




Supporting Information: Individualized text messages about public services fail to sway voters: Evidence from a field experiment on Ugandan elections

Ryan S. Jablonski¹  Mark T. Buntaine² 
Daniel L. Nielson³  Paula M. Pickering⁴

¹London School of Economics and Political Science, R.S.Jablonski@lse.ac.uk

²University of California, Santa Barbara, buntaine@bren.ucsb.edu

³University of Texas at Austin, daniel.nielson@austin.utexas.edu

⁴William & Mary, pmpick@wm.edu

May 12, 2021

Contents

| | | |
|-----------|----------------------------------------------------------------------|-----------|
| 1 | Setting | 2 |
| 2 | Description of Sample and Recruitment | 3 |
| 3 | Example Treatment Messages | 5 |
| 4 | Audit Protocol | 8 |
| 4.1 | Roads | 8 |
| 4.1.1 | Protocol Description | 8 |
| 4.1.2 | Enumerator Instructions | 9 |
| 4.1.3 | Enumerator Survey | 9 |
| 4.2 | Health Centres | 10 |
| 4.2.1 | Protocol Description | 10 |
| 4.2.2 | Enumerator Instructions | 10 |
| 4.2.3 | Enumerator Survey | 10 |
| 4.3 | Primary Education | 11 |
| 4.3.1 | Protocol Description | 12 |
| 4.4 | Water Access | 12 |
| 4.4.1 | Protocol Description | 12 |
| 4.4.2 | Enumerator Instructions | 12 |
| 4.4.3 | Enumerator Survey | 13 |
| 5 | Manipulation Checks & Data Reliability | 13 |
| 5.1 | Post-Treatment SMS-Experience Survey | 13 |
| 5.2 | Manipulation check: belief updating | 14 |
| 6 | Validating Measures of Vote Choice | 16 |
| 7 | Results using Pre-Registered Estimation Strategy | 18 |
| 8 | Heterogeneous Treatment Effects | 19 |
| 8.1 | Heterogeneity analysis not pooled by office | 19 |
| 8.2 | Heterogeneity analysis by prior district chair vote margin | 24 |
| 8.3 | Heterogeneity analysis by ruling party (NRM) incumbency | 27 |
| 8.4 | Heterogeneity analysis by trust in Twaweza | 30 |
| 9 | Attribution for Public Service Outcomes | 33 |
| 10 | Coding Audit Performance and Measuring Respondent Priors | 36 |
| 10.1 | Coding Audit Scores | 36 |
| 10.2 | Measuring Respondent Priors | 36 |
| 11 | Balance and Attrition | 38 |
| 11.1 | Balance | 38 |
| 11.2 | Attrition | 39 |
| 12 | Pre-Registration | 40 |
| 13 | Ethics | 41 |

1 Setting

The study took place in Uganda, a semi-authoritarian country with considerable inequities in public service delivery, poverty, and effective governance (Tripp, 2010; Tumushabe et al., 2010). President Yoweri Museveni and his National Resistance Movement (NRM) party have held power since 1986. Our treatments targeted the 2016 elections for district (or LC5/LCV) and sub-county (LC3/LCIII) offices, which are the two most important levels of government for the delivery of basic public services like education, health and water; and are responsible for between 15 and 18% of the government's budget.

We are specifically focused on elections for council and chairperson seats in LC5 and LC3 elections. LC5s are responsible for a district-level administrative area. The chairs of LC5s are elected in a district-wide election of all eligible voters in simple plurality voting. The LC5 councillor seats represent sub-counties within a district and councillors are elected by eligible voters within each sub-county. LC3 chairs and councillors are elected similarly in sub-county-wide and parish-level elections. Local councils also contain representatives for women, youth, persons with disabilities, and older persons, though we did not study these elections. Chair and councillor elections are held concurrently every five years. In 2016, the LC5 elections were held on February 24 and the LC3 elections were held on March 9. The Presidential and MP elections were held separately on February 28. Incumbents do not face term limits.

In our study, we particularly focus on the local provision of water, roads, health and education. Responsibilities for most of these services is, by law, shared between national, district, and sub-county governments (of Uganda, 1997). However, in practice, sub-county roads are almost entirely managed by LC3s and education is managed by LC5s (Green, 2015). Water supply and health clinics are responsibilities shared by LC3 and LC5 governments. Reforms adopted since 2005 have increased the reliance of the LC5 and LC3s on the central government for both their budgets and key bureaucrats who implement local officials' policies (Manyak and Katono, 2011; Raffler, 2018). Village and ward councils (LC1 and LC2) also have input into the budgetary processes at the LC3 and LC5 level, though in practice, the role of LC1s and LC2s has been hampered by a lack of participation, budgetary resources, institutional authority and delays to elections.

The complexity of local governance creates challenges for accountability, not least because Ugandans frequently have doubts about responsibilities for service provision (Bainomugisha et al., 2015; Martin and Raffler, 2021a; Grossman and Michelitch, 2018). As Martin and Raffler (2021b) demonstrates, the attribution problem is further complicated by the strong role that central government bureaucrats play in the procurement process for local services, making it difficult to allocate blame or praise between elected and non-elected officials. Recognizing this complexity, our intervention attempted to clarify responsibility for public services through a brief sentence attributing *de facto* responsibility for the service the respondent chose as most important (Table S2, row 3). However we recognize that even with this information, attribution may remain ambiguous and complex. In the case of health and water services we point out in the messages that responsibility is shared across LC3 and LC5 offices.

Several characteristics of these local elections make them particularly important and illuminating for understanding the effects of political information. First, while elections are generally uncompetitive or manipulated at the national level, political competition at the local level is more open – a characteristic true in many semi-authoritarian contexts (Gandhi and Lust-Okar, 2009). While NRM candidates retain a decided advantage, in 2016 the vast majority of voters had a choice in local government elections. More specifically, 85.0% of LC5 chair elections and 88.8% of LC5 councillor elections in our sample were contested by at least two candidates, while 86.5% of LC5 chair elections and 83.6% of LC5 councillor elections throughout Uganda were contested. 77.0% of LC5 chair elections and 73.0% of LC5 councillor elections were won by NRM candidates. We provide full statistics on partisanship by office in Table S1.

Second, most respondents (73%) in our sample expected these elections to be free and fair. Likewise citizens are more free to express dissatisfaction with local officials; and many civil society groups work openly to promote effective accountability and engagement in local elections. For example, 54 percent of Ugandans reported that they thought their district councils were doing a bad job at maintaining local roads in the 2015 Afrobarometer survey. In addition, voter decisions on the local level are likely less tied to the partisan and identity politics that shape national elections in Uganda (Carlson, 2015). As such, local elections may be considered a more likely case to identify significant effects of new information on Ugandan voters' choices than national elections, such as for parliament or president, which have been explored by some previous studies that did not find positive evidence of information affecting voters' choices (Humphreys and Weinstein, 2013).

Third, our surveys show large gaps in citizen knowledge of the performance of local governments: only 18% of respondents could accurately rank their local public services relative to district averages. Information scarcity has hindered citizens' ability to participate in local government decision-making and hold local officials accountable (Natamba et al., 2010). This lack of understanding likely contributes to what the Uganda Office of the Auditor General has described as poor service delivery (Office of the Auditor General, 2014). Many Ugandan citizens agree with that description, with less than half (46%) of Afrobarometer respondents in Uganda describing the government provision of services as good (Afrobarometer, 2015). One half of respondents in that survey reported having paid a bribe to obtain water or sanitation services, for example.

Given that Ugandans see local services as poor but lack comparative information on their quality and the responsibility of local officials for delivering them, new mobile technology might help overcome this information deficit. Access to mobile phones has proliferated rapidly in Uganda, like in most other places in the world. The International Telecommunication Union estimates that in 2016 there were 55 mobile phone subscriptions for every 100 people in Uganda.¹

We cooperated with Twaweza, a Ugandan non-governmental organization (NGO) that seeks to convey politically relevant information to citizens around both local elections and more generally as part of public affairs. The overall goal of Twaweza throughout the region is to improve governance by making more information available to the public and by improving the ability of citizens to engage in public life. By focusing on local as opposed to national elections, it was possible to conduct an information campaign without threatening the ruling NRM party in a way that might be

¹Data available at the World Development Indicators, <https://data.worldbank.org/> (Accessed May 2019).

dangerous for our partner (or participants). The environment for elections in Uganda has deteriorated since our 2016 investigation (V-dem 2020).

Table S1: Party competition by office

| Office | Parties Contesting | Parties Winning a Seat |
|-----------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------|
| LC5 Chairs | National Resistance Movement (NRM), Forum for Democratic Change (FDC), the Democratic Party (DP), Uganda People's Congress (UPC), Justice Forum (JEEMA), People's Progressive Party (PPP) and Independents | NRM, FDC, UPC, DP, JEEMA, and Independents |
| LC5 Councillors | NRM, FDC, DP, UPC, JEEMA, PPP, Conservative Party (CP), Social Democratic Party (SDP), Uganda Federal Alliance (UFA), and Independents | NRM, FDC, UPC, DP, and Independents |
| LC3 Chairs | NRM, FDC, DP, UPC, JEEMA, PPP, CP, SDP, UFA, Congress Service Volunteers Organisation (COSEVO), Republican Women and Youth Party (RWYP), Uganda Patriotic Movement (UPM), and Independents | |
| LC3 Councillors | NRM, FDC, DP, UPC, JEEMA, PPP, CP, SDP, UFA, COSEVO, RWYP, UPM, and Independents | |

Data source: (Electoral Commission, 2016) The partisanship of winning candidates for LC3 elections is not available from the Ugandan Election Commission.

2 Description of Sample and Recruitment

We employed a nationally representative sample of 28 out of 111 districts (LC5s). Within each district, our partner, Twaweza, randomly sampled 30 villages in proportion to population. Due to its larger population, Twaweza sampled 60 villages in the district of Kampala. Our research team personally visited each village and held a meeting in which an average of 40 citizens per village were recruited and consented to participate in our study. Recruited citizens had to be voting age and have access to a mobile phone. The villages in our sample are plotted in a map of Uganda in Figure S2.

Challenges encountered in the field prevented us from working in two originally sampled districts, Namutumba and Moyo. We randomly selected Kamuli district as a replacement district for Namutumba.² We did not have sufficient time to replace Moyo. Thus we were only able to conduct research in 27 districts. These were Agago, Amolatar, Arua, Buikwe, Bulambuli, Bushenyi, Butambala, Buyende, Gulu, Hoima, Iganga, Kaabong, Kampala, Kamuli, Kasese, Kiruhura, Kisoro, Kumi, Kyegegwa, Lyantonde, Mityana, Mpigi, Nakapiripirit, Nakasongola, Pallisa, Sironko, Zombo.³

In addition to the 30 Moyo villages excluded, 78 villages were inaccessible, largely due to seasonal road conditions. Since treatment assignment was done following the baseline survey, this attrition does not affect within-sample balance.

Our starting sample at field-based intake included 30,296 citizens and 762 villages (27 districts x 30 villages + 30 additional villages in Kampala - 78 inaccessible villages).⁴ 16,083 of those individuals were included in the baseline survey and randomization. Almost all (98.5%) of the people excluded from the study after field-based intake were unreachable by phone, often due to phones being turned off, lacking charge, being out of range, or being out of service. We excluded additional respondents (1.5%) when they asked not to participate further in the study, despite offering their contact information and initial consent in the field. Of the 16,083 people contacted at baseline, we were able to re-contact 12,972 after the LC5 election and 12,874 after the LC3 election.

Figure S1 is a CONSORT diagram that tracks the study design. Attrition was balanced across treatment, control and news type as discussed below in the "Attrition and Balance" section. Respondents were compensated with 1,000 Ugandan shillings for participating in both baseline and endline surveys.

²Mayuge and Kamuli Districts were resampled based on similarity in predominate language. Kamuli District was chosen by a coin toss to replace Namutumba.

³To protect the identify of villages in the sample, the exact names and locations of villages are not part of replication files.

⁴The sample of villages dropped to 753 at baseline (prior to randomization) and 743 at endline. This attrition was due to unreachable subjects and miscoding of village names by enumerators.

