

Supplemental Appendix:  
Social Accountability and Public Goods Provision  
Testing Informal Mechanisms to Improve Community Welfare in the slums of  
Delhi

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)

April 11, 2023

# Contents

<b>1</b>	<b>Preliminary Qualitative Fieldwork</b>	<b>3</b>
<b>2</b>	<b>Site Selection</b>	<b>6</b>
<b>3</b>	<b>Survey Administration</b>	<b>8</b>
<b>4</b>	<b>Experimental Portion of Survey</b>	<b>11</b>
4.1	Introduction . . . . .	11
4.2	Treatment Conditions Presented in Paper . . . . .	12
4.3	Additional Treatment Conditions . . . . .	13
4.4	Pre-Registration . . . . .	15
4.5	Ethics . . . . .	15
<b>5</b>	<b>Key Variables</b>	<b>17</b>
5.1	Outcome Measure . . . . .	17
5.2	Demographic and Socioeconomic Measures . . . . .	18
5.3	Measures of Religion and Religiosity . . . . .	19
5.4	Measures of Social Ties . . . . .	21
5.5	Measures of Political Factors . . . . .	23
5.6	Measure of Caste . . . . .	24
5.7	Measure of Diversity . . . . .	27
5.8	Measures related to Citizenship, Drainage, and Enumerator . . . . .	28
<b>6</b>	<b>Demographic Variables</b>	<b>30</b>
<b>7</b>	<b>Robustness Checks and Additional Analyses</b>	<b>32</b>
7.1	Supporting Table . . . . .	32
7.2	Balance . . . . .	32

7.3	Survey Purpose and Manipulation Checks . . . . .	34
7.4	Timing of Survey Administration and Treatments . . . . .	38
7.5	All Treatments . . . . .	39
7.6	Enumerator Effects . . . . .	40
7.7	Floor/Ceiling Effects . . . . .	42
7.8	Variations on Diversity Measure . . . . .	46
7.9	Caste and Religion . . . . .	47

# 1 Preliminary 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. Following pre-established interview and data collection protocols that we developed jointly with the CPR, trained field workers gathered qualitative data from members of the community and through observation on six main characteristics of potential study sites: (1) basic demographic information on population estimates through official sources, such as the Delhi Urban Shelter Improvement Board (DUSIB) and the Municipal Corporation of Delhi (MCD) (2) the religious and caste identification of the individuals living in a settlement, based on local knowledge. This information (1 and 2) was also used to corroborate our estimates of total population and the proportion of Muslim population gathered from electoral lists (discussed in Section 2), (3) the age and history of the settlement, including the evolution of public service delivery over time (4) the socio-economic status as well as main occupations of residents in the site, for which there is considerable variation even within slums in Delhi, (5) the level of public goods provision in the site based on the condition of drainage, local roads, water supply, and public toilets, and potential solutions of interest to residents and (6) the types and activities of local political leaders as well as other social organizations in the community, and finally (7) the geographic boundaries for cooperation over providing drain cleaning services. No identifying information of any sort was collected during this process.

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 guided site selection as well as the choice of public good: Based on findings from the fieldwork, we were also able to choose heterogeneous and homogeneous communities that were largely similar to each other in other respects. The reports also informed our choice of drain cleaning as an appropriate focus for the survey. Drainage was identified to be the single most important public good necessitating cooperation among residents at the neighborhood level and the quality of

drainage was an important concern in most settlements. While municipal workers did clean drains along major roads surrounding the settlements, residents in most *bastis* were forced to maintain internal drains through their own efforts (See Table 1).

Table 1: Condition of Drainage in Selected Sites



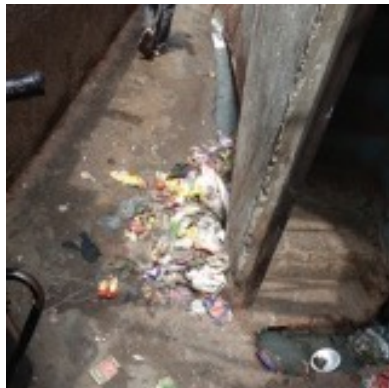
Main drainage clogged with garbage in Site D



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 also informed the design of the instrument. 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. These interviews also provided input on an appropriate monthly fee for the hypothetical drainage company presented in the instrument.

Finally, qualitative evidence on the nature of political and social organization within *bastis* was helpful in interpreting the findings from the survey.

## 2 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 2; Heller et al., 2015). With the exception of resettlement colonies, these sites are unplanned and 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 2: 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 close to a 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 website of the Chief Electoral Officer of Delhi, 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 in Section 3. 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 4

in the subsequent section for an anonymized list of these locations.) The sites were similar in socioeconomic characteristics, quality of public services, and age (ranging from 30-40 years).



### 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 Table 3).

Table 3: 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 target 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. Because we included a screening question for religion, only Muslims and Hindus are

a part of our sample. (In line with our preliminary fieldwork on the sites, those of other religions were very rare.)

