

# Pre-Analysis Plan: Establishing Responsive Linkages between Voters and Politicians in Pakistan\*

*Draft prepared for EGAP 19, Paris, March 3–4, 2017*

Miriam Golden<sup>†</sup>

Saad Gulzar<sup>‡</sup>

Luke Sonnet<sup>§</sup>

February 27, 2017

*Abstract:* Elections are blunt instruments for rewarding well-performing politicians or for punishing those who fail to perform adequately in office. At the same time, a politician’s ability to respond to voters’ concerns is impeded if he lacks ways to collect information between elections. We test the potential for improving responsiveness and accountability by partnering with politicians to use a cellphone-based technology that allows them to record messages and ask questions in their own voice, and then to send these out to large numbers of voters. Voters in turn can respond to specific questions using their phone’s touch pad. We study if this use of Interactive Voice Response (IVR) technology and the ability to provide feedback improves voters’ sense of efficacy and support for democracy and if markers of electoral accountability are improved.

This document is a pre-analysis plan (PAP) for a pilot field experiment to be conducted in Charsadda (in the province of Khyber Pakhtunkhwa), Pakistan. The PAP will be registered once all items are finalized.

---

\*We are grateful for pilot funding from the International Growth Center (IGC) and the Empirical Studies of Conflict (ESOC) Project at Princeton University, as well as the logistical support of the American Institute for Pakistan Studies (AIPS). The pre-test reported here was registered with the UCLA IRB #17-000182.

<sup>†</sup>Professor of Political Science, University of California, Los Angeles, [golden@ucla.edu](mailto:golden@ucla.edu)

<sup>‡</sup>Ph.D. Candidate, Department of Politics, New York University, [saad.gulzar@nyu.edu](mailto:saad.gulzar@nyu.edu)

<sup>§</sup>Ph.D. Student, Department of Political Science, University of California, Los Angeles, [lukesonnet@ucla.edu](mailto:lukesonnet@ucla.edu)

# Contents

<b>1</b>	<b>Motivation and Literature</b>	<b>2</b>
<b>2</b>	<b>Project Overview and Report from Pre-Test</b>	<b>4</b>
<b>3</b>	<b>Research Design</b>	<b>6</b>
3.1	Sampling . . . . .	6
3.2	Treatments . . . . .	8
3.2.1	Treatment 1 – Initial message and questions . . . . .	8
3.2.2	Treatment 2 – Potential follow-up message . . . . .	9
3.3	Embedded Appeal Experiments . . . . .	10
<b>4</b>	<b>Survey Data</b>	<b>11</b>
4.1	Household and Respondent Characteristics (baseline only) . . . . .	12
4.2	Outcome Variables (baseline and endline) . . . . .	12
<b>5</b>	<b>Methodology</b>	<b>15</b>
5.1	Estimands of Interest . . . . .	15
5.2	Hypotheses . . . . .	15
5.2.1	Evaluations of the Incumbent . . . . .	16
5.2.2	State Legitimacy . . . . .	17
5.2.3	Willingness and Ability to Hold Accountable . . . . .	17
5.3	Estimation . . . . .	17
5.3.1	Intent-to-treat . . . . .	18
5.3.2	Local average treatment effect (compliers) . . . . .	18
<b>6</b>	<b>Experimental Challenges</b>	<b>19</b>
6.1	Attrition . . . . .	19
6.2	Spillovers . . . . .	19
<b>7</b>	<b>Power Calculations</b>	<b>19</b>
7.1	Feeling Thermometer Outcomes . . . . .	20
7.2	Binary Outcomes . . . . .	20
7.3	Agreement Scale Outcomes . . . . .	20
<b>8</b>	<b>Bibliography</b>	<b>21</b>

## 1 Motivation and Literature

In the typical developing country, the structures of political parties are weak, elected representatives do not control robust organizational resources, and the opportunity costs to

voters for collectively sanctioning poor political performance are high. As a result, elections are often ineffective for ensuring accountability. Although in principle voters may be able to “throw the rascals out” if they produce unsatisfactory policy outcomes, prior research documents multiple obstacles that prevent voters from exercising this option effectively. Voters need information about how their elected officials have been performing, clear ideas about performance standards, the means to hold politicians accountable, the will to do so, and the belief that other voters will do so as well (Lieberman, Posner and Tsai, 2014). Usually more than one of these essential ingredients is absent.

The first systematic set of studies that used proper methods of causal identification focused on the putative role of information: the idea was that if voters failed to exercise accountability, they must lack information about politician performance. Although some studies show that providing information to voters about incumbent performance influences electoral choice (Reinikka and Svensson, 2005; Ferraz and Finan, 2008; Banerjee et al., 2010; Grossman and Michelitch, 2016), evidence against a clear and unambiguous effect of information is mounting (Lieberman, Posner and Tsai, 2014; Chong et al., 2015). An active area of research therefore relates to identifying the specific steps in the “accountability chain” that drive positive effects of information, as well as those that cause unintended or even perverse effects.

One important line of work tests the importance of information being shared among voters, or what is called “common knowledge” (Chwe, 2001; Dunning et al., 2015). The theoretical motivation lies with the idea that private information alone is inadequate to generate *coordinated* responses by voters who face badly performing politicians. Coordination in turn is necessary to eject a poor performer from public office, since replacing an incumbent demands that many individuals vote for an alternative candidate. Holding a politician accountable requires not only that individual voters have information about the incumbent’s performance but that they know that other voters have the same information and are prepared to act on it as well. But similarly to simple information studies, those that activate common knowledge treatments also appear to generate mixed outcomes (Adida et al., 2016; Bidwell, Casey and Glennerster, 2015).

Other related studies have investigated various second-order reasons that voters might be reluctant to sanction poorly performing incumbents. Some have examined the importance of electoral alternatives (Chong et al., 2015); others, feelings of personal efficacy (Lieberman, Posner and Tsai, 2014). Some have examined the credibility of the information about politician performance (Malhotra, Michelson and Valenzuela, 2012; Weitz-Shapiro and Winters, 2017), and others the importance of collective expectations (Gottlieb, 2015). Finally, work specifically on credit-claiming demonstrates that it may be the quantity of messages rather than their exact content that is effective in shifting opinion on elected representatives (Grimmer, Messing and Westwood, 2012). All these studies contribute to understanding why simply handing voters information about politicians — such as whether they are doing a good job or whether they are delivering services to constituents — often does not seem to trigger effective and appropriate electoral responses.