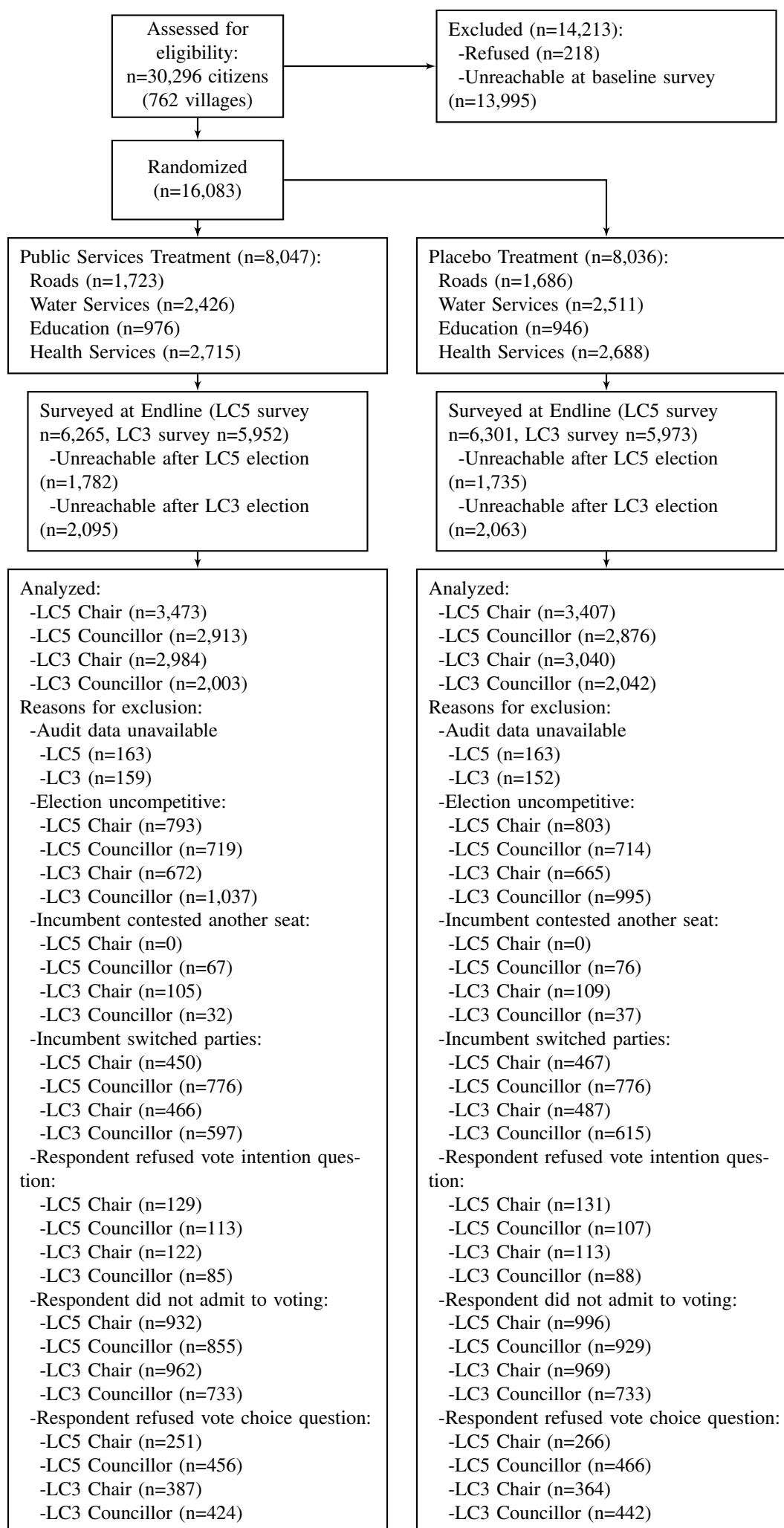


Figure S1: **CONSORT diagram tracking study design.** This diagram shows the number of subjects from intake through analysis, including key steps that where subjects are excluded from the sample, including an inability to reach them for the baseline survey, attrition at the endline survey, village-level exclusion because of party switching that renders incumbency theoretically unclear, or village-level because of uncontested elections. Though we originally sampled 870 villages, we were refused permission to work in 30 villages and rain made an additional 78 villages inaccessible during subject recruitment, which affected the total number of subjects assessed for eligibility.

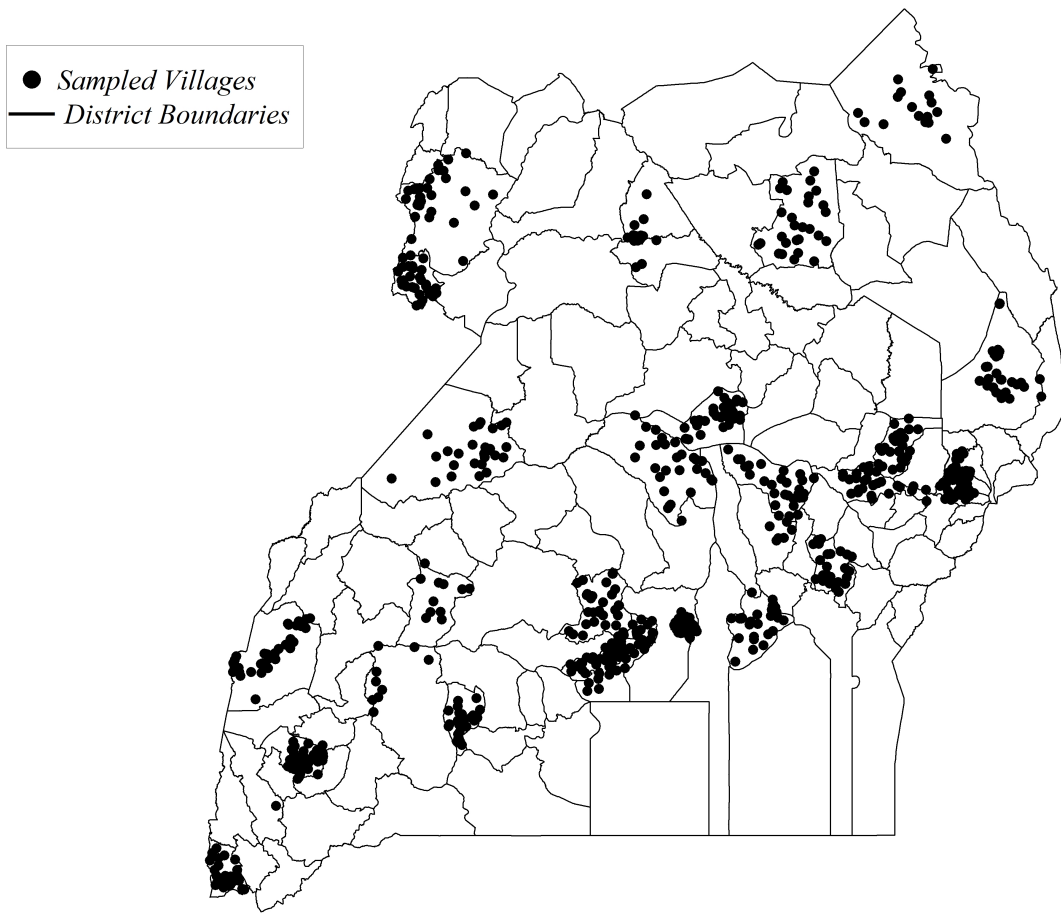


Figure S2: **Map of sampled villages in Uganda.** Population-weighted sample of 30 villages within each of 27 randomly sampled districts. Black dots indicate sampled villages.

3 Example Treatment Messages

A list of example treatment and placebo messages sent to subjects is in Tables S2 and S3. Illustrations of how these messages appeared on phones are in Figure S3.

Table S2: Example Treatment SMS Messages

| | Roads (Much Better) | Water (Much Worse) | Health (Better) | Education (A Little Worse) |
|----|---------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | Dear Francis, welcome to the Twaweza Information Service. We enjoyed meeting you and talking on the phone. | Dear George, welcome to the Twaweza Information Service. We enjoyed meeting you and talking on the phone. | Dear Margaret, welcome to the Twaweza Information Service. We enjoyed meeting you and talking on the phone. | Dear Emma, welcome to the Twaweza Information Service. We enjoyed meeting you and talking on the phone. |
| 2 | Over the next few days we are going to be sending you information about public services as we explained. | Over the next few days we are going to be sending you information about public services as we explained. | Over the next few days we are going to be sending you information about public services as we explained. | Over the next few days we are going to be sending you information about public services as we explained. |
| 3 | You mentioned that you are interested in road quality. Your sub-county roads are managed by your LC3. | You mentioned that you are interested in water access. Protected water access is managed by both your LC3 and LC5. | You mentioned that you are interested in health care. Your Health Centre III is maintained by both your LC3 and LC5. | You mentioned that you are interested in primary education. Primary education in your village is the responsibility of your LC5. |
| 4 | Based on our audits, sub-county roads in your village are much better than other roads in your district. | Based on our audits, water access in your village is much worse than other water access in your district | Based on our audits, the nearest health centre to your village is better than other health centres in your district | Based on Twaweza's Uwezo Study, your village's primary school is a little worse than other primary schools in your district. |
| 5 | A reason your roads are much better is that the number of major obstacles are lower compared to others in your district. | One reason your water access much worse is that the walk time to the water source is a little worse compared to others in your district. | One reason your health centre is better is that cleanliness is a little better compared to others in your district. | A reason your primary school is a little worse is that pupils' math skills are a little worse than average. |
| 6 | Most other roads in your district have 44.3 major obstacles per mile. Your sub-county road has 6 major obstacles per mile. | Most other villages have a walk time of 14 minutes or more to the water source. Your walk time to the water source is 16 minutes. | Most other health centres in your district have cleanliness scores of 11 out of 16. Your centre has 14 out of 16. | Other schools have 4 out of 10 pupils from P3 to P7 who have P2 math skills. Your village's primary school has 3 out of 10 pupils. |
| 7 | A reason your roads are much better is that the road roughness is much better compared to others in your district. | One reason your water access much worse is that the fee per jerrycan is worse compared to others in your district. | One reason your health centre is better is that cleanliness is a little better compared to others in your district. | A reason your primary school is a little worse is that pupils' math skills are a little worse than average. |
| 8 | Road roughness in your road has a score of 6.8/10. Most other roads in your district have a score of 6/10 | Most other water access in your district have a fee of 69 UGX per jerrycan. Your fee per jerrycan is 500 UGX. | Most other health centres in your district have cleanliness scores of 12 out of 16. Your centre has 13 out of 16. | Most other primary schools in your district have 4 out of 10 pupils absent. Your primary school has 3 out of 10 pupils absent. |
| 9 | A reason your roads are much better is that the maintenance of your road is much better compared to others in your district. | Although your water access is much worse the number of people waiting for water is a little better compared to others in your district. | Although your health centre is better, drug availability is much worse compared to others in your district. | Although your village's primary school is a little worse, 2 out of 10 pupils from P3 to P7 can read and understand a P2-level story in English. |
| 9 | Most other roads in your district have a score of at least 3.8/10 in road maintenance. Your sub-county road has a score of 8/10 | Most other villages in your district have 4 people waiting at the water source. Your water source had 4 waiting in line | Most other health centres in your district have 5 recommended drugs available. Your centre has 5. | Most other primary school in your district have 2 out of 10 pupils who can understand a P2 level story in English |
| 10 | We have provided you with information about the quality of roads managed by your LC3. | We have provided you and many of your neighbors with information about the quality of water access maintained by your LC3 and LC5. | We have provided you with information about the quality of health centres maintained by your LC3 and LC5. | We have provided you with information about the quality of primary schools managed by your LC5. |
| 11 | According to our measurements, these roads are much better than other roads in your district. | According to our measurements your water access is much worse than other water access in your district. | According to our measurements your health centre is better compared to other health centres in your district. | According to Twaweza's Uwezo study, your primary school is a little worse than other primary schools in your district. |

Table shows example public services messages sent to respondents in the treatment group. All messages were customized to individual districts based on the results of public service audits. Additional introductory and concluding messages were sent which are not listed here.

Table S3: Placebo SMS Messages

| | Message |
|----|-------------------------------------------------------------------------------------------------------------------------------------|
| 1 | Dear Francis, welcome to the Twaweza Information Service. We enjoyed meeting you and talking on the phone. |
| 2 | Over the next few days we are going to be sending you information about public services as we explained. |
| 3 | Washing hands with soap after using the toilet or before meals can reduce the risk of children getting diarrhea diseases. |
| 4 | Studies in Uganda show that when parents read in English to their children on a regular basis, their students receive higher marks. |
| 5 | Boiling your water for at least three minutes will kill most harmful bacteria from drinking water |
| 6 | Careless or reckless driving contributes to 82 percent of road accidents in Uganda. |
| 7 | Studies in East Africa show that when parents encourage their children to read in English, their students receive higher marks. |
| 6 | Boiling your water for at least three minutes will also kill most harmful parasites and viruses from your drinking water. |
| 7 | Excessive speeding more than doubles the risk of being involved in a fatal accident in Uganda. |
| 8 | Studies in East Africa show that people with strong math skills better manage their family budgets. |
| 9 | Studies in Uganda show that strong math skills are important for economic development. |
| 10 | Studies in Uganda show that when people wash their hands in clean water, they are less likely to contract diseases. |
| 11 | Research suggests that attending regular medical check-ups is important for good health. |
| 12 | Regular check-ups save money and lives in the long-run by testing blood pressure, cholesterol, and for diseases such as HIV. |

Table shows example messages sent to respondents in the placebo group. The order of the messages was randomized across respondents in the placebo group. Additional introductory and concluding messages were sent which are not listed here.