Table 4 provides an overview of our sample by settlement. In total, our enumerators estimated that there were 12,310 households in the areas to which they administered the survey. Out of those, 5,814 individuals began a survey, of whom 14 were never assigned treatment and an additional 3 never had their settlement recorded, reducing the total sample size to 5,797. This puts our overall response rate at 47%, though as seen in Table 4, the rate varies from location to location, with the lowest at 27% (in Basti E1) and the highest at 78% (in Basti C5).

Table 4: Number of Participants Surveyed and Response Rate by *Basti*

	Site	Basti	Households	Surveyed	Closed House/Refused	Response Rate
1	A	A1	1,400	1,005	395	72%
2	B	B1	1,260	541	719	43%
3	C	C1	1,550	673	877	43%
4	C	C2	130	66	64	51%
5	C	C3	1,080	634	446	59%
6	C	C4	1,065	491	574	46%
7	C	C5	125	98	27	78%
8	C	C6	850	436	414	51%
9	D	D1	900	381	519	42%
10	D	D2	400	212	188	53%
11	D	D3	400	211	189	53%
12	D	D4	300	158	142	53%
13	E	E1	2,400	669	1,731	28%
14	E	E2	140	106	34	76%
15	E	E3	230	62	168	27%
16	E	E4	80	54	26	68%

However, we do not use this entire sample in the paper. First, some participants were missing data on important variables, eliminating an additional 124 participants from the total sample size. The variables for which missing data led to exclusion include geocoordinates (used to determine the diversity of the respondent’s immediate neighborhood), religious identification, caste identification, and the outcome index. This is a small number of participants (approximately 2%), and we treat those units lacking these data as missing completely at random (MCAR).

Second, out of the remaining 5,673 participants, we keep only 2,516 that fall into the control condition and three experimental treatments: Horizontal Accountability, Vertical Accountability, and the Black Sheep treatment. For this paper, we exclude two combination treatments as well as three treatments we considered as alternate controls. For more information on all of the treatments, please see Section 4 of this appendix.

## 4 Experimental Portion of Survey

Participants were randomized into the groups shown in Table 5. Because over 20 enumerators each used their own tablet to administer the survey, participants were assigned to a treatment group using simple randomization, which results in slightly different sample sizes from treatment group to treatment group. Note that, as indicated by the first column in Table 5, only participants from the first four treatment groups were included in the analysis presented in the paper. Note that this table lists the sizes of each treatment group after eliminating participants with data missing on important variables, as noted in Section 3, adding up to a total of 5,673 participants.

Table 5: Treatment Groups

In Paper?	Treatment	N
Yes	Control	654
Yes	Horizontal Accountability	622
Yes	Vertical Accountability	637
Yes	Black Sheep	603
No	Combination Condition: Horizontal Accountability + Black Sheep	616
No	Combination Condition: Vertical Accountability + Black Sheep	635
No	No Names Conditions: Control	675
No	No Names Conditions: Horizontal Accountability	605
No	No Names Conditions: Vertical Accountability	626

### 4.1 Introduction

*The experimental portion of the survey begins with a description of a hypothetical drain cleaning service. That description read as follows:*

We are now going to present to you an offer from a drain cleaning company. Please note that this offer is entirely hypothetical, and no service is going to be implemented as a result of this study.

This company would regularly clean the drains and make sure that they are not clogged. In case of any problems with the drainage system, especially during the rainy season, the company would send its staff to take care of the issue. We would like to know if residents of your neighborhood, [basti name inserted here], would be interested in this offer.

There are two conditions for this offer:

- A monthly subscription fee of Rs.50 per household.
- 2/3 or 67% of the residents in your neighborhood indicate their agreement with this contract.

## **4.2 Treatment Conditions Presented in Paper**

Participants from the first four treatments listed in Table 5 are analyzed in the paper. Below is the text of each of the treatment groups. Where necessary, italics are used to comment on the implementation of that particular treatment.

**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 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 be reported to the other residents in your area.

**Horizontal Accountability** *Same as control until the final paragraph, which reads as follows.*

... 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

**Vertical Accountability** *Same as control until the final paragraph, which reads as follows.*

... 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

**Black Sheep (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.

### **4.3 Additional Treatment Conditions**

There were five treatment conditions for which we do not present results in the paper. Two of these, which we refer to as the “combination conditions,” combined the accountability treatments

with the black sheep treatment. The other three were alternative versions of the control, horizontal accountability, and vertical accountability treatments that did not include the names in either the first paragraph (i.e. “among them Sunil and Abdul”). These names are not necessary to this treatment, but they are for the black sheep condition, which uses manipulation of these names to emphasize ingroup underperformance. We didn’t know if just the inclusion of these names by itself would change the results, so we included both version of these conditions. Because results do vary based on this, our main analysis uses the control and accountability treatments with names because it is the most similar to the black sheep condition. An analysis in Section 7.5 provides results for all conditions.