Thus far, research has focused almost exclusively on voters, placing heroic burdens on their abilities to receive, process, and use information in a politically effective manner. Little attention has been paid to whether it is possible to use information in ways that align with politician incentives rather than as a blunt instrument of electoral retaliation. The elected official — even one who wants to be responsive to voter interests — confronts an absence of actionable information about and an inability to respond in a timely fashion to voter preferences. The incentive of the politician to be responsive to voters hinges on extracting electoral credit, and thus on credit-claiming; not only does this require that he know what voters prefer, it also requires that he be able to deliver and put his specific stamp on whatever he delivers.

In this study, we aim to move beyond one-way communication between politicians and citizens to study how an on-going *interaction* between the two players might affect outcomes of interest. We test whether improving politician information about and responsiveness to voter preferences substantially affects the ability and willingness of voters to hold their elected officials accountable. Our treatment seeks to leverage information to alter both voter and politician incentives. The intervention is designed to be strategic, in that we (1) provide new channels for voters to communicate with elected officials, giving them a greater incentive to exercise voice; and (2) provide informational “teeth” to the capacity of local officials to respond (Fox, 2015). We experimentally test the importance of responsiveness and related mechanisms of electoral accountability in Khyber Pakhtunkhwa (KP), Pakistan.

## 2 Project Overview and Report from Pre-Test

This pre-analysis plan is for a pilot experiment to be conducted in February–March 2017 that draws on results from a limited pre-test that was conducted over four days in February 2017. The objectives of the pre-test were as follows: first, we wanted to test the technology to ensure that the partners were able to deliver the IVR without problem, and that voters and politicians were able to receive the output. Second, we wanted to check if the low rates of take-up by voters that have been highlighted in the literature on the use of information technology in less developed countries could be improved. We pre-tested the experimental intervention in partnership with one Member of the KP Provincial Assembly (MPA Sultan Mohammed Khan, who represents the constituency of PK-18 Charsadda-II and is affiliated with the Qaumi Watan Party (QWP)). On the technology side, we partnered with VOTO Mobile (<https://www.votomobile.org/>), an organization that operates in developing countries around the world to use cell phones to connect citizens to the organizations that serve them. We report the take-up rates of the pre-test below; other results are not yet available.

The pilot experiment outlined in this pre-analysis plan involves two main waves of treatments, both of which are based on IVR calls recorded in the voice of the sitting MPA.<sup>1</sup> The **first**

---

<sup>1</sup>For the pilot, we continue our partnership with MPA Khan.

set of treatments are an IVR call that either provides a contact between the MPA and his constituents that is largely used to claim credit for services or that simply provides an opportunity for contact between the politician and constituents. In addition, the call optionally allows the MPA to ask questions of constituents regarding upcoming decisions (we label this *contact + questions*). These questions will focus on development spending and priorities over legislative duties or other responsibilities of the MPA. Each treatment condition (*contact* and *contact + questions*) will involve a series of phone calls over time that strengthen the treatment and allow us to test for dosage effects. The **second** set of treatments involves a follow-up call by the same MPA that either is generic in that it does not mention specific activities performed by the MPA, or is responsive, meaning it references the aggregate feedback results available to the politician after the first stage. Again, the number of calls placed in this stage will be varied to study dosage effects.

The pre-test already conducted involved only the the first set of treatments. Field activities proceeded as follows. First, we sent trained enumerators fluent in Pashto to canvass two villages to collect phone numbers of and administer informed consent to 224 male heads of households.<sup>2</sup> Second, we recorded the IVR treatments in collaboration with MPA Khan. Third, working with VOTO, we robo-called the 224 constituents with the IVR recordings. (The scripts of the treatments appear below, in Section 3.2.) In the pre-test, 49 people got *contact*, and 150 people got *contact + questions*. The remainder (25) were controls. Fourth, we asked voters to use their cell phone keypads to respond to the questions. The pilot revealed that 70.3 percent of respondents answered their phones after one or two attempts to reach them. Around 30 percent of respondents who picked up the phone and were asked a question completed the call. We judge this an adequate response rate.

Two components of our design are constructed specifically to mitigate the low response rates experienced by prior SMS-based and other feedback interventions (Humphreys and Weinstein, 2012; Grossman and Michelitch, 2016; Blair, Littman and Paluck, 2017; Open Parliament, N.d.): (1) the use of IVR technology and (2) the request to voters for high-value, actionable information. (In Section 3.3, we also outline some experiments we plan to embed in the pilot that we hope will improve response rates.) We discuss each.

1) **The use of IVR:** IVR offers advantages to both voters and politicians. It permits the politician to initiate a financially costless feedback loop with voters, rather than requesting voters initiate a potentially costly data transfer. With IVR, voters press a number on the cell phone touch pad in response to a verbal question, mitigating issues with literacy (Leo and Morello, August 2015). IVR allows politicians to script the precise information that will be disseminated. This improves its value to them.<sup>3</sup> By collecting cell phone numbers of villagers (which politicians do not have) and structuring an IVR treatment, the intervention

---

<sup>2</sup>We work only with men in this iteration because the enumerators we have been able to recruit thus far are all male.

<sup>3</sup>Surveys using mobile phones also have the potential to be statistically representative (Leo et al., April 2015). However, the absence of a census frame in Pakistan will make it difficult to examine the representativeness of our samples.