Figure S3: **Treatment messages on typical phones.** When displayed on a smart phone, subjects can read the entire message without scrolling. When displayed on a simple feature phone, messages would often be broken across multiple screens and require scrolling. While we do not have data on recipient experiences with these messages, these considerations should be taken into account when comparing the results of our intervention with other types of information disseminated by mobile phone.

4 Audit Protocol

The co-PIs worked with the field team to develop protocols for auditing each of the public services our preliminary focus groups and Afrobarometer surveys indicated were important to citizens. All enumerators were provided with protocols for auditing roads, health centers, and water access. The audit for these services were conducted between November 29-December 19, 2015 by enumerators after completing recruitment in the villages in the study sample. Enumerators then input audit results into their tablets in the field. The audits of primary education outcomes had already been conducted in 2014 by Twaweza, our local partners. The audits took considerable care and time to conduct. Some audits were not completed as planned because roads to villages were washed out during the field season or because the village-specific public service was not applicable (e.g., we could not locate a relevant health facility for auditing).

4.1 Roads

4.1.1 Protocol Description

The goals of the road audits were to provide: 1) a reliable sample of local road quality that is representative at the district level; 2) a sample of road quality that is representative within target enumeration areas; 3) reliable comparisons between similar road types between enumeration areas within a district; and 4) comparable measurements between enumerators, vehicles and tablets.

Enumerators used an application called RoadLab in order to measure the surface conditions of roads. The app uses the accelerometer, gyroscope and GPS of a tablet or smartphone to measure and evaluate road roughness using the International Roughness Index (IRI). The app can reliably record road conditions when the vehicle is moving at 20 km/hr or greater. For comparable results using multiple devices, the same model of tablet - Nexus 7 (2013 edition) - were used.

The enumerators conducted two measurements: a roughness measurement and a visual inspection. Roughness was measured using RoadLab, which provides a 1-14 measurement where higher values indicate a rougher road. Enumerators also conducted a visual inspection, noting the condition of features such as road-side drainage, shoulders, rail crossing, sidewalks, culverts and bridges. Using a protocol manual, enumerators rated such features on a 0-10 scale, on which higher values indicate more problematic conditions.

Using these two indexes, we created a condition score. The condition score = $100 - (\text{Average Value for Distress} * 5 + \text{Average Value for Roughness} * 10)$. This condition score was the primary means for comparing different enumeration areas.

One challenge we faced is that road types differ between enumeration areas. This is both a measurement challenge and an inferential challenge. At a measurement level, the interval between two condition scores for a paved road is

going to be orthogonal to the interval between two condition scores for a dirt road. At an inferential level, it is not entirely clear whether it is meaningful to compare scores for a dirt road and a paved road as this is likely to correlate very strongly with whether a village is located in a rural or urban area; thus, this may not represent a meaningful distinction for voters.

To address these challenges, we divided roads into three categories: (1) asphalt concrete surfaced roads, (2) bituminous surface-treated roads, (3) gravel roads, and (4) dirt roads. We then calculated condition scores separately for each of these road types. This allows us to compare the condition of similar road types across enumeration areas. Or alternatively, this allows us to weight the index based upon the amount of variation within each road type.

4.1.2 Enumerator Instructions

Enumerators used the following protocol to audit sub-county road quality:

Within each enumeration area (village) drive to the location of the largest trading center. You may need to consult with the LC1 or another credible source to identify the largest trading center. If there are multiple trading centers, select the largest one. If more than one trading center exists that are of similar size in the enumeration area, randomly select a trading center using a coin toss.

Identify a local government road (LC3 road) that begins at the trading center and continues towards the largest population center. Confirm with a local government official or other credible source that the road you plan to audit is a local government (LC3) road. The road you audit should not be a national road. Conduct the road audit – including the Roughness Measurement and Visual Inspection – over a distance of 3 km along this road in the most direct route from the trading center to the next population center. Once you have completed auditing 3 km of road, repeat the road audit on the opposite side of the road on the way back to the trading center.

To take measurements of roughness using the RoadLab application, turn on the RoadLab app. In settings, select the district name and village name where you are conducting the audit. Remember to change the district and village name before you take Roughness Measurements in a new village. Next, place the tablet on the floor of the vehicle between your feet. The tablet should be facing up and as close to horizontal as possible. Just before beginning the drive from the trading center towards the identified population center, manually start tracking by pressing the start-tracking button in the Roughness Evaluation/Bump Detection tab. You can confirm that the app is tracking by checking if the Roughness Evaluation/Bump Detection tab lists your current speed. Do not pick up or tilt the tablet while Roughness Measurements are being taken. This will cause the app to stop taking measurements. Take Roughness Measurements during the entire 6-km round trip (from the trading center out 3 km, and the 3 km back to the trading center). Once you return to the trading center, press the stop recording button to conclude the Roughness Measurement.

Conduct a Visual Inspection of the road from inside the vehicle while the Roughness Measurements are being taken. For the Visual Inspection, record on a piece of paper the number of major obstacles on the road that force the car to slow down or swerve. However, only record the number of major obstacles you encounter when the automobile is moving at speeds less than 25 km/hr. It may help if you ask the driver to inform you when the vehicle is moving at speeds below 25 km/hr. Record major obstacles for the entire 6-km round trip. These obstacles might include any features – such as potholes, washed out shoulders (a road that has been narrowed), stagnant water, uncleared vegetation – that cause the vehicle to slow down or swerve.

Also, when you reach the 3 km turning point, record an estimate of the width of the entire road in meters. Once you return to the trading center after completing the entire Roughness Measurement, start the KoboToolbox app and answer all the questions in the survey.

One of the significant challenges in the field was the poor quality of roads, which made measurements using Roadlab impossible for 40 percent of roads audited. If a road was not passable by car, team leaders conducted a visual inspection by foot and entered information into the KoboCollect road audit survey. Enumerators did this through the following supplemental procedures:

To conduct the visual inspection by foot, walk down the selected sub-county road for 5 minutes, then return to the trading center. During the walk, take note of the following: 1) approximate width of the path at the trading center and turn-around point; 2) any new construction or road/path improvements (if any); 3) the general quality of the road. Once you finish the walk, open up the KoboCollect and answer all the questions that apply. For the question, What type of road is the one being audited?, enter “other”, and describe the road on the next data entry page. For example, you might describe the road as a footpath, or extremely muddy road, or steep muddy road. Mark 0 for the question regarding major obstacles. Answer the rest of the questions based on your visual inspection during your walk. On the last data entry page, indicate that you had issues using RoadLab, and briefly explain why the road was impassable by car.

4.1.3 Enumerator Survey

In addition to the RoadLab data, enumerators completed the following visual survey:

Koboconnect Visual Inspection Survey

Team Leader’s first name:

Name of village (as listed on the village assignment sheet):

Name of the trading center where the road audit begins:

Name of the nearest population center where the sub-county road leads:

1.1. What type of road is the one being audited?

a. tarmac

b. marram

c. dirt

d. Could not determine. Explain []

1.2. How many major obstacles did you encounter on the road over the entire 6-km round trip drive at speeds less than 25 km/hr?

These include any features – such as potholes, washed out shoulders, standing water, uncleared vegetation – that forced you to slow down or swerve.

a. obstacles

d. Could not determine. Explain

1.3. On a scale of 1 to 10 (1 being the best and 10 being the worst), how well would you rate the maintenance of this road relative to other roads you have seen in this district?

b. Could not determine. Explain

1.4. What is the approximate width of the road in meters at the starting point (trading center)?

a. meters

d. Could not determine. Explain:

1.5. What is the approximate width of the road in meters at the turn-around point (3 km from trading center)?

a. meters

d. Could not determine. Explain

1.6. Did you notice any ongoing construction or improvement work on this stretch of road?

a. Yes, explain:

b. No

d. Could not determine. Explain:

1.7. Did you notice any recently completed construction or improvement work on this stretch of road?

a. Yes, explain:

b. No

d. Could not determine. Explain:

1.8. Did you have any challenges operating or calibrating the RoadLab application?

No

Yes, please explain the problem in detail:

4.2 Health Centres

4.2.1 Protocol Description

The goals of audits of sub-county health centres were to provide: 1) a reliable sample of local health care quality that is representative at the district level; 2) a sample of local health care quality that is representative within target enumeration areas; 3) reliable comparisons between similar local health care institutions between enumeration areas within a district; and 4) comparable measurements between enumerators

4.2.2 Enumerator Instructions

Conduct the following three measurements and record the results in KoboCollect using the VoteChoice Healthcare survey:

- **Drug Supply:** Seek out the dispensary within the Health Centre 3. Ask a dispensary staff member to show you what drugs are available in the following drug categories: antimalarial, antibiotic (e.g., erythromycin, septrine), deworming, anti-retroviral, pre-natal vitamins. For each category list whether or not the drug is available or unavailable. Record results in Koboconnect.
- **Clinic Cleanliness:** On a three-point scale in which 1 = clean, 2 = neither particularly clean nor dirty, 3 = dirty, rate the cleanliness of the health clinic's main reception and waiting area, and nearest washroom (for your gender) for the following: (1) floors, (2) walls, (3) furniture (facilities in the washrooms), (4) smell. Record your results in Koboconnect while you are auditing cleanliness.
- **Waiting Time:** Over a 30-minute period – only on Mondays through Fridays and after 10 am – survey all individuals who are leaving the Health Centre asking them how long they waited for their examination. Record results in Koboconnect while you are auditing waiting time.

4.2.3 Enumerator Survey

The Health Centre Survey Instrument is found below:

Team Leader's first name:

Name of the District:

Name of the Sub-county:

Name of the Health Centre 3:

GPS of the dispensary:

Considering the following drug categories – antimalarial, antibiotic, deworming, anti-retroviral, pre-natal vitamins – which of these drugs were available or missing from the dispensary?

Antimalarial: Available Unavailable

Antibiotic (erythromycin, septrine): Available Unavailable

Deworming: Available Unavailable

Anti-retroviral: Available Unavailable

Pre-natal vitamins: Available Unavailable

6) How clean were the floors of the main reception and waiting area?

1 - Clean

2 - Neither particularly clean nor dirty

3 - Dirty

Unable to determine. Reason:

7) How clean were the walls of the main reception and waiting area?

- 1 - Clean
 2 - Neither particularly clean nor dirty
 3 - Dirty
 Unable to determine. Reason: []
- 8) How clean were the furniture of the main reception and waiting area?
 1 - Clean
 2 - Neither particularly clean nor dirty
 3 - Dirty
 Unable to determine. Reason: []
- 9) What was the smell of the main reception and waiting area?
 1 - Good
 2 - Neither particularly good nor bad
 3 - Bad
 Unable to determine. Reason: []
- 10) How clean were the floors of the washroom?
 1 - Clean
 2 - Neither particularly clean nor dirty
 3 - Dirty
 Unable to determine. Reason: []
- 11) How clean were the walls of the washroom?
 1 - Clean
 2 - Neither particularly clean nor dirty
 3 - Dirty
 Unable to determine. Reason: []
- 12) How clean were the facilities of the washroom?
 1 - Clean
 2 - Neither particularly clean nor dirty
 3 - Dirty
 Unable to determine. Reason: []
- 13) What was the smell of the washrooms?
 1 - Good
 2 - Neither particularly good nor bad
 3 - Bad
 Unable to determine. Reason: []
- 14) What was the wait time of the patient 1 in minutes?
 a) []
- 15) What was the wait time of the patient 2 in minutes?
 a) []
- 16) What was the wait time of the patient 3 in minutes?
 a) []
- 17) What was the wait time of the patient 4 in minutes?
 a) []
- 18) What was the wait time of the patient 5 in minutes?
 a) []
- 19) What was the wait time of the patient 6 in minutes?
 a) []
- 20) What was the wait time of the patient 7 in minutes?
 a) []
- 21) What was the wait time of the patient 8 in minutes?
 a) []
- 22) What was the wait time of the patient 9 in minutes?
 a) []
- 23) What was the wait time of the patient 10 in minutes?
 a) []

4.3 Primary Education

Our study's goals of using audit data of primary education outcomes that were conducted and provided by our research partner, Twaweza, were to provide: 1) a reliable measure of educational outcomes for students in grades 3-7 in each village in the sample; 2) reliable comparisons between educational outcomes for students in grades 3-7 across villages within a district; and 3) comparable measurements between enumerators. In addition, we sought to further disseminate data on education that Twaweza had collected.