**Combination Condition: Horizontal Accountability and Black Sheep (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

**Combination Condition: Vertical Accountability and Black Sheep (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

**No Names Conditions: Control, Horizontal Accountability, and Vertical Accountability** *Same as the text already presented for these conditions but without the parts of the first paragraph that refer to specific names.*

## 4.4 Pre-Registration

We pre-registered the design of our experiment with EGAP, now migrated to Open Science Foundation. Crucially, we indicated that we would examine heterogeneous effects by diversity as a main hypothesis, as well as by religious identification (Muslim versus Hindu) for one of the treatments. Below, we provide clarification where necessary to explain divergence from the pre-analysis plan.

*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 to conduct.

*Second:* As already noted, we do not include all treatment arms of the experiment in the analysis in this paper. We exclude the combined treatments for space concerns and also because there is no noticeable difference in the combined effects, and we exclude the alternative “no names” conditions because we view our use of the “with names” control to provide a closer comparison with the treatments.

*Third:* When considering heterogeneous effects by religious identification, we only hypothesized an effect for the black sheep treatment. However, results are similar across treatments, and therefore we present heterogeneous effects across all treatments.

## 4.5 Ethics

Our research project adheres to all APSA Principles and Guidelines for Human Subjects Research and was approved by the 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, the imperative 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.

## 5 Key Variables

Our instrument draws upon a number of existing surveys, including the World Values Survey (Inglehart et al., 2012), surveys conducted for 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) at Gothenburg University in Sweden (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 survey projects. Questions on religious participation, caste, and financial assets were based on the National Election Survey templates, while those on social ties and networks relied heavily on GLD questionnaires.

### 5.1 Outcome Measure

We use an index of five questions to measure the outcome of our experiment. All five questions 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 answered. Section 7.7 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 among the entire sample, indicating a high internal consistency for this index.

## 5.2 Demographic and Socioeconomic Measures

Our survey includes a number of measures that we primarily use to address similarity between Muslims and Hindus in our sample and as control variables in our models. Below is a list of basic demographic variables.

- **Gender:** Coded as 1 for females and 0 for males.
- **Married:** Coded as 1 if married and 0 otherwise.
- **Age:** Age in years.

Variables that capture socioeconomic status include the following. In addition, we generate measures to capture financial hardship and economic condition through level of assets.

- **Home Ownership:** Coded as 1 if they own their place of residence and 0 otherwise.
- **People per room in Residence:** Number of members of household divided by the number of rooms in their residence.
- **Education Level:** 1–10 scale, with higher numbers indicating higher education.
  - 1 = No formal education, illiterate;
  - 2 = No formal education, literate;
  - 3 = Incomplete primary school;
  - 4 = Completed primary school;
  - 5 = Middle pass/Matric fail;
  - 6 = Matric pass;
  - 7 = 12th pass/Intermediate;

- 8 = Graduate;
  - 9 = Post-graduate;
  - and 10 = Professional Degree.
- **Employment Status:** Coded as 1 if employed and 0 if not.
  - **Financial Hardship Index:** An index of four measures assessing financial hardship, standardized on a 0–1 scale with higher numbers indicating more hardship.
    - First measure is of trouble paying rent or house payments in the last year, coded as 1 if yes and 0 if no;
    - Second measure is of having to take out a loan in the last year, coded as 1 if yes and 0 if no;
    - Third measure is of having to sell valuables in the last year, coded as 1 if yes and 0 if no;
    - Fourth measure is of savings in the past year, with a 1 indicating that the household saved money, a 2 indicating that they just got by, and a 3 indicating that they spent some savings; standardized on a 0–1 scale before combining with the dichotomous variables.
  - **Level of Assets:** To capture the variation in
    - **Mid-Tier Assets:** Dichotomous variable generated from a set of asset measurements, equal to 1 if they have a bike, cooler, or fridge and a 0 otherwise.
    - **High-Tier Assets:** Dichotomous variable generated from a set of asset measurements, equal to 1 if they have a vehicle (2-, 3-, or 4-wheel), computer, or AC and 0 otherwise.

### 5.3 Measures of Religion and Religiosity

Nine questions address religion and religiosity in our survey. 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 site selection, the vast majority of respondents either Hindu or Muslim. Those from other religions were dropped from the survey. Responses to this question are used to identify respondents as Muslim or Hindu at the individual level and are later used, at aggregated levels, to define the religious composition of neighborhoods and of areas around each respondent. 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.

We use these five questions to generate measures of religiosity. The first question about prayer is called **Private Religious Practice**; it is not combined into an index with anything else. We combine the remaining four items into a single index, referred to as **Public Religious Practice**.<sup>1</sup>

---

<sup>1</sup>Note that, contextually, donations and fasting are both fairly public religious activities for Hindus and Muslims.

Respondents were also 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. These two questions are combined into an index we refer to as a measure of **Religious Ties**; this index is used to explore the mechanisms of the effect of the treatments.

## 5.4 Measures of Social Ties

In the survey, 16 questions relate to social ties with others in their neighborhood. The first is a standalone question, with a three-point scale of possible responses: Few people know each other, some people know each other, or most people know each other.

- In your neighborhood, would you say:

The second social ties question is an independent question about the frequency with which respondents talk to their neighbors, with a five-point scale of possible responses: Never, only when necessary, rarely, sometimes, or daily. We combine the two into a single index, called **General Social Ties**.