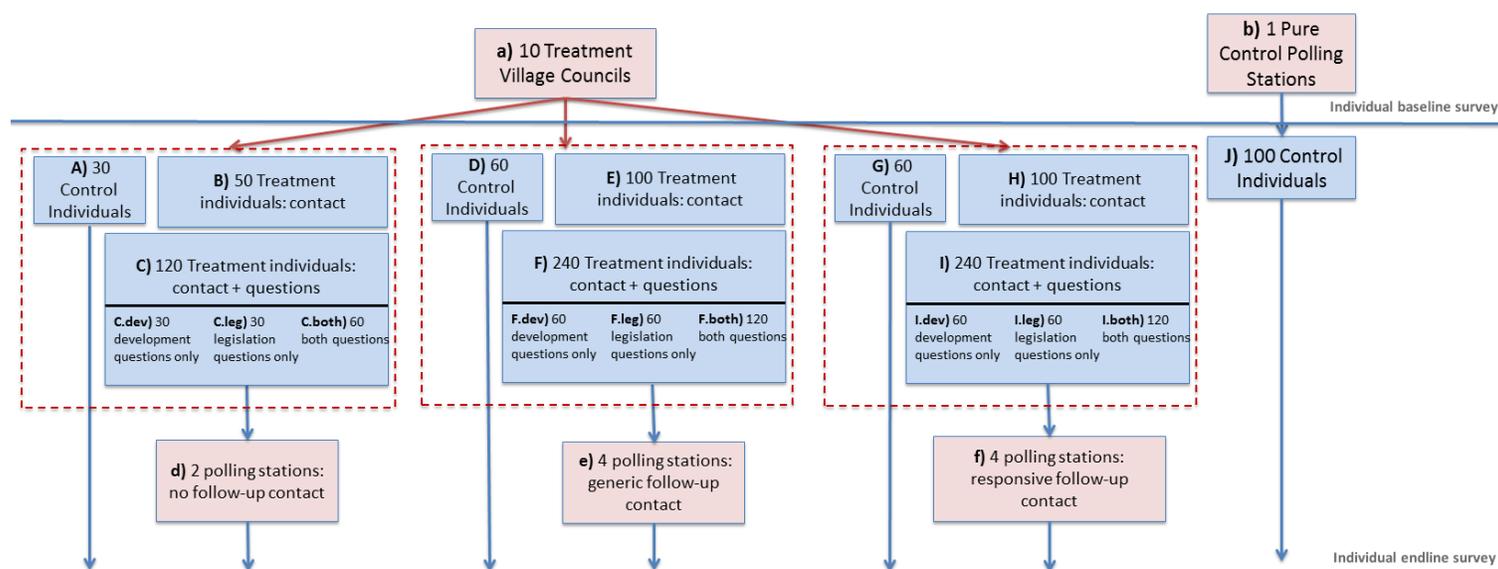
offers them a new technology to collect information they judge of electoral relevance. By allowing politicians to set the informational agenda, we focus the treatment on information to which the politician can easily respond, rather than on ad hoc individualized voter requests that require staff and time, and that reinforce traditional patron-client interactions between politicians and voters.

2) **The provision of high-value actionable information:** In addition, we build on studies that show that crucial (rather than low-value) information can be used successfully to engage citizens (Rezaee, Hasanain and Khan, 2015). Our setup empowers voters around issues of immediate concern, such as the short-term use of development funds, and provides them voice to their elected representatives around these issues. The salience of the issues is baked into our design, because we work with MPs to construct scripts that allow them to credit-claim for activities that they believe are of most interest to voters.

### 3 Research Design

Figure 1 outlines our research design, which we refer to below using the bold letters denoting different treatment conditions.

Figure 1: Pilot Experimental Design



*Note:* This diagram shows the randomization scheme for our pilot. Randomization is at the polling station level. Red boxes denote polling station conditions and blue boxes denote individual level conditions.

### 3.1 Sampling

First, we sample 11 of the Village Council delimited areas out of the 25 within the constituency of the MPA we partner with for the pilot. (We refer to these interchangeably as Village Councils or Village Council areas.) We randomly select one of the Village Council areas to be a pure control (**b**) area. The other ten are assigned to receive some combination of treatments **a**.

We presently lack much data about each of the sampled Village Councils. We have a list of the names of the Village Councils, and for each of these, we have a list of the one to three largest villages. There are unknown numbers of additional smaller settlements within each Village Council.

We are now working to devise a method (using data from Google Maps) to geographically sample within each Village Council so that we randomize the selection of settlement areas. That method will be identified prior to registration of this PAP.

We will send enumerators to administer baseline surveys in each of the sampled settlements, instructing them to use a random walk, such that 100 respondents are interviewed in total in each Village Council area. Within each household, we construct our sample from adult males who possess cell phones, randomly selecting one to be our respondent.<sup>4</sup> (We will provide exact information on the randomization procedure we train our enumerators to use prior to registration of this PAP.) Enumerators conduct a baseline survey to collect phone numbers and pre-treatment covariates, as well as to administer informed consent. (The survey instrument is detailed in Section 4.)

Within each of the ten treatment Village Councils, 15 respondents are sorted into a *control* condition (**A, D, G**), 25 are sorted into *contact* condition (**B, E, H**), and the remaining 60 are sorted into the *contact + questions* condition (**C, F, I**). The *contact* condition involves receiving a call from the MPA that is largely credit-claiming in nature but does not ask for feedback about respondent preferences. The *contact + questions* treatment asks one or two questions of the respondent, encouraging him to provide feedback using his keypad about the decisions the MPA must make in the near future. Within the *contact + questions* condition, (**C, F, I**), we randomize the number and type of questions people get to respond to. Specifically, we ask only a question about development of some (**C.dev, F.dev, I.dev**), we ask only about legislative priorities for others (**C.leg, F.leg, I.leg**), and we ask both questions to the remainder (**C.both, F.both, I.both**). The full text of the questions appears in Section 3.2.1.

Finally, the ten treatment Village Council areas are sorted into one of three follow-up treatment conditions: *control*, *generic*, and *responsive*. Respondents in **A,B,C,D,G** receive no follow-up call, while respondents in **E, F** receive a *generic* follow-up call, and respondents in **H, I** receive a *responsive* follow-up call that directly addresses the demands of voters in

---

<sup>4</sup>Most adult males in the area possess cell phones.

those Village Council districts, acknowledging their input. The three treatment conditions are randomized at the Village Council level and delivered only to individual respondents who received a treatment call in the first stage. Drafts of the scripts for the follow-up calls appear in Section 3.2.2.

## 3.2 Treatments

### 3.2.1 Treatment 1 – Initial message and questions

Our first treatment is a recorded call by the MPA of the constituency we are working in and involves questions about upcoming decisions and credit-claiming for other activities. Voters will either not receive the call at all (control), receive question 1 only (on development spending), question 2 only (on legislative priorities), receive both questions, or receive no questions at all.

- **Introduction:** Assalam U Alaikum. I am Sultan Muhammad Khan, your elected MPA. I am calling you as your elected representative in the provincial assembly. This phone call is a part of a new effort to reach out to my constituents to get their opinions and inform them about my activities. This phone call will not take long.
- **Questions**
  1. As you know, many development works in the constituency have been completed and some are underway. Recently, I have approved some more funds for the constituency. I would like to know your opinion on how to spend this money. I am going to read a list of options to you. Please press the number that corresponds to the option you would like to recommend. Please wait until I have finished the list before choosing.
    - If you prefer roads, press 1
    - If you want to have funeral services area, press 2
    - If you prefer electricity infrastructure like transformers and electric poles, press 3
    - If you want to have paved streets, press 4
  2. As your MPA, I have participated actively in the Khyber Pakhtunkhwa assembly’s discussions. Besides other successes, I have passed a bill to prohibit interest on loans. Secondly, I have also passed a bill to do away with the housing tax in the rural areas of the province. Do you want me to continue spending time in the assembly discussions or to concentrate more on your other issues and concerns?
    - If you prefer I spend more time participating in the assembly discussions, press 1