The data were part of Twaweza's annual Uwezo initiative, which monitors basic literacy and numeracy levels of children in a nationally representative sample of districts across Uganda by conducting an annual household-based survey of 20 randomly selected households in each of the 30 villages sampled per district Twaweza (2017). Though the Uwezo survey collects educational data on all children in the household aged 6-16 years, our investigation uses data on children only in grades 3-7.

Twaweza advised us to use data they collected on pupil attendance, numeracy and English literacy to calculate scores on the quality of primary school education. More specifically, these data included information on whether children attended school on the day before the household survey was conducted; whether the children were able to perform Primary 2 level numeracy tasks up to division; and whether the children were able to read a Primary 2 level story in English.

We used education data collected from the 2013/2014 Uwezo study, with the exception of data for Kamuli district, where we used data collected from the 2011/2012 Uwezo study.

4.3.1 Protocol Description

Uwezo volunteers in each village were provided with a household data sheet. Uwezo enumerators were instructed to list all children aged 3-16 years and to test all children of years 6 to 16 only for reading in literacy (English and local language) and numeracy (mathematics) and the bonus test.

The first set of questions in the household survey on demographics of the family and the schooling status of the children was to be answered by the head of household or other adult.

The second set of questions about attendance at school was to be answered by a child. Specifically, the survey asked:

Did you attend school on Friday?

yes

No

The third part of the survey recorded the child's performance on literacy and numeracy tests administered by the Uwezo enumerator. For the tests used in our study, the survey asked enumerators to tick the highest level for:

English literacy:

1. nothing

2. letter

3. word

4. paragraph

5. story

Numeracy:

1. nothing

2. counting

3. num rec. 10-99

4. addition

5. subtraction

6. multiplication

7. division

4.4 Water Access

4.4.1 Protocol Description

The goals of the water access audit were to provide: 1) a reliable measure of access to a protected public water point from the home of the 3rd subject recruited in each village in the sample, 2) reliable comparisons between access to protected public water points across villages within a district, and 3) comparable measurements between enumerators.

Team leaders were instructed to audit a water access point in each of the villages they visited. The audit included three parameters: 1) the round-trip walking time between the home of the 3rd person recruited and the nearest water access point, 2) the number of people waiting in line at the water access point, and 3) the cost per jerrycan if there was a fee. The instruments used were a smart phone/tablet with GPS and a timer/clock.

4.4.2 Enumerator Instructions

Enumerators used the following protocol to audit water access:

Within each village, ask the third person that is recruited to take you to their neighborhood. Ask the subject for directions to the nearest protected public water access point that is currently functioning. A protected water access point can include a borehole, well (deep hole with a bucket, or pipe delivering water from a spring), tap, or spring box. Start the timer on your phone when you leave the subject's neighborhood and walk at a normal walking speed to the water access point. Make sure to walk all the way to the water access point. While at the water access point, quickly inquire about the fee amount if the water is not free. Quickly count the number of people standing in line (if any), return to the starting point at the subject's neighborhood, and stop the timer. Make sure you record the time required to complete the entire walk from the neighborhood to the water access point and back to the neighborhood. Open up the KoboToolbox app and answer all the questions in the VoteChoiceWaterAccess survey.

One of the challenges enumerators confronted in the field regarding water access was that some of the water access points were constructed by Non-Governmental Organizations (NGO) or built and maintained by individuals. Below are additional instructions provided to team leaders with guidance on this issue:

If the protected water access point is a privately owned, we will not audit it. However, we will audit water access points installed by NGOs since they often work in partnership with the local government to plan and install them. For water access points installed by NGOs, follow the same protocol as before. In KoboCollect, answer the Type of Protected Water Source with "other." Then on the next data entry page, enter in the type of water source and add "NGO." For example, if the NGO-installed water source was a borehole, you would type in "borehole NGO." Fill in the rest of the survey as you normally would.

Another problem was that several of the water access points selected for the audit were located in adjacent villages. Since the local government might allocate water access resources based on the availability of water access points outside of the village, team leaders were directed to audit the closest public or NGO-managed water access point to the selected household, even if it was located outside of the village.

4.4.3 Enumerator Survey

Enumerators then completed the Water Access Survey Instrument below:

Team Leader first name:

Village name (as listed on the village assignment sheet):

Type of protected water source: a) borehole, b) well, c) tap, d) springbox, e) other (explain): []

Total time to walk to the water access point and return to the subject's neighborhood: [] minutes

Number of people waiting in line at the water access point at the time of arrival: []

Cost (if any) for the price/20 litre jerry can: []

GPS of subject's neighborhood: []

5 Manipulation Checks & Data Reliability

We performed several checks to validate subjects' actual participation in voting and to check for bias in reported vote choice. First, we asked respondents about the basin color in their polling station and checked for internal consistency of responses.⁵ Second, due to social desirability, respondents may have over-reported voting for incumbents who performed well and under-reported voting for incumbents who performed poorly. However, assuming the treatment and placebo groups were equally likely to over-report votes for the ruling party, this dynamic would not have altered estimated effects. Our check for reporting bias toward the election winner did not show any differences between responses conducted before and after the release of official results (SI Figure S5).

Of the 16,083 subjects in our experimental sample, we were able to re-contact 12,972 after the LC5 election and 12,874 after the LC3 election. Approximately 20% of the 12-question endline surveys were completed before the official LC5 election results were announced and approximately 35% of the endline surveys were completed prior to the release of the official results of the LC3 elections, with subjects contacted in random order at endline. Findings for surveys conducted before and after election results were released did not differ substantively (SI Table S5). Of respondents who completed the LC5 endline survey, only 26% indicated that they did not see the messages we sent to them; only 17% of subjects in the LC3 endline survey said they did not see our messages. In a call-back survey of 100 randomly selected respondents, most subjects (> 70%) reported that they could easily distinguish our messages from election-related messages sent from political parties and other organizations during the same period (see SI Figure S4).

To maximize institutional memory and retention of respondents, we retained 80% of the enumerators throughout the recruitment, baseline, and endline phases.⁶ Enumerators worked to build trust with subjects by striving to interface with the same subjects throughout all phases of the study. Refusal rates were higher in Kampala, where support for opposition parties is strong. Respondents in the capital were suspicious that our study was affiliated with the ruling party or meant to undermine the election, suspicions that our enumerators worked to dispel in communication.

5.1 Post-Treatment SMS-Experience Survey

To assess the experience of subjects with the messages sent out as part of the treatment, we randomly selected a quota sample of 100 subjects for a post-experimental survey about their experience with Twaweza messages. While we do not distinguish between subjects who received budget and public service messages in this analysis, we find that the large majority of subjects in the experiment found the messages sent out to be memorable, understandable, and valuable.

⁵77-82% of responses were consistent with the modal color selected by respondents in a village. The non-modal answers may also happen due to difference in polling station assignment or the presence of multiple basins. The percentage correct does not vary meaningfully between treatment and control, or between good- and bad-news-eligible subjects (SI Figure S7).

⁶To further encourage respondents to continue participation, we provided 1,000 Ugandan Shillings of airtime for each survey completed. We also had the same RA speaking with the respondent through each round of communication.

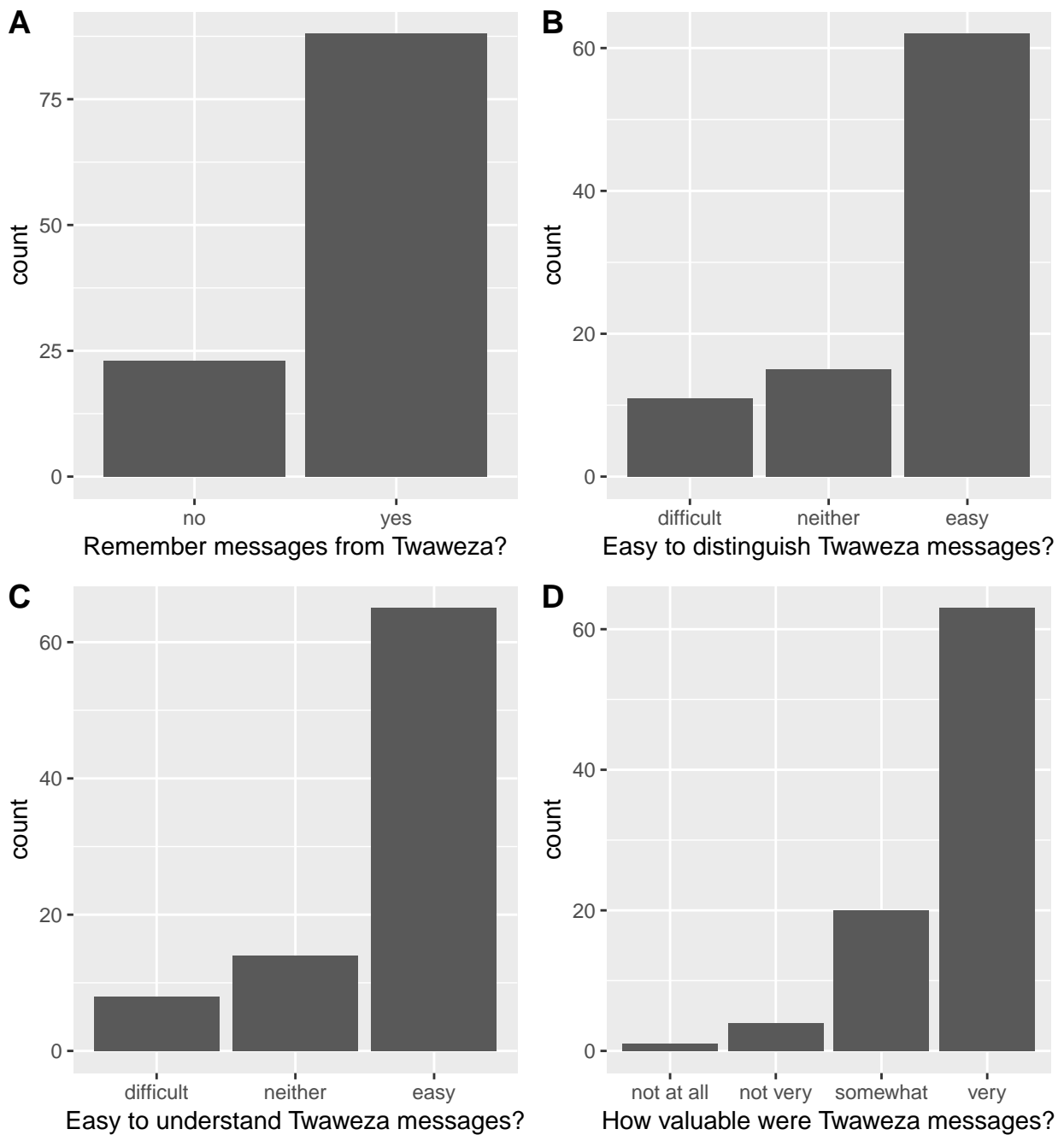


Figure S4: Descriptive data from post-experiment survey on subject experiences with SMS messages.

5.2 Manipulation check: belief updating

We hypothesized that the public service treatment would work by causing subjects to update their beliefs about the comparative quality of public services that their elected officials provided. There are a number of ways that subjects might update their beliefs in response to treatment and we did not pre-specify a single way of testing belief updating. Accordingly, we examine several types of updating outcomes. We report on the rates of four types of updating among treated subjects relative to control subjects. Perfect updating is having correct posterior beliefs in relation to the treatment information. Partial updating is having posterior beliefs that are closer to the treatment information than prior beliefs or having correct prior and posterior beliefs. Loose updating is having posterior beliefs that are not further away from the treatment information than prior beliefs. The eligible group excludes subjects at the extremes of the scale that cannot display more divergent beliefs for loose updating. Directional updating is moving posterior beliefs in the direction of the treatment information or having correct prior and posterior beliefs. Since we only have information about the prior beliefs of subjects who received information about their most preferred public service, we exclude all subjects from this analysis who were assigned to a second-choice service for treatment because of a missing audit.