- How often do you talk to your neighbors?

Respondents are then asked to answer a set of five questions about social ties with neighbors. Respondents could answer one of the following on a four-point scale: Never, rarely, sometimes, or always. We combined these to create form an index measure. This index had a high internal consistency, with a Cronbach's alpha of 0.86. We refer to this measure as the **Helpfulness Index**.

- 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

Lastly, an additional nine-question battery of questions measures 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. We use these questions to make two indices. The first we refer to as the **Obligation to Family Index**, and for it we include just the measures for immediate and extended

family. The second we refer to as the **Social Obligation Index**, and it includes all the remaining measures.

## 5.5 Measures of Political Factors

We generate a number of variables to gauge the nature of political engagement in respondents.

- **Political Participation:** Proportion of six listed political activities that the participant engaged in during 2015 elections, if they have a voter card with a Delhi address (if they didn't then coded as a 0).
  - The six activities are attending election meetings/rallies, participating in processions/nukkad natak etc, participating in door-to-door canvassing, contributing or collecting money, distributing election leaflets or putting up posters, and transporting neighbors to polling booths.
- **Networks in State Institutions:** A dichotomous variable coded as a 1 if they noted relations with any of seven local figures.
  - The seven local figures asked about were a government official in the area, a party worker from a party in the area, the pradhan, the MLA, the MCD Councilor, an NGO worker in the area, an a religious leader in the area;
  - they were counted as yes if one of these individuals was a relative, a close friend, or both.
- **Involvement in Local Organizations:** A dichotomous variable coded as a 1 if they noted participation in any of five local volunteer organizations.
  - The five local volunteer organizations asked about were an NGO, a Resident Welfare Association (RWA), caste organization, religious organization, and a non-caste and non-religious organization.



The survey includes six measures of participant perceptions of efficacy of state and political institutions. 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. We generate a measure, called **General Political Trust Index**, based on the first five institutions. *Pradhan* is not included in this index since it is an informal institution. The political trust index does not have very good internal consistency (Cronbach's alpha = 0.57), but is of theoretical importance. We also combine measures of two of these—the MCD and *pradhan*—with a separate measure in the survey measuring *pradhan* popularity (also on a 1–4 scale) into a measure called **Local Political Trust Index** that is directly relevant for the cleaning and maintenance of drains.

## 5.6 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, in which 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). Note that, for this section, we are using the full sample size of 5,797 to report caste statistics from the entire sample (see Section 3). This resulted in 2,519 respondents identified as belonging to a low caste, 1,400 identified as belonging to another backward class (OBC), and 1,813 identified as belonging to a high caste. Using this categorization, approximately 45% of Hindus and 39% of Muslims were categorized as low caste (see Table 6). When using caste as a variable in our 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; Kawar; Korwa; Nagesia; Sawar; Ho; Kharia; Bhumij; Chero; Kharwar; Lohra; Mahli; Mal Pahariya; Mahali; Patelia; Baiga; Bhilala; Mahadev-Koli; Korku; Barela; Bhotia; Buksa; Jannsari; 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	1,981 (45%)	1,400 (32%)	1,003 (23%)	42 (1%)
Muslim	538 (39%)	-	810 (59%)	23 (2%)

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

## 5.7 Measure of Diversity

We measure ethnic diversity through fine-grain data on GPS coordinates of the household location as the survey was administered. We used two methods to collect this data. First, the tablets used to administer the survey automatically recorded the enumerator's geo-location 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.

Since cellular service is sometimes spotty in the areas of Delhi we surveyed, even using both of these methods sometimes generated inaccurate geocoordinates – the enumerator is either unable to access the Google Maps application or the device records the coordinates of the closest cell phone tower rather than the enumerator's precise location. Approximately 7% of our surveys had evidence that this occurred (e.g. the geocoordinates were not as fine-grained as the others or dozens of participants had the exact same number). 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 in such cases. We deemed it better to provide an approximate estimate rather than eliminate these data. For five respondents, none of these methods provided reliable coordinates.

We then use the geolocation of respondents, combined with the measure of religious identification as Hindu or Muslim, to construct our measure of diversity. Our measure of diversity is based

on the polarization index from [Montalvo and Reynal-Querol \(2005\)](#), represented by the following 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. We use the data on geocoordinates to compute the polarization index for each participant based on the religious identification of other participants within 100 meters (equivalent of a few streets). This provides us with a near-continuous measure of diversity at the local level. We further generate alternative measures of diversity with 50 meters and 150 meters radius respectively.

## 5.8 Measures related to Citizenship, Drainage, and Enumerator

- **Years in Settlement:** Measured in years.
- **Consider Self Citizen of Delhi:** Coded as 1 if yes and 0 if no.
- **Quality of Drainage:** Rating of quality of drainage in neighborhood, 1–5 scale.
  - 1 = Extremely bad;
  - 2 = Somewhat bad;
  - 3 = Okay;
  - 4 = Somewhat good; and
  - 5 = Very good.