– If you prefer I spend more time looking at other issues, press 2

- **Thanks:** Thank you for your feedback
- **Closing:** In the end I would like to bring to your notice that in terms of education, I have opened two degree colleges in our constituency, one for boys and another for girls; different schools have been upgraded; and new schools are being constructed. In terms of health care, we are going to build a large hospital in the constituency, and many BHUs [Basic Health Units] have been upgraded to RHCs [Rehabilitation Hospital Units]. In terms of electricity infrastructure we have done a record number of works; installed transformers and electric poles and overhauled the old transformers. In terms of roads, the construction of Takht Bhai's main road, from Razaro to Saro Shah, has been approved and the work is already underway. For the farmers of this constituency, we are building watercourses. To prevent Razaro and Uthmanzai from flooding, we are strengthening the banks of River Jindi. Similarly, we are making progress in other areas such as paving the streets and sewers. As you know I am the only MPA in this constituency who is trying to bring the general public and MPA closer. I have started different ways of communicating with my constituents. I promise you that in coming days I will do everything for the prosperity and development of my constituency. I look forward to your support. Good Bye.

### 3.2.2 Treatment 2 – Potential follow-up message

After the results from the first treatment have been collected and aggregated at the village level to anonymize them, they will be shared with the MPA and his responses will be crafted. The responses will come in two different formats, and respondents will either not receive this follow-up, will receive a follow up with detailed information, or will receive a generic follow-up. The follow-up message will be written with the MPA. The following are the draft scripts we will provide to MPA Khan. We will file

- **Introduction:** Assalam U Alaikum. This is Sultan Muhammad Khan, your elected MPA. I recently called you about my activities in your constituency and I am following up to let you know more about my upcoming plans.
  - **Generic:** Soon I will be allocating my special funds across the sectors of transportation, funeral services, sanitation, and electricity infrastructure. In the near future you should see this spending making real differences in your community. Also, I will continue to spend time both enforcing the legislation I have worked hard to pass as well as on constituency services.
  - **Responsive:** I have carefully reviewed the feedback from my constituents in the last call that I placed. Most people wanted me to spend the special funds on [TOP ANSWER] and therefore I will spend most of the money I have on this sector. I will allocate most of the rest to [SECOND ANSWER]. In another poll,

most of you want me to spend [MORE/LESS] time enforcing the legislation I have passed. I am happy to do so and will update my schedule to make this possible.

- **Closing:** Thank you for your time, and I hope that I will receive your continued support going forward.

### 3.3 Embedded Appeal Experiments

We also embed two smaller experiments into our intervention. These treatments are delivered via SMS messages that will be sent out to some respondents a few hours before they receive the MPA’s phone call. The rationale for the SMS message is that it alerts the respondent to the fact that he should expect to receive a call from his MPA, and it primes him to answer it. The SMS messages will be sent in Pashto, using the latin alphabet.

The evidence from SMS mobilization campaigns on electoral turnout indicates positive effects in developed contexts in both high and low salience elections (Dale and Strauss, 2009; Malhotra et al., 2011) as well as in developing countries (Aker, Collier and Vicente, 2016). However, it appears that SMS messages may be insufficient to prompt citizens to report public service deficiencies or corruption (Blair, Littman and Paluck, 2017; Grossman, Michelitch and Santamaria, 2016). Thus, there is mixed evidence about the effectiveness of SMS messaging, depending on the purpose. We craft messages to study whether some types of messages are more likely to encourage citizens to spend a small amount of time responding to their MPA’s question.

We randomize the content of the SMS messages to estimate effects of two different types of appeals on response rates: 1) an individualistic appeal and 2) a community appeal. We also randomly assign respondents in each condition to receive a message that contains a peer component.

The individualistic/community appeal design is as follows. First, we randomly assign three possible SMS messages to respondents: one that is neutral, one that states the respondent will receive a call from his MPA because the MPA wants to help the respondent and the respondent’s family, and a third that instead states the respondent will receive a call from his MPA because the MPA wants to help everyone in the constituency. (The exact wording is below.) The distinction resonates with the difference between clientelistic and programmatic appeals (Wantchekon, 2003), although it is not identical to it. (We cannot send out pure clientelistic appeals because the subsequent treatment does not involve the delivery of individualized and reversible benefits; that is, it is designed to exclude clientelistic benefits. For terminology, see (Diaz-Cayeros, Estévez and Magaloni, 2016, ch. 1).) We expect that individualized appeals will be of higher value to poorer respondents and community appeals more valued by wealthier respondents, and that we will see take-up rates vary accordingly. These hypotheses draw on a broad literature claiming that clientelism is especially attractive to poorer voters (Stokes et al., 2013).

The scripts of the SMS messages are as follows:

- **Individualistic/Community Appeal:**

- **Neutral:** Hi [name from baseline survey]. In a few hours you will receive a call from your MPA Sultan Muhammed from this number. We hope you can take the call.
- **Individualistic/Community:** Hi [name from baseline survey]. In a few hours you will receive a call from your MPA Sultan Muhammed from this number so that he can [talk to you about how best to help you and your family.]/[talk to you about how to do a good job for everyone in his constituency.] We hope you can take the call.

Second, we randomize whether we mention that the MPA has called other constituents and that a majority of them have responded, predicting that peer effects will push citizens to be more likely to participate.<sup>5</sup> This is a variant of the social pressure model of turnout (Gerber, Green and Larimer, 2008). Specifically, to the above statements, we randomly add the following statement, identified by double brackets:

- **Peer Appeal:**

- Hi [name from baseline survey]. In a few hours you will receive a call from your MPA Sultan Muhammed from this number so that he can [talk to you about how best to help you and your family.]/[talk to you about how to do a good job for everyone in his constituency.] [[Sultan Muhammad has also called a lot of other people and most of them have responded.]] We hope you can take the call.

We measure three different outcomes of the embedded experiments. First, do respondent’s answer the call from the MPA, which comes from the same phone number as the number sending the SMS? Second, what percentage of the call do they listen to before hanging up? Third, do respondent’s call the phone number that sent the SMS before receiving the IVR call?

## 4 Survey Data

In Figure 1, we show two measurement periods at the individual level: a baseline survey and an endline survey.<sup>6</sup> Surveys focus on four separate categories of questions. First, we gather information about assets and political knowledge in order to have some important pre-treatment covariates that will predict the outcomes of interest. Second, we ask respondents

---

<sup>5</sup>The statement reflects the response rates in the pre-test.

