders, H. (2020). The Local Governance and Performance Index (LGPI) Household Survey 2019: Kenya, Malawi, Zambia. Technical report, The Program on Governance and Local Development, University of Gothenburg.