- **Drainage Problem Requiring Help:** Dichotomous measure of household experiencing drainage problems requiring help outside of the household in the last two years.
- **% Hindu Enumerator:** The proportion of participants given a Hindu enumerator.

## 6 Demographic Variables

Table 8 provides demographic information on the 2,516 participants included in the analyses in the paper. On the whole, our modal participant is male, married, and middle-aged; has lived two decades in a house with two rooms that they own; is employed, experiences moderate financial hardship, and owns a few valuable assets (e.g., motorbike, cell phone, and TV); has a low level of education; is Hindu and engaged in their religion; has low political involvement but strong social ties as evidenced by their willingness to help others (even if it requires them to forgo wages); is lower caste; and lives in a neighborhood with very poor drainage. Overall, about a quarter of the sample is Muslim.

Table 8: Sample Demographic Characteristics

Variable	Mean	SD
Demographics		
Gender (1 = Female, 0 = Male)	0.19	0.39
Married (1 = Married, 0 = Other)	0.91	0.29
Age	38.48	11.74
Home Ownership (1 = Own Residence, 0 = Other)	0.83	0.38
People Per Room in Residence	3.36	1.87
Education Level (1-10 Scale)	3.94	2.31
Employment Status (1 = Employed)	0.83	0.38
% Mid-Tier Assets (Bike, Cooler, Fridge)	0.49	0.50
% High-Tier Assets (Vehicle, Computer, AC)	0.19	0.39
Financial Hardship Index (1 = More Hardship)	0.25	0.23
Religiosity/Social Ties		
Private Religious Practice (1 = Never, 5 = Daily)	4.25	1.08
Public Religious Practice (4 Items, 1–5 Scale)	3.79	0.86
Religious Ties Index (2 Items, 1–3 Scale)	2.19	0.58
General Social Ties (0-1 Scale)	0.89	0.20
Helpfulness Index (1-4 Scale)	3.11	0.78
Forgo Wages Index, Family (1-3 Scale)	2.79	0.35
Forgo Wages Index, Others (1-3 Scale)	2.25	0.53
Political Factors		
General Political Trust Index (6 Items, 1–4 Scale)	2.18	0.53
Local Political Trust Index (3 Items, 1-4 Scale)	1.91	0.77
Network in State Institutions (7 Items, 1 = Any)	0.14	0.35
Political Participation, 2015 (6 Items)	0.16	0.25
Involvement in Local Organizations (1 = Any)	0.06	0.24
Caste		
Caste Level (0 = Lower, 1 = Upper)	0.32	0.47
Diversity		
Hindu-Muslim Diversity within 100 Meters (0–1 Index)	0.40	0.31
Neighborhood and Enumerator		
Years Residing in Settlement	20.73	10.51
Consider Self Citizen of Delhi (1 = Yes)	0.94	0.24
Quality of Drainage (1-5)	1.81	1.05
Drainage Problem Requiring Help (1 = Yes)	0.23	0.42
% Hindu Enumerator	0.96	0.19



## 7 Robustness Checks and Additional Analyses

### 7.1 Supporting Table

Table 9 is the regression corresponding to Figure 1 in the paper.

Table 9: Effect of Treatments on Favorability toward Drainage Program, by Minority Religion

	Index of Favorability toward Drainage Program
Constant	2.76*** (0.04)
Horizontal Accountability	-0.01 (0.05)
Vertical Accountability	-0.03 (0.05)
Black Sheep	-0.02 (0.05)
Muslim	-0.13* (0.07)
Horizontal Accountability x Muslim	0.18* (0.11)
Vertical Accountability x Muslim	0.25** (0.10)
Black Sheep x Muslim	0.22** (0.11)
N	2,516
R <sup>2</sup>	0.003
Adjusted R <sup>2</sup>	0.001
Residual Std. Error	0.81 (df = 2508)
F Statistic	1.25 (df = 7; 2508)

\*p < .1; \*\*p < .05; \*\*\*p < .01

### 7.2 Balance

To test for balance, we use an omnibus balance test (Hansen and Bowers, 2008), reported in Table 10. Under this test, the null hypothesis is that the sample is balanced, and covariates are included as predictors in a model with assignment to treatment as the outcome. The test is an omnibus one because it simultaneously tests for the likelihood of seeing the estimated coefficients

on all covariates, given the null hypothesis of a balanced sample. The outcome is a binary covariate, so we test each treatment group against the control separately. We included 24 covariates to test for imbalance (the same ones included in our demographics table); any discordance between that and the degrees of freedom reported in Table 10 is due to the presence of NAs in these covariates. Missing values are treated as their own dichotomous covariate for the purpose of these tests. The p-values in the final column of Table 10 indicate that we cannot reject the null hypothesis of a balanced sample for the three treatment groups whose results we examine in this paper.

Table 10: Omnibus Balance Tests

Treatment	$\chi^2$	DF	P-Value
Horizontal Accountability	31.634	30	0.385
Vertical Accountability	24.762	31	0.778
Black Sheep	27.609	30	0.591

### 7.3 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 implementing the drain 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 2,243 respondents in our four treatment groups, or 89% 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 2,388 respondents, or 95% 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 Tables 11, 12, and 13.