<sup>6</sup>The version of this intervention that we hope to undertake at scale (i.e. partnering with almost all the MPAs in the province of KP) will also collect election data at the polling station level from the 2018 general election as well as administer a midline survey.

about their political preferences and levels of support for certain political actors. Third, we ask them about the legitimacy of the state and the democratic process. Fourth, we ask them about their ability and willingness to hold incumbents accountable. We ask most of these questions in both the baseline and the endline surveys, permitting a difference-in-difference analysis for the most important questions in our survey.

The baseline survey is administered in person by a trained enumerator. Resource constraints mean that for the pilot, we administer the endline survey on the phone rather than in person.

#### 4.1 Household and Respondent Characteristics (baseline only)

- *[household\_material]* What type of material is the house made of? Is it fully concrete, concrete with a metal roof, fully mud, mud with a metal roof, does it have a thatched roof, or something else?
- *[household\_motorcycle]* Does the household have a motorcycle or some other motorized vehicle?
- *[household\_television]* Does the household have a television?
- *[know\_pmln]* Is the PML-N currently a member of the KPK government?
- *[know\_president]* Who is currently the president of Pakistan?

#### 4.2 Outcome Variables (baseline and endline)

- Political Preferences
  - *[therm\_mpa]* A feeling thermometer from 1-10 for the MPA
  - *[therm\_party]* A feeling thermometer from 1-10 for the party of the MPA
  - *[party\_support]* Which party do you support the most?
  - *[join\_party, endline only]* Are you interested in signing up for a political party? Which party would you join?
  - *[endorsement\_exp, endline only]* A recent bill [from MPA Sultan Muhammed Khan] did away with housing taxes in rural areas. Some argue that this is fair as it relieves a burden on rural citizens, but others argue it hurts the ability for the KP government to provide services. How do you feel about this bill?
    - \* Strongly agree
    - \* Somewhat agree
    - \* Somewhat disagree

- \* Strongly disagree
- [*legislative\_priorities, endline only*] There are recent efforts to have MPAs focus more on passing laws and improving Khyber Pakhtunkhwa rather than focusing on providing specific projects for their constituency. If a politician has 5 days to spend time on legislation or providing goods for their constituency, how many should be spent on legislation?

- State Legitimacy

- [*agree\_services*] How much do you agree with the following statement: Generally the leaders of provincial government are competent at providing/delivering services such as building roads, schools etc. to the people of their area.
  - \* Strongly agree
  - \* Somewhat agree
  - \* Somewhat disagree
  - \* Strongly disagree
- [*important\_elected*] Is it important to you that Pakistan is governed by representatives elected by the people?
  - \* Very important
  - \* Somewhat important
  - \* Somewhat unimportant
  - \* Very unimportant
- [*peaceful\_action, endline only*] How much do you agree with the following statement: The best way to improve the situation of people like you is through the peaceful development of the state.
  - \* Strongly agree
  - \* Somewhat agree
  - \* Somewhat disagree
  - \* Strongly disagree
- [*state\_looksafter, endline only*] How much does the state look after people like you?
  - \* Not at all
  - \* A little
  - \* A moderate amount

- \* A lot
- \* A great deal
- [*party\_looksafter, endline only*] How much does the QWP look after people like you?
  - \* Not at all
  - \* A little
  - \* A moderate amount
  - \* A lot
  - \* A great deal
- [*mpa\_looksafter, endline only*] How much does MPA Sultan Muhammed Khan look after people like you?
  - \* Not at all
  - \* A little
  - \* A moderate amount
  - \* A lot
  - \* A great deal
- Will and ability for accountability
  - [*follow\_through*] How likely do you think it is that MPA Sultan Muhammed Khan will be able to accomplish his development plans for the constituency?
    - \* Very unlikely
    - \* Unlikely
    - \* Likely
    - \* Very likely
  - [*efficacy*] How much can people like you affect what the government does?
    - \* Not at all
    - \* A little
    - \* A moderate amount
    - \* A lot
    - \* A great deal

- *[voting\_decision, endline only]* Next year, when you go to vote in the general elections for the seat of the MPA, what is the single most important factor that will determine who you vote for? [Note: we will be pretesting this question with open-ended answers before settling on several fixed categories]

## 5 Methodology

The design depicted in Figure 1 allows us to make several comparisons.

### 5.1 Estimands of Interest

The bold letters in the material below refer to the units in Figure 1. Allowing for a slight abuse of notation,  $\mathbb{E}[\mathbf{B}, \mathbf{C}, \mathbf{E}]$  is the mean outcome for units in those blocks.

The randomization of the first IVR call allows us to make the following comparisons:

- 1.1  $\tau = \mathbb{E}[\mathbf{B}, \mathbf{C}, \mathbf{E}, \mathbf{F}, \mathbf{H}, \mathbf{I}] - \mathbb{E}[\mathbf{A}, \mathbf{D}, \mathbf{G}, \mathbf{J}]$ : The effect of any IVR contact vs no contact.
- 1.2  $\tau = \mathbb{E}[\mathbf{B}, \mathbf{E}, \mathbf{H}] - \mathbb{E}[\mathbf{A}, \mathbf{D}, \mathbf{G}, \mathbf{J}]$ : The effect of receiving a credit-claiming contact from MPA.
- 1.3  $\tau = \mathbb{E}[\mathbf{C}, \mathbf{F}, \mathbf{I}] - \mathbb{E}[\mathbf{B}, \mathbf{E}, \mathbf{H}]$ : The effect of contact + questions conditional on contact by MPA.
- 1.4  $\tau = \mathbb{E}[\mathbf{C.dev}, \mathbf{F.dev}, \mathbf{I.dev}] - \mathbb{E}[\mathbf{C.leg}, \mathbf{F.leg}, \mathbf{I.leg}]$ : The effect of being asked about development instead of legislation.
- 1.5  $\tau = \mathbb{E}[\mathbf{A}, \mathbf{D}, \mathbf{G}] - \mathbb{E}[\mathbf{J}]$ : The effect of any spillovers from treated individuals to control individuals in treated Village Council areas.

The randomization of the second IVR call allows us to make the following comparisons:

- 2.1  $\tau = \mathbb{E}[\mathbf{H}, \mathbf{I}, \mathbf{E}, \mathbf{F}] - \mathbb{E}[\mathbf{A}, \mathbf{D}, \mathbf{G}, \mathbf{J}]$ : The effect of receiving both IVR treatments vs the controls.
- 2.2  $\tau = \mathbb{E}[\mathbf{H}, \mathbf{I}, \mathbf{E}, \mathbf{F}] - \mathbb{E}[\mathbf{B}, \mathbf{C}]$ : The effect of receiving a follow-up call conditional on contact by MPA in first stage.
- 2.3  $\tau = \mathbb{E}[\mathbf{E}, \mathbf{F}] - \mathbb{E}[\mathbf{B}, \mathbf{C}]$ : The effect of a generic follow-up call.
- 2.4  $\tau = \mathbb{E}[\mathbf{H}, \mathbf{I}] - \mathbb{E}[\mathbf{E}, \mathbf{F}]$ : The effect of a responsive follow-up call conditional on getting some follow-up call.
- 2.5  $\tau = \mathbb{E}[\mathbf{I}] - \mathbb{E}[\mathbf{A}, \mathbf{D}, \mathbf{G}, \mathbf{J}]$ : The total effect of contact + questions initially and a responsive follow-up.