We find significant evidence of all types of belief updating at the LC5 endline (Table S4), an expansion on the results presented in the main text Figure 3. This provides strong evidence that individuals messages changed relevant political beliefs when they were immediately novel. In contrast, we do not find any evidence of updating at the LC3 endline, indicating that the information in the treatments did not have durable effects on beliefs about the comparative quality of public services (Table S5). As discussed in the main text, this pattern is consistent with related research about the short-lived effects of information within the time period of active campaigns (Gerber et al., 2011).

Table S4: Evidence of subject belief updating at the LC5 endline survey

| Updating Type | (1) | (2) | (3) |
|-------------------------|---------|---------|----------|
| Perfect | 0.093 | 0.089 | 0.089 |
| | (0.009) | (0.009) | (0.018) |
| | p=0.000 | p=0.000 | p=0.000 |
| Partial | 0.087 | 0.085 | 0.085 |
| | (0.010) | (0.010) | (0.017) |
| | p=0.000 | p=0.000 | p=0.000 |
| Loose Updating | 0.048 | 0.048 | 0.048 |
| | (0.009) | (0.009) | (0.014) |
| | p=0.000 | p=0.000 | p=0.001 |
| Loose Updating Eligible | 0.058 | 0.058 | 0.058 |
| | (0.010) | (0.010) | (0.017) |
| | p=0.000 | p=0.000 | p=0.001 |
| Directional | 0.080 | 0.078 | 0.078 |
| | (0.010) | (0.010) | (0.016) |
| | p=0.000 | p=0.000 | p=0.000 |
| Clustering | Indiv. | Indiv. | District |
| Village FEs | No | Yes | Yes |

Two-sided p-values

Table S5: Evidence of subject belief updating at the LC3 endline survey

| Updating Type | (1) | (2) | (3) |
|-------------------------|---------|---------|----------|
| Perfect | -0.001 | -0.000 | -0.000 |
| | (0.009) | (0.009) | (0.011) |
| | p=0.879 | p=0.972 | p=0.978 |
| Partial | 0.011 | 0.011 | 0.011 |
| | (0.010) | (0.010) | (0.010) |
| | p=0.282 | p=0.269 | p=0.280 |
| Loose Updating | 0.001 | -0.000 | -0.000 |
| | (0.010) | (0.009) | (0.011) |
| | p=0.904 | p=0.981 | p=0.984 |
| Loose Updating Eligible | -0.001 | 0.000 | 0.000 |
| | (0.010) | (0.010) | (0.012) |
| | p=0.955 | p=0.961 | p=0.969 |
| Directional | 0.013 | 0.013 | 0.013 |
| | (0.010) | (0.010) | (0.009) |
| | p=0.212 | p=0.207 | p=0.139 |
| Clustering | Indiv. | Indiv. | District |
| Village FEs | No | Yes | Yes |

Two-sided p-values

6 Validating Measures of Vote Choice

An important concern is that respondents may not have accurately reported their vote choices. We provide several figures and tables to assess the trustworthiness of our data. First, we compare self-reported results before and after the announcement of the official results in Figure S5.⁷ The almost perfectly linear relationship suggests there is limited bias in reported voting in favor of the announced winner of the election⁸.

Next, we compare reported results within our sample to the official district and sub-county vote counts in Figure S6. Consistent with accurate reporting, survey-reported voting for the incumbent party is well correlated with official results data.⁹

We also validate the self-reported vote choices by asking subjects to name the color of the water basin used to collect paper ballots within their polling booth. Since subjects did not have time to consult with others in their village about the color, we expect that, on average, the self-reported color should match the modal color named in each village, except in rare instances where voters were assigned to different polling stations. As shown in Figure S7, reports from the LC3 and LC5 elections appear consistent with accurate reporting: 92% and 93% of respondents could name the color of the water basin at their polling station for LC3 and LC5 elections respectively. Of these, 82% and 77% were able to accurately name the color in an internally consistent manner.¹⁰ The accuracy of reporting does not vary meaningfully between treatment and control.

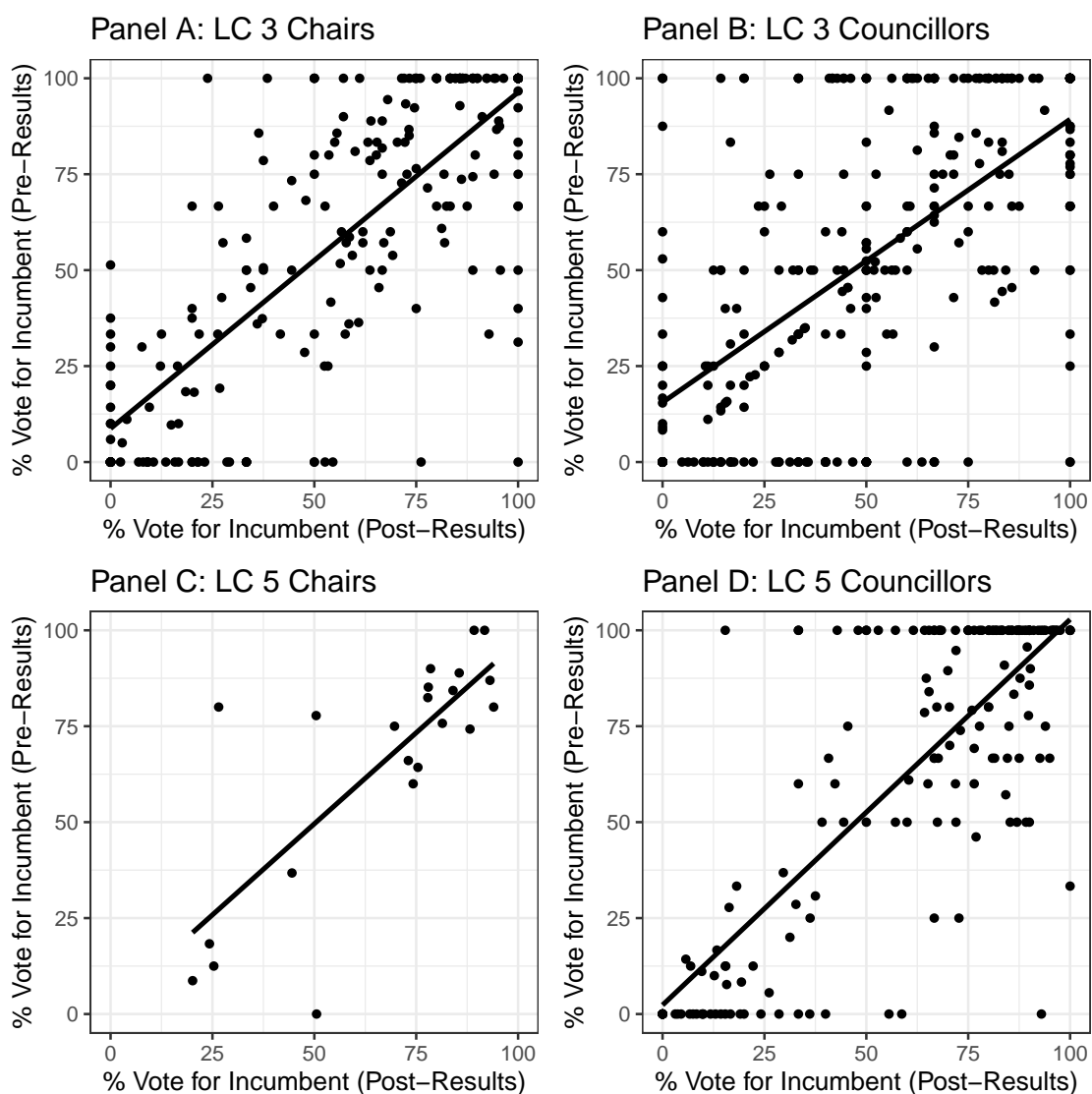


Figure S5: **Comparison of incumbent vote share, pre- and post-election results being announced by politician.** The graphs show the proportion of voters that reported voting for the incumbent in our endline survey, comparing responses after the election results had been announced to responses before the election results had been announced. Note that pre-results reported vote choice proportions cluster at 0% and 100% because of low sample size.

⁷Locally, the release of polling information differed across constituency; however based on discussion with local sources and a review of radio transcripts, we conclude that most results were released in the evening of the day after the election. The release dates for these figures correspond to 17:00 on the day after the election.

⁸lc5 chairs: $\rho = 0.77$; lc5 councillors: $\rho = 0.80$; lc3 chairs: $\rho = 0.78$; lc3 councillors: $\rho = 0.64$.

⁹For LC3 chairs, $\rho = 0.63$ ($p < 0.0001$). For LC3 councillors, $\rho = 0.40$ ($p < 0.0001$). For LC5 chairs, $\rho = 0.50$ ($p = 0.09$). For LC5 councillors, $\rho = 0.61$ ($p < 0.001$).

¹⁰The accurate color is assumed to match the modal color selected by respondents in the village. This may not be accurate when respondents in a village are assigned to different polling stations or when there are multiple basins. While the majority of basins were black (75%), water basins could be blue, brown, green, grey, orange, purple, red, white or yellow.

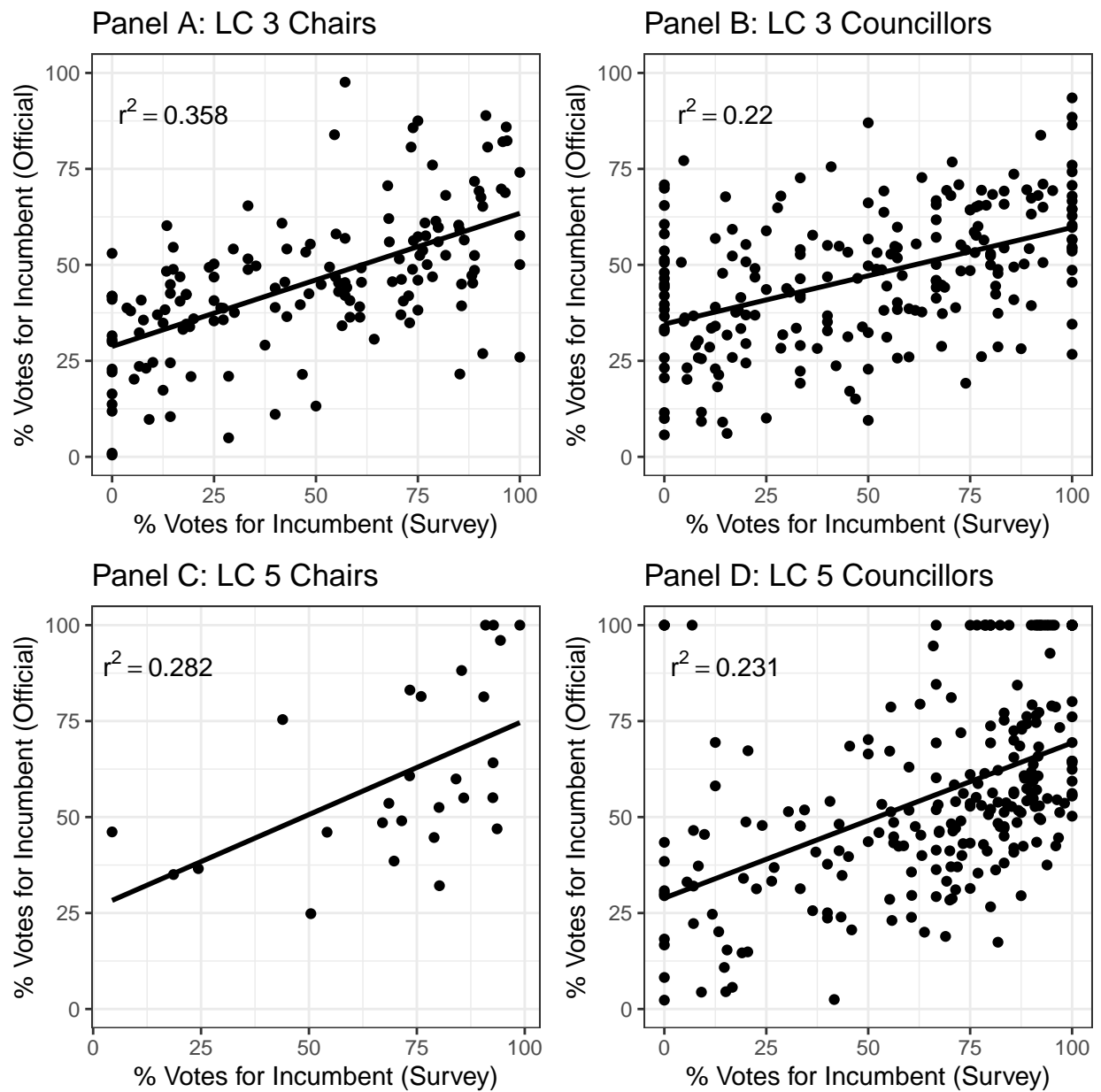


Figure S6: **Official versus reported votes for incumbents** The graphs show the proportion of voters that reported voting for the incumbent party in our endline survey compared the proportion from official voting figures in the relevant constituency for each seat. The LC3 graphs exclude elections where individual incumbents did not run; whereas the LC5 graphs include all competitive elections. We limit the scope of the LC3 analysis due to ambiguities in matching constituencies across electoral commission and census files.