Table 11: Effect of Treatments, Excluding if Failed Manipulation Check

	Index of Favorability toward Drainage Program
Constant	2.79*** (0.03)
Horizontal Accountability	0.02 (0.05)
Vertical Accountability	0.04 (0.05)
Black Sheep	0.03 (0.05)
N	2,228
R <sup>2</sup>	0.0003
Adjusted R <sup>2</sup>	-0.001
Residual Std. Error	0.79 (df = 2224)
F Statistic	0.25 (df = 3; 2224)

\*p < .1; \*\*p < .05; \*\*\*p < .01

Table 12: Effect by Diversity of Neighborhood, Excluding if Failed Manipulation Check

	Index of Favorability toward Drainage Program
Constant	2.80*** (0.05)
Horizontal Accountability	0.07 (0.07)
Vertical Accountability	-0.01 (0.07)
Black Sheep	-0.01 (0.07)
High Diversity	-0.02 (0.07)
Horizontal Accountability x High Diversity	-0.11 (0.09)
Vertical Accountability x High Diversity	0.09 (0.09)
Black Sheep x High Diversity	0.09 (0.10)
N	2,228
R <sup>2</sup>	0.003
Adjusted R <sup>2</sup>	-0.0003
Residual Std. Error	0.79 (df = 2220)
F Statistic	0.90 (df = 7; 2220)

\*p < .1; \*\*p < .05; \*\*\*p < .01

Table 13: Effect by Minority Religion, Excluding if Failed Manipulation Check

	Index of Favorability toward Drainage Program
Constant	2.83*** (0.04)
Horizontal Accountability	-0.03 (0.05)
Vertical Accountability	-0.02 (0.05)
Black Sheep	-0.03 (0.05)
Muslim	-0.16** (0.07)
Horizontal Accountability x Muslim	0.21* (0.11)
Vertical Accountability x Muslim	0.24** (0.11)
Black Sheep x Muslim	0.23** (0.11)
N	2,228
R <sup>2</sup>	0.003
Adjusted R <sup>2</sup>	0.0003
Residual Std. Error	0.79 (df = 2220)
F Statistic	1.09 (df = 7; 2220)

\*p < .1; \*\*p < .05; \*\*\*p < .01

## 7.4 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>2</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 half a second, 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
Black Sheep	7.38

---

<sup>2</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, among whom there was more variation (minimum of 9 minutes, maximum of 32 minutes).

## 7.5 All Treatments

Table 15 shows the overall effects (null) of all treatments, including those not a part of the analyses in the paper. Note the lack of a pattern with the combination treatments, which is why they were not a focus in the paper. Additionally, note that the difference between the “no names” control and the main control is as big or bigger than any other effect, indicating that just mentioning names might have an effect. For this reason, we use the control with names, which is closer to the treatments texts and therefore offers a tighter control.

Table 15: Effect of All Treatments

	Index Outcome
Constant	2.73*** (0.03)
Horizontal Accountability	0.04 (0.05)
Vertical Accountability	0.04 (0.04)
Black Sheep	0.04 (0.05)
Combination Condition: Horizontal Accountability + Black Sheep	-0.02 (0.05)
Combination Condition: Vertical Accountability + Black Sheep	0.04 (0.05)
No Names Conditions: Control	0.04 (0.04)
No Names Conditions: Horizontal Accountability	-0.02 (0.05)
No Names Conditions: Vertical Accountability	0.01 (0.05)
N	5,673
R <sup>2</sup>	0.001
Adjusted R <sup>2</sup>	-0.0005
Residual Std. Error	0.81 (df = 5664)
F Statistic	0.65 (df = 8; 5664)

\*p < .1; \*\*p < .05; \*\*\*p < .01



## 7.6 Enumerator Effects

The third model presented in our paper hinges on religious identification; for that reason, we may be concerned about how the religion of the survey enumerator affects our results. In Table 16 we calculate this 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 extremely speculative. Out of the twenty-five total enumerators who administered our survey, only two of them were Muslim. Out of our four treatment groups, there were only 97 respondents who took the survey from a Muslim enumerator. Additionally, 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 16 indicates that participants with Muslim enumerators have estimates that are different than those with Hindu enumerators, but those coefficients are estimated quite imprecisely and there is no clear pattern across treatments, with for instance Horizontal Accountability having a huge negative interaction with participant religious identity while the Vertical Accountability treatment has a huge positive interaction. Additionally, baselines might differ based on enumerator identity. For instance, the baseline in the control group for Hindu respondents is much lower when they have a Muslim enumerator (2.15,  $SE = 0.20$ ) than when they have a Hindu enumerator (2.78,  $SE = 0.04$ ). On the other hand, the opposite is true for Muslim respondents; the baseline for Muslims is much higher with a Muslim enumerator (3.17,  $SE = 0.35$ ) than it is with a Hindu enumerator (2.61,  $SE = 0.07$ ).