## 5.2 Hypotheses

We expect our interventions to affect three classes of outcomes. First, we expect contact, especially contact that is part of responsive linkages between politicians and their constituents, to improve the constituents' evaluations of the incumbent. Second, we expect similar effects of our intervention on preferences for peaceful, democratic governance. Third, we expect that the portions of our intervention specifically designed to give voice to constituents will do the most to increase their willingness and ability to hold incumbents accountable.

### 5.2.1 Evaluations of the Incumbent

H1.1 Hearing from your MPA improves perceptions of MPA and incumbent party because it demonstrates they are interested in reaching out and because they successfully claim credit for good behavior.

- Estimands: 1.1 (effect of any contact), 2.1 (effect of getting called in both rounds)
- Outcomes: *therm\_mpa*, *therm\_party*, *endorsement\_exp*, *join\_party* = QWP, *party\_support* = QWP
- Direction:  $\tau > 0$

H1.2 Hearing from an MPA whom you already support will be more effective at increasing support for the incumbent, since politically opposed respondents will discount the credibility of the messages by the incumbent

- Estimands: 1.1 (effect of any contact), 2.1 (effect of getting called in both rounds)
- Outcomes: *therm\_mpa*, *therm\_party*, *endorsement\_exp*, *join\_party* = QWP, *party\_support* = QWP
- Direction:  $\tau(\text{QWP Supporter}) > \tau(\text{Non-QWP Supporter})$

H1.3 Being asked about development, the issue respondents probably care most about, will have greater effects than being asked about legislation

- Estimands: 1.4 (effect of being asked about development rather than legislation)
- Outcomes: *therm\_mpa*, *therm\_party*, *endorsement\_exp*, *join\_party* = QWP, *party\_support* = QWP
- Direction:  $\tau > 0$

H1.4 Being asked questions and getting a direct response to those questions by your MPA makes you feel as though he is not just letting you know about the good work he is doing, but interested in working more for you in the future.

- Estimands: 1.3 (effect of being asked questions), 2.4 (effect of a responsive follow-up versus generic follow-up)
- Outcomes: *therm\_mpa*, *therm\_party*, *endorsement\_exp*, *join\_party = QWP*, *party\_support = QWP*
- Direction:  $\tau > 0$

### 5.2.2 State Legitimacy

H2.1 Hearing from your MPA normalizes peaceful interactions with democratically elected officials.

- Estimands: 1.1 (effect of any contact), 2.1 (effect of getting called in both rounds)
- Outcomes: *important\_elected*, *peaceful\_action*
- Direction:  $\tau > 0$

H2.2 Hearing from your MPA increases the belief that respondents are being cared for by the state

- Estimands: 1.1 (effect of any contact), 2.1 (effect of getting called in both rounds)
- Outcomes: *state\_looksafter*
- Direction:  $\tau > 0$

### 5.2.3 Willingness and Ability to Hold Accountable

H3.1 Being able to provide input, as well as hearing that this input was considered, increases a respondent's sense of political efficacy as well as willingness to judge an incumbent based on how he performs in responding to a constituent's concerns

- Estimands: 1.3 (effect of being asked questions versus any contact), 2.4 (effect of a responsive follow-up versus generic follow-up)
- Outcomes: *efficacy*, *voting\_decision = performance*
- Direction:  $\tau > 0$

## 5.3 Estimation

There are six ways we estimate our treatment effects,  $\tau$ . The first three estimate the intent-to-treat (ITT) effect and do not consider that there may be issues with compliance. The

second three use the same tools in an instrumental variables framework to estimate the treatment effect among compliers, which is a local average treatment effect.

### 5.3.1 Intent-to-treat

For all outcomes, we use a simple difference in means estimator to get the ITT effect

$$Y_i = \tau D_i + \epsilon_i, \quad (1)$$

where  $D_i$  is coded according to the estimand of interest.

For outcomes where we do not have the outcome measured at baseline, we use simple pre-treatment covariate adjustments and block (Village Council) fixed effects to increase the power of our tests:

$$Y_i = \tau D_i + \beta_1 D_i \bar{X}_i + \beta_2 \bar{X}_i + \theta_i + \epsilon_i, \quad (2)$$

where  $\bar{X}_i$  is a vector of centered pre-treatment variables (following [Lin et al. \(2013\)](#)) and  $\theta_i$  is a Village Council fixed effect. While the Village Council fixed effects maximize the value from block randomizing within the 10 treatment areas, we lose any leverage from the control Village Council. Thus we estimate a similar specification that omits Village Council fixed effects.

For outcomes where we have the outcome measured at baseline, we use a difference-in-differences approach to further improve our power, continuing to adjust for pre-treatment covariates:

$$Y_{it} = \alpha_0 D_i + \alpha_1 T_t + \tau D_i T_t + \beta_1 \bar{X}_i + \beta_2 T_t \bar{X}_i + \theta_i + \epsilon_i, \quad (3)$$

where  $Y_{it}$  is the outcome at time  $t = 0$ , the baseline, or time  $t = 1$ , the endline, and  $E_t$  is a dummy for that is 1 at the endline and 0 at the baseline. By interacting the pre-treatment covariates with the time indicator, we allow them to differentially predict the control potential outcomes both pre- and post-treatment. As with the second estimator we consider, we also specify a difference-in-differences model that omits Village Council fixed effects.

### 5.3.2 Local average treatment effect (compliers)

Although we are taking steps to boost compliance, including the advance SMS appeal to respondents that is described in our embedded experiment (see [Section 3.3](#)) and a call-back line, compliance will still be imperfect. To deal with this, we replicate the above specifications exactly but instrumenting for compliance using assignment to treatment. To be precise, we replace  $D_i$  with  $C_i$ , which is whether or not participant  $i$  complied with the experiment, and then instrument for  $C_i$  with  $D_i$ , assignment to treatment. We define compliance as whether the respondent received the treatment by answering the phone call that puts them in a certain treatment condition. This provides us with the local average treatment effect among compliers.