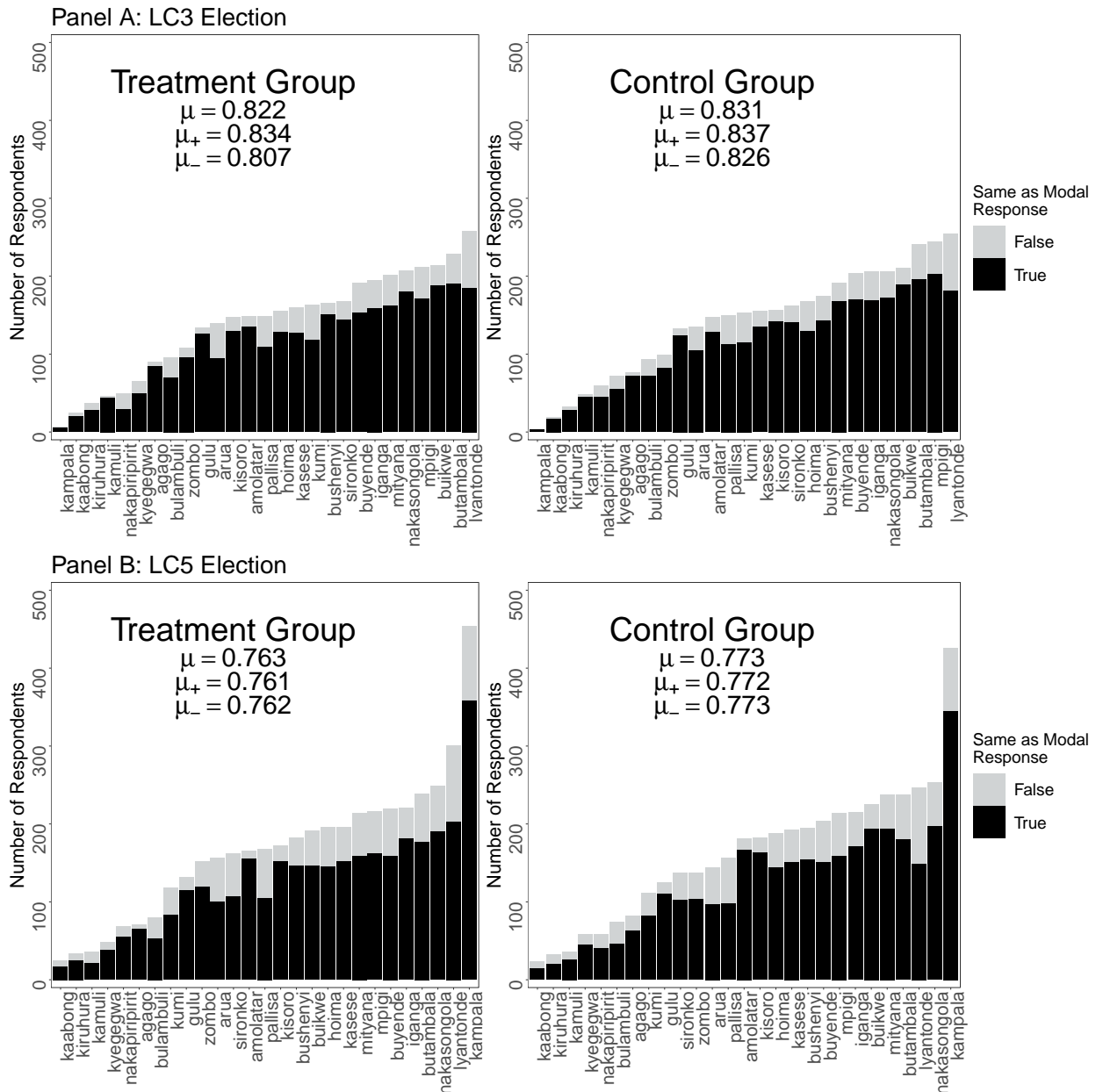


Figure S7: **Reported Polling Station Basin Color for LC3 and LC5 Elections** Our data appear consistent with accurate reporting: 92% and 93% of voting respondents recalled a basin or basin color in their polling station, and among these 82% and 77% were able to accurately name the color in an internally consistent manner with other voters at their polling station. Consistent with low social desirability effects, this figure demonstrates that these proportions do not vary meaningfully across treatment and control, or good and bad news eligibility. μ_+ , μ_- , μ_+ indicate the means of correct responses for good news eligible, bad news eligible and all respondents in each group.

7 Results using Pre-Registered Estimation Strategy

The results that we present in the main text depart from the pre-registered estimation strategy outlined in our pre-analysis plan for the main vote choice and turnout outcomes. In particular, the pre-registered strategy directly transformed the outcome variable by its pre-treatment measure, rather than including that pre-treatment measure as a covariate. As it has been shown that this estimation strategy reduces power Van Breukelen (2006), we deviated from our original plan and instead of directly re-scaling the outcome, used the pre-treatment outcome variable as a covariate in our main results. The pre-registered estimation strategy was instead Eq. 1. As displayed in Figure S8, none of the main results on vote choice or turnout, where pre-treatment measures are available, change with this strategy.

$$y_{ij,t=1} - y_{ij,t=0} = \alpha + \tau_1 T_{ik}^+ + \beta \mathbf{Z}_i + \nu_j + \epsilon_{jh} \quad (1)$$

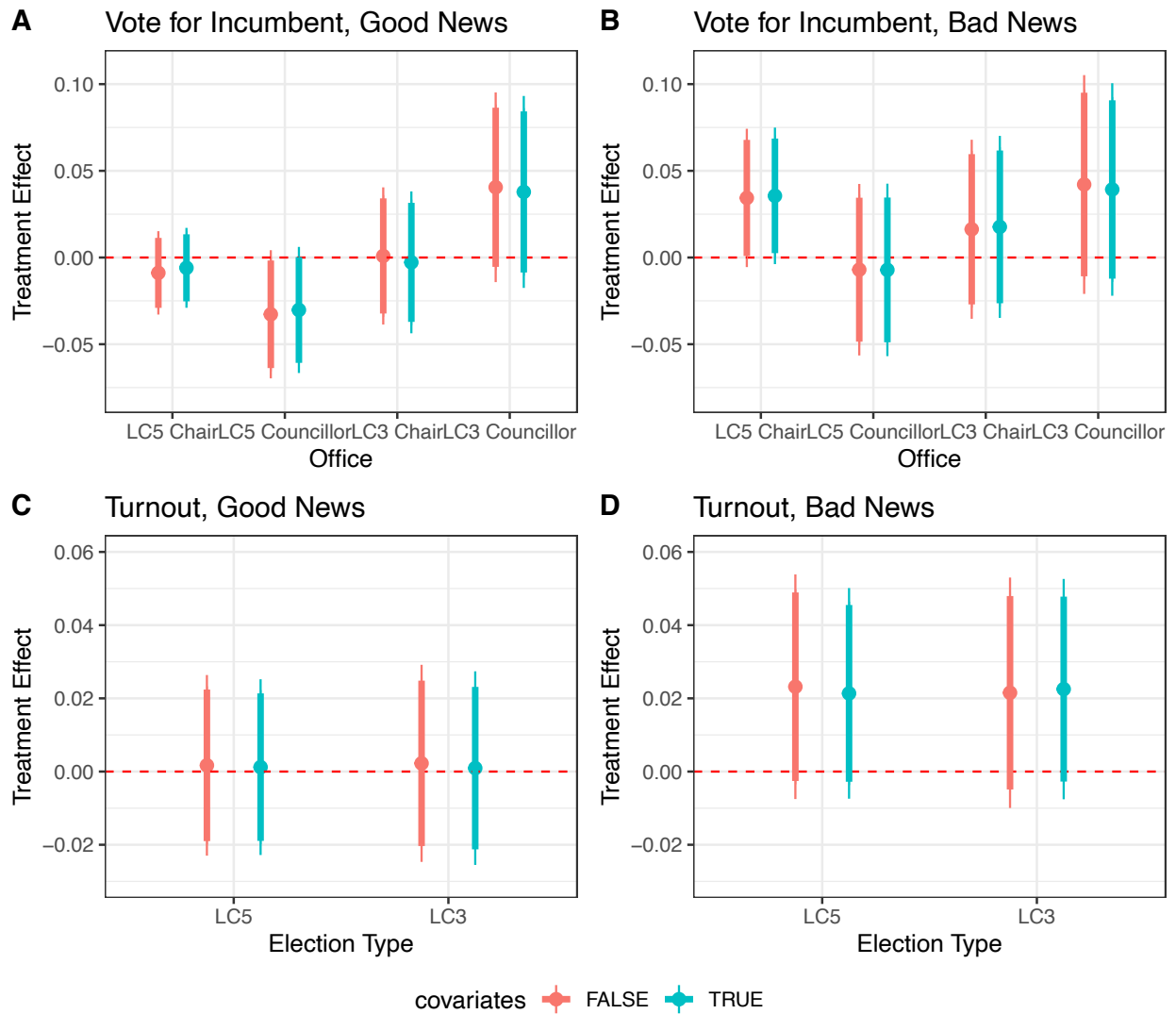


Figure S8: Treatment effects for main vote choice and turnout outcomes estimated as pre-specified.

8 Heterogeneous Treatment Effects

8.1 Heterogeneity analysis not pooled by office

In the main text, Figure 2 shows the effects of treatment in subgroups defined by pre-registered moderators pooled across offices. Our pre-registered analysis strategy did not specify pooling by office. We only pooled to make the totality of the results easier to display succinctly. Figures S9 - S12 display the same results disaggregated to individual offices. There are no notable instances where the results for particular offices diverge from the pooled results in the main text. In no case do the effects of treatment grow in the predicted direction of the moderator, which in the figures is left to right.

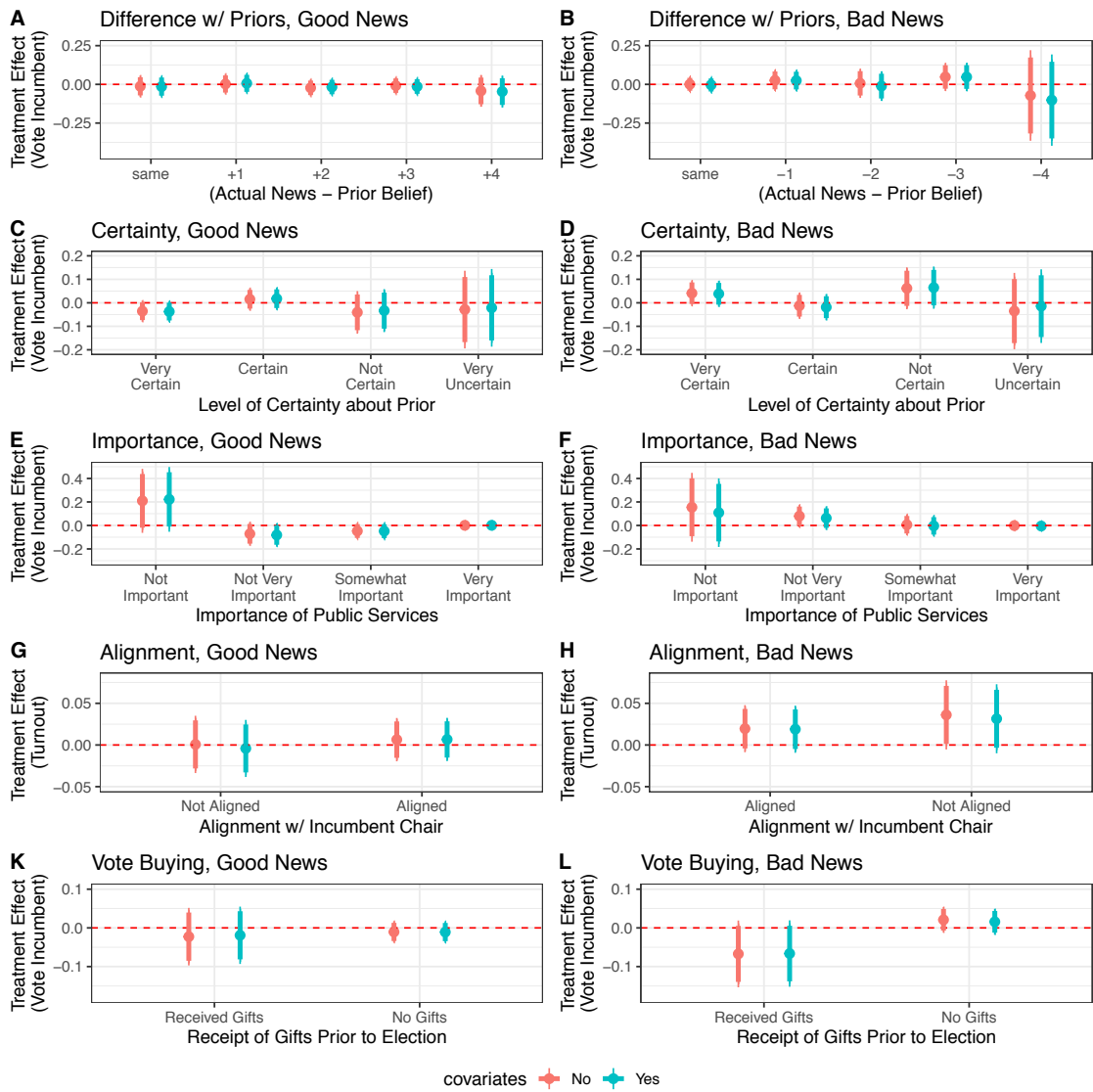


Figure S9: **Conditional effects of treatment with good and bad news subgroups for LC5 Chairperson elections.** *Notes:* Figure includes turnout for LC5 elections in panels **G** and **H**. 95% confidence intervals derived from robust standard errors without clustering. Sample used for estimation of panels **A-F** and **K-L** exclude uncontested elections, elections where the incumbent switched parties, and redistricted constituencies. This sample deviates slightly from the one specified in our pre-analysis plan.