As already noted, we do not think we can draw much, if any, conclusion from these models. The estimates in the right column of Table 16 are entirely dependent on two Muslim enumerators who only administered the survey to 97 respondents, and those respondents are not representative of our sample as a whole. 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 16: Effect by Minority Religion, for Hindu and Muslim Enumerators

	Index of Favorability toward Drainage Program	
	Hindu Enumerators	Muslim Enumerators
	(1)	(2)
Constant	2.78*** (0.04)	2.15*** (0.20)
Horizontal Accountability	-0.01 (0.05)	-0.01 (0.27)
Vertical Accountability	-0.002 (0.05)	-0.08 (0.24)
Black Sheep	-0.02 (0.05)	-0.05 (0.27)
Muslim	-0.17** (0.07)	1.02*** (0.35)
Horizontal Accountability x Muslim	0.21** (0.11)	-0.76 (0.59)
Vertical Accountability x Muslim	0.23** (0.11)	0.51 (0.54)
Black Sheep x Muslim	0.25** (0.11)	-0.57 (0.55)
N	2,414	97
R <sup>2</sup>	0.004	0.23
Adjusted R <sup>2</sup>	0.001	0.17
Residual Std. Error	0.80 (df = 2406)	0.76 (df = 89)
F Statistic	1.34 (df = 7; 2406)	3.76*** (df = 7; 89)

\*p < .1; \*\*p < .05; \*\*\*p < .01

## 7.7 Floor/Ceiling Effects

Our main outcome variable in this study is an index composed of five questions about participant interest in the drain 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 2.95, 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.

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 1, we see that that respondents pool at the lower and upper ends of our scale, with 6.9% of respondents recording a 1 for all five outcome questions and 8.9% 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 for each of our models, the results of which can be found in Tables 17, 18, and 19. The estimates from these models are very similar to those in the paper; if anything, the coefficients from the tobit model in Table 19 would indicate slightly higher interaction effects for Muslims.

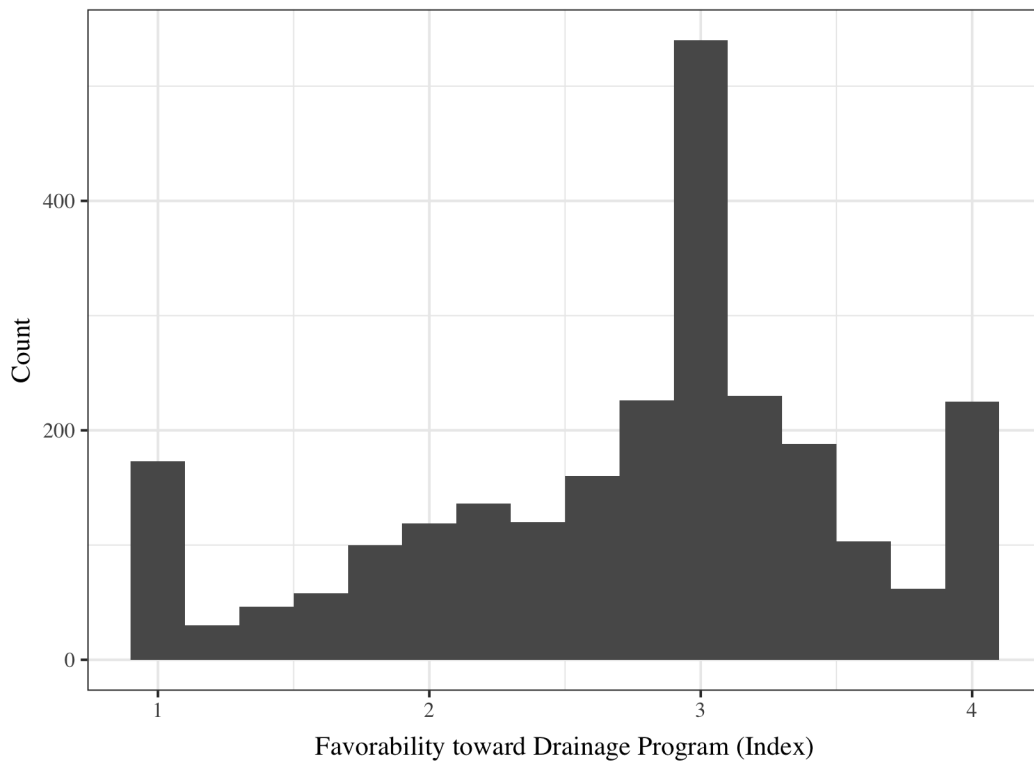


Figure 1: Histogram of Outcome Index, All Respondents

Table 17: Effect of Treatments, Tobit Model

	Index of Favorability toward Drainage Program
Constant	2.73*** (0.04)
Horizontal Accountability	0.04 (0.05)
Vertical Accountability	0.04 (0.05)
Black Sheep	0.04 (0.05)
N	2,516
Log Likelihood	-3,375.19
Wald Test	0.99 (df = 3)