## 6 Experimental Challenges

### 6.1 Attrition

We take two approaches to dealing with attrition. First, we estimate the local average treatment effect among always-reporters if attrition is not predicted by treatment status. Second, because we have quite a few baseline covariates that are relevant to our potential outcomes of interest for all participants in our study, we weight our non-attriters by the inverse probability of attrition, regardless of whether we believe attrition is predicted by treatment status.

### 6.2 Spillovers

While our randomization will be done at the individual level, we are able to use the location of all treated respondents to test for effects of the treatment status of his neighbors. We take two approaches to handling spillovers. First, we take a design-based approach. By assigning one of the eleven sampled Village Councils to be a pure control, we can compare outcomes among the 100 respondents in the pure control Village Council with the outcomes among the 150 control respondents who are in Village Councils with treated units. This allows us to account for the effect of being near treated units when you are a control unit. Second, we also include logged counts of each treatment condition that are within 100 meters of the household. This allows us to estimate if there are spillover effects from neighbors who have a particular composition of treatments.

## 7 Power Calculations

We have completed one set of power calculations via simulation with many of the parameters calibrated using data from the pre-test, and are currently finishing two further sets of power calculations. In this document we demonstrate our ability to uncover treatment effects on feeling thermometers from 1 to 10. Before the final pre-analysis plan is registered, we will also demonstrate our power to detect treatment effects on binary outcomes, such as party support, and treatment effects on four-point scales of agreement (e.g. “Strongly agree”, “Somewhat agree”, “Somewhat disagree”, “Strongly disagree”).

We show our power to detect effects using our simple difference in means specification as well as our second specification that includes baseline pre-treatment covariates as controls.<sup>7</sup> We also consider two different estimands of interest, Estimand 1.1 and Estimand 2.4, corresponding to the effect of any IVR call and the effect of a responsive follow-up call rather

---

<sup>7</sup>The lack of information about the strength of the dependency between baseline and endline measurements of our outcomes of interest precludes meaningful difference-in-differences power calculations.

than a generic follow-up call, two of our main comparisons of interest. We also consider the case where only 70 percent of respondents are compliers, in line with response rates in our pre-test.

## 7.1 Feeling Thermometer Outcomes

Feeling thermometers range from 1 to 10 and are discrete in nature. To capture the data generating process for the feeling thermometer, we draw control and treatment feeling thermometer potential outcomes from a beta-binomial distribution, with nine trials. We add one to make the range go from one to ten. We carefully calibrate this distribution by setting  $\alpha$  and  $\beta$  so that the mean and standard deviation of the control potential outcomes closely mirror the empirical mean and standard deviation of a feeling thermometer for the incumbent from a survey of 224 respondents from Charsadda-II, the pilot district. Furthermore, we simulate baseline pre-treatment covariates, two knowledge questions and two feeling thermometers, to mirror the pre-test data. Then, we use potential outcomes models fit using the pre-test data to simulate outcome data. This allows us to replicate the efficiency gains from using highly predictive pre-treatment covariates. We also add village fixed effects to the simulated data, using the variation across two villages in our pre-test to calibrate the variance of these fixed effects.

Our use of calibration data from the experimental constituency as well as the faithful replication of our outcome data make it likely that these power calculations are quite accurate. Furthermore, because we will also employ difference-in-differences and more than the four pre-treatment covariates used in these simulations, we consider these detectable effect sizes to be conservative.

Thus our minimum detectable effect on a feeling thermometer outcome when considering both main comparisons of interest is around 0.24 standard deviations of our feeling thermometer, or about 0.65 points on the feeling thermometer.

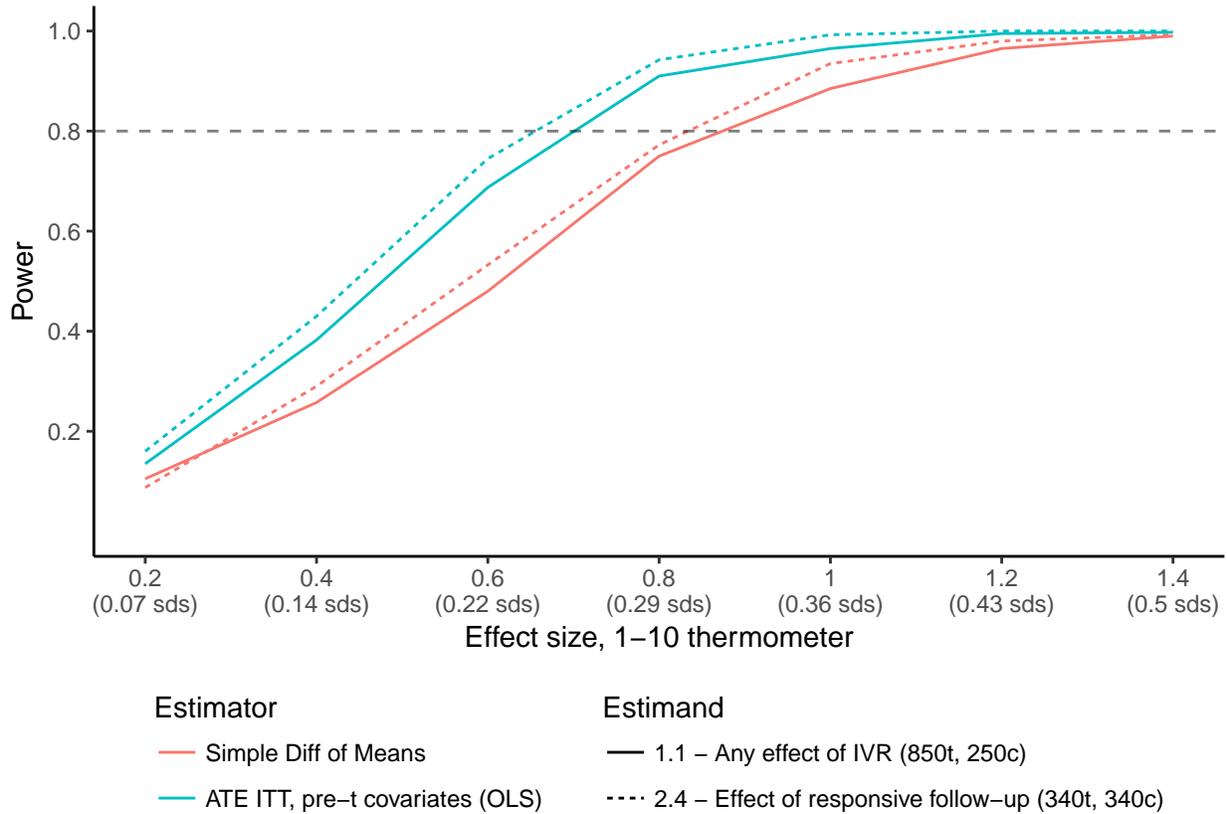
## 7.2 Binary Outcomes

[Forthcoming]