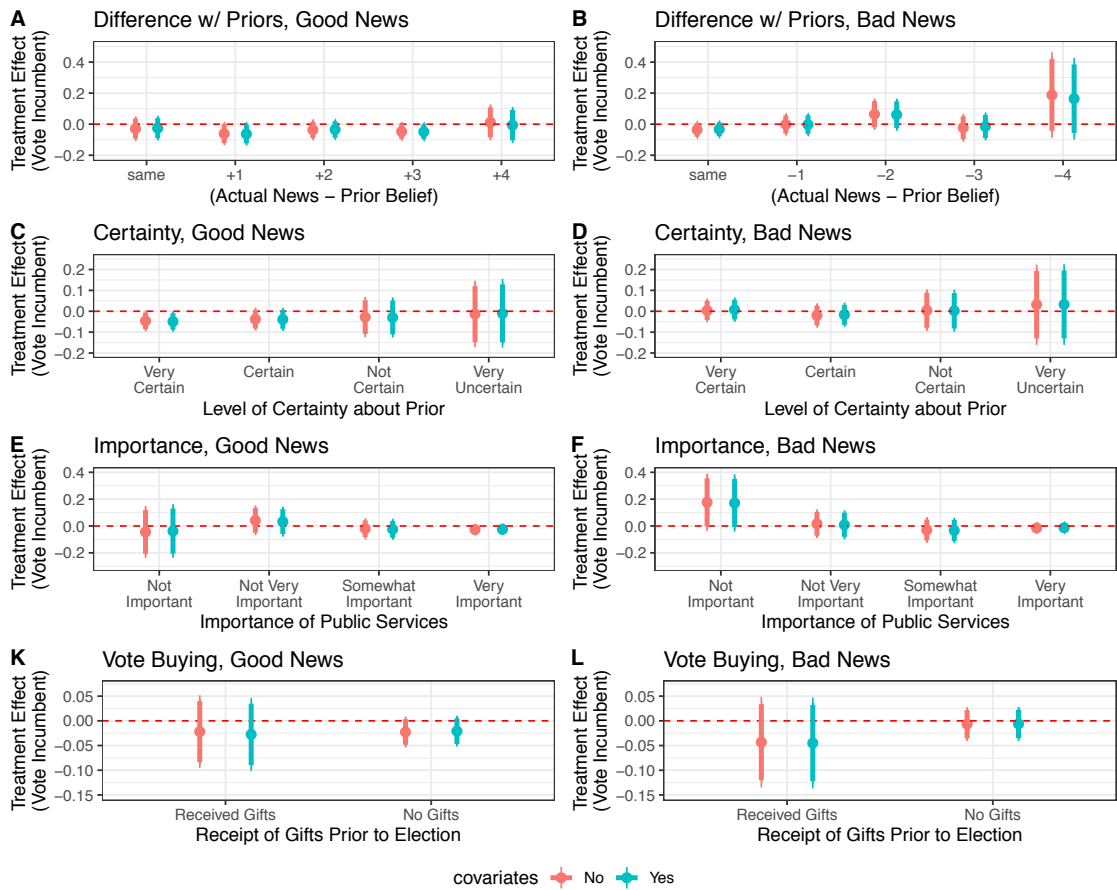


Figure S10: **Conditional effects of treatment with good and bad news subgroups for LC5 Councillor elections.** *Notes:* 95% confidence intervals derived from robust standard errors without clustering. Sample used for estimation of panels A-F and K-L exclude uncontested elections, elections where the incumbent switched parties, and redistricted constituencies. This sample deviates slightly from the one specified in our pre-analysis plan.

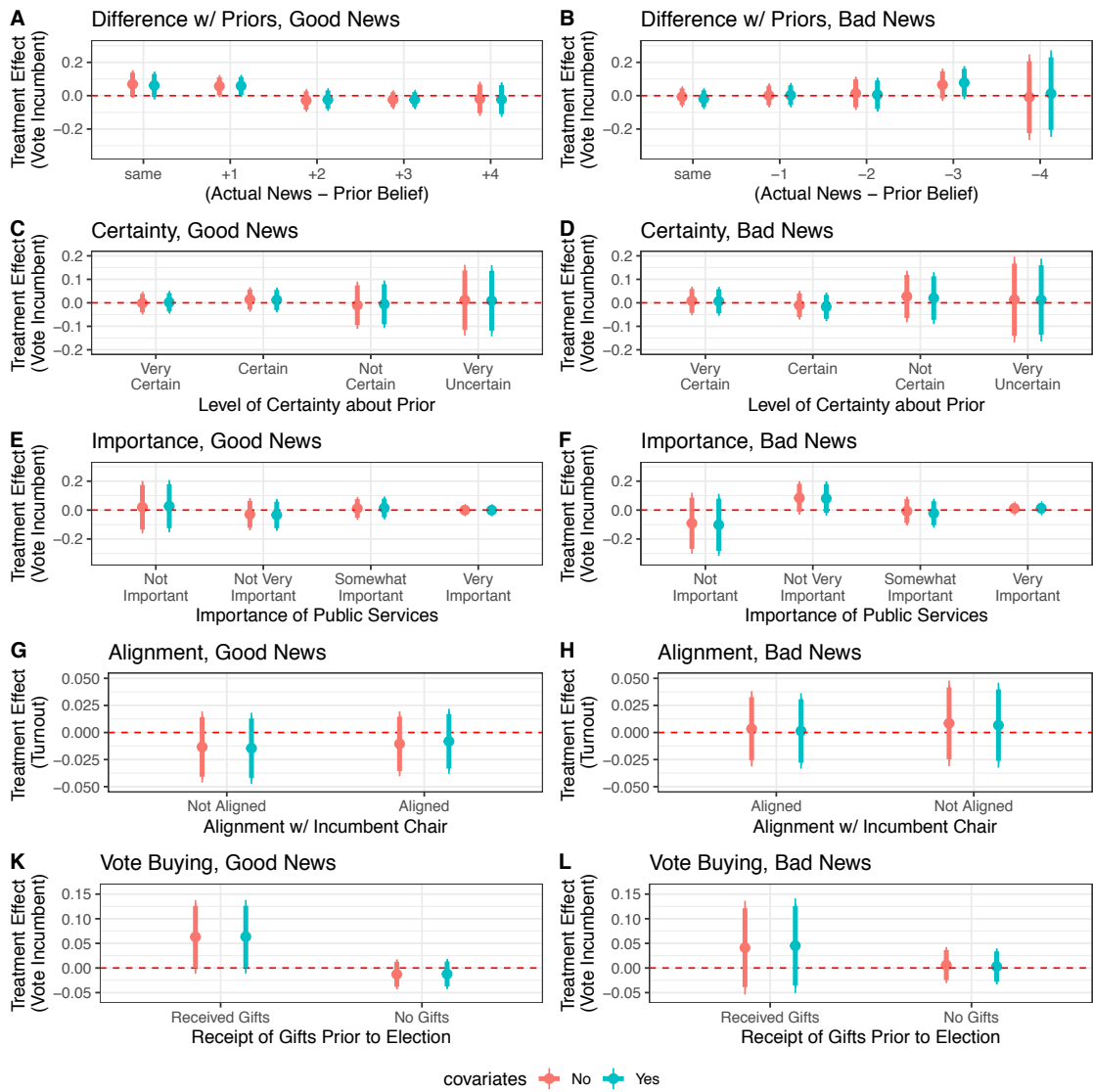


Figure S11: **Conditional effects of treatment with good and bad news subgroups for LC3 Chair-person elections.** *Notes:* Figure includes turnout for LC3 elections in panels **G** and **H**. 95% confidence intervals derived from robust standard errors without clustering. Sample used for estimation of panels **A-F** and **K-L** exclude uncontested elections, elections where the incumbent switched parties, and redistricted constituencies. This sample deviates slightly from the one specified in our pre-analysis plan.

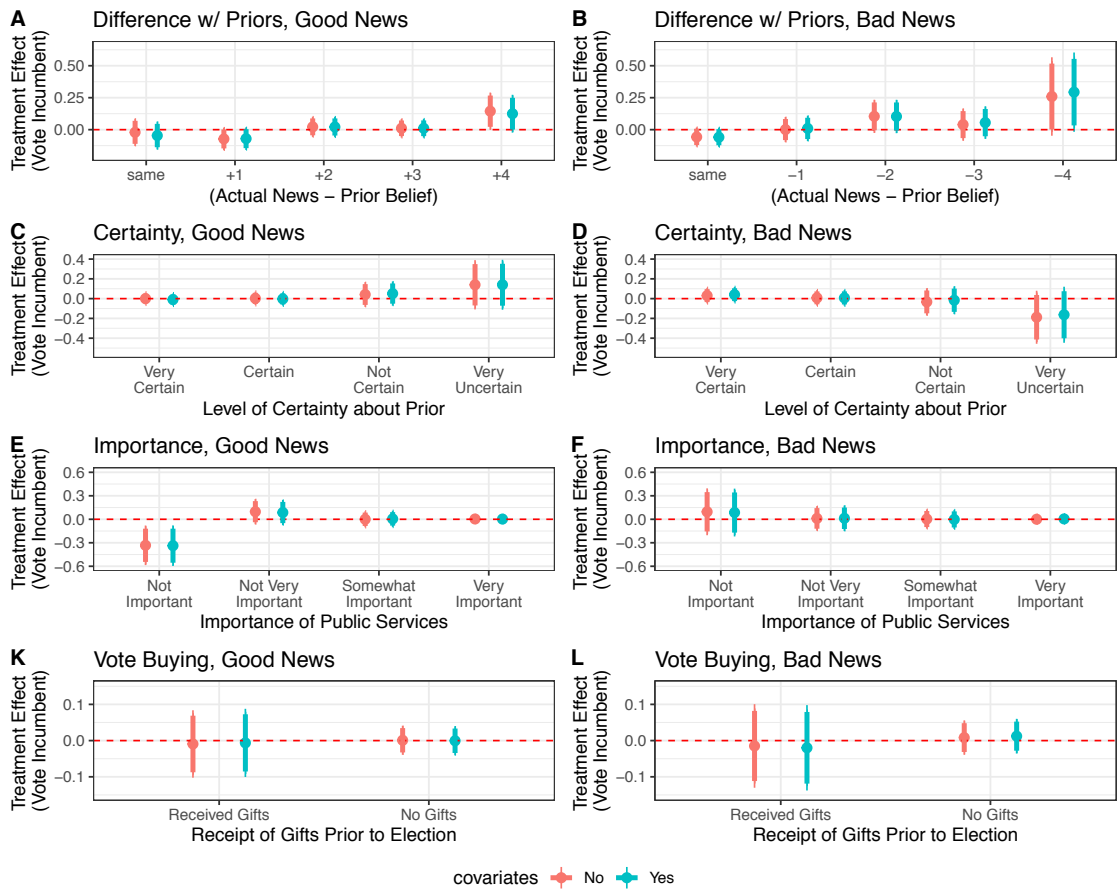


Figure S12: **Conditional effects of treatment with good and bad news subgroups for LC3 Councillor elections.** *Notes:* 95% confidence intervals derived from robust standard errors without clustering. Sample used for estimation of panels **A-F** and **K-L** exclude uncontested elections, elections where the incumbent switched parties, and redistricted constituencies. This sample deviates slightly from the one specified in our pre-analysis plan.