\*p < .1; \*\*p < .05; \*\*\*p < .01

Table 18: Effect by Diversity of Neighborhood, Tobit Model

	Index of Favorability toward Drainage Program
Constant	2.71*** (0.05)
Horizontal Accountability	0.10 (0.07)
Vertical Accountability	0.01 (0.07)
Black Sheep	-0.01 (0.08)
High Diversity	0.05 (0.07)
Horizontal Accountability x High Diversity	-0.12 (0.11)
Vertical Accountability x High Diversity	0.07 (0.11)
Black Sheep x High Diversity	0.11 (0.11)
N	2,516
Log Likelihood	-3,370.92
Wald Test	9.54 (df = 7)

\*p < .1; \*\*p < .05; \*\*\*p < .01

Table 19: Effect by Minority Religion, Tobit Model

	Index of Favorability toward Drainage Program
Constant	2.77*** (0.04)
Horizontal Accountability	-0.01 (0.06)
Vertical Accountability	-0.03 (0.06)
Black Sheep	-0.02 (0.06)
Muslim	-0.13 (0.08)
Horizontal Accountability x Muslim	0.23* (0.12)
Vertical Accountability x Muslim	0.29** (0.12)
Black Sheep x Muslim	0.25** (0.12)
N	2,516
Log Likelihood	-3,370.98
Wald Test	9.42 (df = 7)

\*p < .1; \*\*p < .05; \*\*\*p < .01

## 7.8 Variations on Diversity Measure

We chose a 100 meter radius when constructing our measure of diversity at the individual level. However, the results of our analysis do not substantively change if we use other radii, as seen in Table 20. Though there are slight variations, there are no substantive shifts in the treatment effect for other radii.

Table 20: Effect of Treatments, by Diversity at Different Radii

	Index of Favorability toward Drainage Program		
	50 meters	100 meters	150 meters
	(1)	(2)	(3)
Constant	2.71*** (0.05)	2.70*** (0.05)	2.69*** (0.05)
Horizontal Accountability	0.10 (0.06)	0.09 (0.06)	0.11* (0.06)
Vertical Accountability	0.005 (0.06)	0.01 (0.06)	0.02 (0.06)
Black Sheep	0.01 (0.07)	-0.02 (0.07)	-0.01 (0.07)
High Diversity	0.02 (0.06)	0.04 (0.06)	0.06 (0.06)
H.A. x High Diversity	-0.11 (0.09)	-0.11 (0.09)	-0.14 (0.09)
V.A. x High Diversity	0.06 (0.09)	0.06 (0.09)	0.03 (0.09)
B.S. x High Diversity	0.08 (0.09)	0.10 (0.09)	0.09 (0.09)
N	2,468	2,516	2,516
R <sup>2</sup>	0.003	0.004	0.004
Adjusted R <sup>2</sup>	0.0000	0.001	0.002
Residual Std. Error	0.81 (df = 2460)	0.80 (df = 2508)	0.80 (df = 2508)
F Statistic	1.00 (df = 7; 2460)	1.36 (df = 7; 2508)	1.59 (df = 7; 2508)

\*p < .1; \*\*p < .05; \*\*\*p < .01

## 7.9 Caste and Religion

Although caste does not have an overall interaction with the treatments, it does appear that high caste Muslims are the ones primarily driving the result we find in our final model in the paper, as shown in Table 21. This is not pre-registered or theorized, and there are few participants in some of the cells of this analysis for Muslim respondents. For these reasons, we can only speculate on the reasons behind this result. Perhaps it is the case that high status Muslims are more driven to project a positive image of their ingroup and, therefore, have a stronger motivation to respond to the treatments compared to lower status Muslims.

Table 21: Effect by Caste, for Hindus vs. Muslims

	Index of Favorability toward Drainage Program	
	Hindus (1)	Muslims (2)
Constant	2.73*** (0.04)	2.78*** (0.11)
Horizontal Accountability	0.0005 (0.06)	-0.05 (0.16)
Vertical Accountability	-0.02 (0.06)	-0.10 (0.15)
Black Sheep	0.02 (0.06)	-0.01 (0.15)
High Caste	0.11 (0.08)	-0.24* (0.13)
Horizontal Accountability x High Caste	-0.04 (0.12)	0.36* (0.20)
Vertical Accountability x High Caste	-0.03 (0.12)	0.53*** (0.20)
Black Sheep x High Caste	-0.19 (0.12)	0.35* (0.20)
N	1,914	602
R <sup>2</sup>	0.002	0.03
Adjusted R <sup>2</sup>	-0.001	0.01
Residual Std. Error	0.79 (df = 1906)	0.86 (df = 594)
F Statistic	0.63 (df = 7; 1906)	2.19** (df = 7; 594)

\*p < .1; \*\*p < .05; \*\*\*p < .01



## References

- Ali, S. (2002). Collective and Elective Ethnicity: Caste among Urban Muslims in India. *Sociological Forum*, 17(4):593–620.
- 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 & WVSA Secretariat, Madrid, Spain & Vienna, Austria.
- 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., Tengtanga, 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.
- Montalvo, J. G. and Reynal-Querol, M. (2005). Ethnic diversity and economic development. *Journal of Development Economics*, 76(2):293–323.