## 7.3 Agreement Scale Outcomes

[Forthcoming]

Figure 2: Power Calculations for Feeling Thermometer



## 8 Bibliography

### References

- Adida, Claire, Jessica Gottlieb, Eric Kramon and Gwyneth H. McClendon. 2016. “When Good News is Bad . . . and When it Isn’t: Voter Coordination, Preferences, and Electoral Behavior.” Unpublished paper.
- Aker, Jenny C., Paul Collier and Pedro C. Vicente. 2016. “Is information power? Using mobile phones and free newspapers during an election in Mozambique.” *Review of Economics and Statistics* (0).
- Banerjee, Abhijit V., Selvan Kumar, Rohini Pande and Felix Su. 2010. “Do Informed Voters Make Better Choices? Experimental Evidence from Urban India.” Unpublished paper.
- Bidwell, Kelly, Katherine Casey and Rachel Glennerster. 2015. “Debates: Voting and Expenditure Responses to Political Communication.” Unpublished paper.
- Blair, Graeme, Rebecca Littman and Elizabeth Levy Paluck. 2017. “Motivating The Adop-

- tion of New Community-Minded Behaviors: An Empirical Test in Nigeria.” Unpublished paper.
- Chong, Alberto, Ana L. De La O, Dean Karlan and Leonard Wantchekon. 2015. “Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice and Participation.” *Journal of Politics* 77(1):55–71.
- Chwe, Michael Suk-Young. 2001. *Rational Ritual: Culture, Coordination, and Common Knowledge*. Princeton: Princeton University Press.
- Dale, Allison and Aaron Strauss. 2009. “Don’t forget to vote: Text message reminders as a mobilization tool.” *American Journal of Political Science* 53(4):787–804.
- Diaz-Cayeros, Alberto, Federico Estévez and Beatriz Magaloni. 2016. *The Political Logic of Poverty Relief: Electoral Strategies and Social Policy in Mexico*. New York: Cambridge University Press.
- Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan D. Hyde, Craig McIntosh, Claire Adida, Eric Arias, Taylor Boas, Mark T. Buntaine, Sarah Bush, Simon Chauchard, Jessica Gottlieb, F. Daniel Hidalgo, Marcus E. Holmlund, Ryan Jablonski, Eric Kramon, Horacio Larreguy, Malte Lierl, Gwyneth H. McClendon, John Marshall, Daniel Nielson, Melina Platas Izama, Pablo Querubin, Pia Raffler and Neelanjan Sircar. 2015. “Political Information and Electoral Choices: A Pre-meta-analysis Plan.” Available at <http://egap.org/registration/736>.
- Ferraz, Claudio and Frederico Finan. 2008. “Exposing Corrupt Politicians: The Effect of Brazil’s Publicly Released Audits on Electoral Outcomes.” *The Quarterly Journal of Economics* 123(2):703–45.
- Fox, Jonathan A. 2015. “Social Accountability: What Does the Evidence Really Say?” *World Development* 72:346–61.
- Gerber, Alan, Donald P. Green and Christopher W. Larimer. 2008. “Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment.” *American Political Science Review* 102(1):33–48.
- Gottlieb, Jessica. 2015. “Greater Expectations: A Field Experiment to Improve Accountability in Mali.” *American Journal of Political Science* 60(1):143–57.
- Grimmer, Justin, Solomon Messing and Sean J. Westwood. 2012. “How words and money cultivate a personal vote: The effect of legislator credit claiming on constituent credit allocation.” *American Political Science Review* 106(4):703–19.
- Grossman, Guy and Kristin Michelitch. 2016. “Transparency and Accountability Initiatives Targeting Politician Performance between Elections: Evidence from a Field Experiment in Subnational Uganda.” Unpublished paper.
- Grossman, Guy, Kristin Michelitch and Marta Santamaria. 2016. “Texting Complaints to

- Politicians: Name Personalization and Politicians' Encouragement in Citizen Mobilization." *Comparative Political Studies* .
- Humphreys, Macartan and Jeremy M. Weinstein. 2012. "Policing Politicians: Citizen Empowerment and Political Accountability in Uganda." Unpublished paper.
- Leo, Ben, Robert Morello, Jonathan Mellon, Tiago Peixoto and Stephen Davenport. April 2015. Do Mobile Phone Surveys Work in Poor Countries? Working Paper N. 398 Center for Global Development.
- Leo, Benjamin and Robert Morello. August 2015. Asking What the People Want: Using Mobile Phone Surveys to Identify Citizen Priorities. Working Paper N. 418 Center for Global Development.
- Lieberman, Evan S., Daniel N. Posner and Lily L. Tsai. 2014. "Does Information Lead to More Active Citizenship? Evidence from an Education Intervention in Rural Kenya." *World Development* 60:69–83.
- Lin, Winston et al. 2013. "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique." *The Annals of Applied Statistics* 7(1):295–318.
- Malhotra, Neil, Melissa R. Michelson and Ali Adam Valenzuela. 2012. "Emails from Official Sources can Increase Turnout." *Quarterly Journal of Political Science* 7(3):321–32.
- Malhotra, Neil, Melissa R. Michelson, Todd Rogers and Ali Adam Valenzuela. 2011. "Text Messages as Mobilization Tools: The Conditional Effect of Habitual Voting and Election Salience." *American Politics Research* 39(4):664–681.
- Open Parliament. N.d. Bridging the gap between citizens and elected representatives. Technical report Available at: <http://openparliament.pk/mps-sindh/>.
- Reinikka, Ritva and Jakob Svensson. 2005. "Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda." *Journal of the European Economic Association* 3(2-3):259–267.
- Rezaee, Arman, Ali Hasanain and Yasir Khan. 2015. "Crowdsourcing government accountability: Experimental evidence from Pakistan." Unpublished paper.
- Stokes, Susan C., Thad Dunning, Marcelo Nazareno and Valeria Brusco. 2013. *Brokers, Voters, and Clientelism: The Puzzle of Distributive Politics*. New York: Cambridge University Press.
- Wantchekon, Leonard. 2003. "Clientelism and voting behavior: Evidence from a field experiment in Benin." *World Politics* 55(03):399–422.
- Weitz-Shapiro, Rebecca and Matthew S. Winters. 2017. "Can Citizens Discern? Information Credibility, Political Sophistication, and the Punishment of Corruption in Brazil." *Journal of Politics* 79(1):60–74.