8.2 Heterogeneity analysis by prior district chair vote margin

After pre-registration, we added additional analyses to rule out the possibility that the average effects we report in the main text are obscuring positive treatment effects in subgroups that are more likely to be influenced by information about the programmatic performance of politician. One possibility is that information effects are most likely to emerge in districts with competitive elections, as information might not be persuasive unless deployed in a political climate with viable challengers. We already include only contested elections in the main analysis, but other degrees of competitiveness might also be important. We examine whether the effect of treatment is larger in districts where the previous (in 2011) election for LC5 chair was won by a lower margin, since the data on these elections is complete. We modify the Eq. 1 to include an interaction between Treatment (T_{ik}^+) and the 2011 vote margin for LC5 chair in the district. Note that village fixed-effects absorb the direct effect of prior LC5 vote margin on vote choice. Effectively, this turns the treatment into a dosage based on past vote margin. The direct effect of treatment should be positive in the good news group and negative in the bad news group, since this is the estimate of the information effect when the 2011 LC5 vote margin is held at zero (i.e., when district elections are most competitive). The interaction effect should show a diminishing treatment effect (negative in the good news group and positive in the bad news group) as the previous vote margin grows, if heterogeneity by competitiveness is present.

As displayed in Tables S6 and S7, we do not find evidence consistent with the prediction that information effects are largest in districts with historically competitive elections. The direct effects of treatment (interpretable as treatment effects when the 2011 LC5 chair vote margin is held at zero) are not distinguishable from zero, and the interaction effects between treatment and vote margin do not show diminished treatment effects in the predicted direction (negative for the good news group and positive for the bad news group). Overall, this provides greater confidence that the main results are not averaging over relevant heterogeneity.

Table S6: Heterogeneous effects of treatment by LC5 chair vote margin in 2011, without covariates

| | DV: Vote Choice for the Incumbent; Intention as Control | | | | | | | |
|-------------------------|---------------------------------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|
| | LC5 Chair | LC5 Councillor | LC3 Chair | LC3 Councillor | LC5 Chair | LC5 Councillor | LC3 Chair | LC3 Councillor |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | | | Good News | | Bad News | | | |
| Treatment | -0.026 (0.020) p = 0.211 | -0.020 (0.029) p = 0.482 | 0.021 (0.019) p = 0.276 | -0.014 (0.034) p = 0.687 | 0.014 (0.026) p = 0.602 | -0.015 (0.024) p = 0.542 | -0.028 (0.027) p = 0.298 | 0.012 (0.038) p = 0.742 |
| Treatment X Margin | 0.055 (0.051) p = 0.295 | -0.027 (0.076) p = 0.726 | -0.102 (0.069) p = 0.139 | 0.050 (0.102) p = 0.624 | -0.005 (0.078) p = 0.948 | 0.021 (0.073) p = 0.772 | 0.153 (0.083) p = 0.066 | -0.035 (0.111) p = 0.750 |
| Village fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Covariates | No | No | No | No | No | No | No | No |
| Observations | 3,884 | 3,248 | 3,454 | 2,295 | 2,996 | 2,541 | 2,570 | 1,750 |
| Adjusted R ² | 0.283 | 0.342 | 0.369 | 0.288 | 0.254 | 0.363 | 0.331 | 0.238 |

Note: SEs clustered by politician and two-tailed *p*-values

Table S7: Heterogeneous effects of treatment by LC5 chair vote margin in 2011, with covariates

| | DV: Vote Choice for the Incumbent; Intention as Control | | | | | | | |
|-------------------------|---------------------------------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|
| | LC5 Chair | LC5 Councillor | LC3 Chair | LC3 Councillor | LC5 Chair | LC5 Councillor | LC3 Chair | LC3 Councillor |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | | Good News | | | Bad News | | | |
| Treatment | -0.029 (0.021) p = 0.177 | -0.021 (0.029) p = 0.470 | 0.023 (0.019) p = 0.246 | -0.019 (0.036) p = 0.589 | 0.008 (0.028) p = 0.776 | -0.013 (0.026) p = 0.608 | -0.032 (0.027) p = 0.230 | 0.016 (0.038) p = 0.681 |
| Treatment X Margin | 0.066 (0.053) p = 0.226 | -0.018 (0.076) p = 0.818 | -0.106 (0.067) p = 0.115 | 0.066 (0.107) p = 0.538 | -0.004 (0.081) p = 0.959 | 0.012 (0.075) p = 0.870 | 0.162 (0.082) p = 0.049 | -0.036 (0.109) p = 0.741 |
| Village fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3,884 | 3,248 | 3,454 | 2,295 | 2,996 | 2,541 | 2,570 | 1,750 |
| Adjusted R ² | 0.289 | 0.346 | 0.371 | 0.293 | 0.265 | 0.365 | 0.336 | 0.241 |

Note: SEs clustered by politician and two-tailed *p*-values

8.3 Heterogeneity analysis by ruling party (NRM) incumbency

Uganda is an electoral autocracy and the ruling National Resistance Movement (NRM) party dominates national politics, holding the presidency and a super-majority of elected offices across the country. During the 2016 elections investigated in this study, opposition candidates for national-level offices were routinely harassed and international observers reported many irregularities in the elections.

Because of this political climate, it is possible that the informational treatment would have less of an effect when the incumbent was a member of the ruling NRM party. For these contests, voters could have perceived voting against the incumbent as a safety risk, diminishing the role that information about programmatic performance played in determining vote choice.

We examine whether the effect of treatment is larger in contests where the incumbent is not a member of the NRM party. We modify the Eq. 1 to include an interaction between Treatment (T_{ik}^+) and a binary indicator of whether the incumbent is a member of NRM. Note that village fixed-effects absorb the direct effect of incumbent type on vote choice, so it is not displayed. The direct effect of treatment should be positive in the good news group and negative in the bad news group, since this is the estimate of the information effect when the incumbent is not a member of the NRM. The interaction effect should show a diminishing treatment effect (negative in the good news group and positive in the bad news group) when the incumbent is NRM, if safety concerns blunt the effect of treatment.

As displayed in Tables S8 and S9, we do not find evidence consistent with the prediction that information effects are largest when the incumbent is not a member of the NRM. The direct effects of treatment (interpretable as treatment effects the incumbent is not a member of NRM) are not distinguishable from zero and the interaction effects between treatment and vote margin do not consistently show diminished treatment effects in the predicted direction (negative for the good news group and positive for the bad news group). Overall, this provides greater confidence that the main results are not averaging over heterogeneity in incumbent type.

Table S8: Heterogeneous effects of treatment by NRM incumbency, without covariates

| | DV: Vote Choice for the Incumbent; Intention as Control | | | | | | | |
|-------------------------|---------------------------------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|
| | Good News | | | | Bad News | | | |
| | LC5 Chair (1) | LC5 Councillor (2) | LC3 Chair (3) | LC3 Councillor (4) | LC5 Chair (5) | LC5 Councillor (6) | LC3 Chair (7) | LC3 Councillor (8) |
| Treatment | -0.025 (0.009) p = 0.011 | -0.055 (0.034) p = 0.111 | 0.018 (0.017) p = 0.286 | 0.033 (0.029) p = 0.259 | 0.035 (0.018) p = 0.070 | -0.006 (0.027) p = 0.825 | 0.0003 (0.026) p = 0.990 | -0.016 (0.039) p = 0.687 |
| Treatment X NRM | 0.019 (0.015) p = 0.199 | 0.036 (0.038) p = 0.341 | -0.038 (0.024) p = 0.116 | -0.045 (0.037) p = 0.226 | -0.029 (0.026) p = 0.284 | -0.005 (0.033) p = 0.891 | 0.021 (0.032) p = 0.514 | 0.025 (0.046) p = 0.596 |
| Village fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Covariates | No | No | No | No | No | No | No | No |
| Observations | 3,884 | 3,248 | 3,454 | 2,295 | 2,996 | 2,541 | 2,570 | 1,750 |
| Adjusted R ² | 0.283 | 0.342 | 0.369 | 0.289 | 0.254 | 0.363 | 0.331 | 0.238 |

Note: SEs clustered by politician and two-tailed *p*-values

Table S9: Heterogeneous effects of treatment by NRM incumbency, with covariates

| | DV: Vote Choice for the Incumbent; Intention as Control | | | | | | | |
|-------------------------|---------------------------------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|
| | LC5 Chair (1) | LC5 Councillor (2) | LC3 Chair (3) | LC3 Councillor (4) | LC5 Chair (5) | LC5 Councillor (6) | LC3 Chair (7) | LC3 Councillor (8) |
| Treatment | -0.031 (0.012) p = 0.017 | -0.048 (0.034) p = 0.164 | 0.020 (0.017) p = 0.233 | 0.033 (0.031) p = 0.280 | 0.033 (0.020) p = 0.119 | -0.001 (0.028) p = 0.973 | -0.004 (0.026) p = 0.869 | -0.009 (0.039) p = 0.817 |
| Treatment X NRM | 0.027 (0.018) p = 0.139 | 0.029 (0.038) p = 0.444 | -0.041 (0.024) p = 0.092 | -0.047 (0.038) p = 0.219 | -0.034 (0.027) p = 0.230 | -0.012 (0.034) p = 0.732 | 0.025 (0.032) p = 0.435 | 0.020 (0.047) p = 0.677 |
| Village fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3,884 | 3,248 | 3,454 | 2,295 | 2,996 | 2,541 | 2,570 | 1,750 |
| Adjusted R ² | 0.289 | 0.346 | 0.371 | 0.294 | 0.265 | 0.365 | 0.335 | 0.241 |

Note: SEs clustered by politician and two-tailed *p*-values

8.4 Heterogeneity analysis by trust in Twaweza

One possible reason why information might not have an effect is that recipients do not trust the source of the information. In this section we evaluate the possibility that our results are averaging over heterogeneity in this trust.

Our implementing partner for this study was Twaweza Uganda. Voters were informed that messages were derived from audits conducted in partnership with Twaweza when providing informed consent; as well as in the introduction to the treatment messaging. It is possible that respondents who trust Twaweza less are also less likely to respond to treatment. Twaweza is, on average, fairly well respected in our sample: 35 % of respondents reported a lot of trust in information received from Twaweza; 45% reported a little trust; 4% reported no trust; 15% did not know.

We examine heterogeneity in responses to the messages based on baseline trust in Twaweza as a source of information. To do so, we set the baseline category as any response other than “trust a lot” when asked about trust in Twaweza as an information source. The moderator variable *Trust Twaweza* is positive when a respondent reported a lot of trust in Twaweza as an information source. We then interact this binary variable with the treatment indicator. If trust in the information source is important in determining voters’ responses, the results should show null effects in the direct treatment indicator, positive effects in the interaction between treatment and trust in the good news group, and negative effects in the interaction between treatment and trust in the bad news group.

Tables S10 and S11 show no evidence that trust in Twaweza as an information source moderates the effect of treatment on vote choice for the incumbent for any office. Overall, this provides greater confidence that the main results are not averaging over heterogeneity in trust in the information source.

Table S10: Heterogeneous effects of treatment by trust in Twaweza as an information source, without covariates

| | DV: Vote Choice for the Incumbent; Intention as Control | | | | | | | |
|---------------------------|---------------------------------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|
| | Good News | | | | Bad News | | | |
| | LC5 Chair | LC5 Councillor | LC3 Chair | LC3 Councillor | LC5 Chair | LC5 Councillor | LC3 Chair | LC3 Councillor |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Treatment | -0.004 (0.014) p = 0.797 | -0.035 (0.020) p = 0.090 | -0.019 (0.016) p = 0.238 | 0.023 (0.022) p = 0.282 | 0.017 (0.021) p = 0.432 | -0.001 (0.020) p = 0.942 | 0.026 (0.021) p = 0.226 | 0.026 (0.026) p = 0.327 |
| Treatment X Trust Twaweza | -0.019 (0.024) p = 0.437 | 0.019 (0.035) p = 0.580 | 0.036 (0.028) p = 0.202 | -0.061 (0.035) p = 0.086 | -0.015 (0.024) p = 0.544 | -0.022 (0.031) p = 0.472 | -0.034 (0.036) p = 0.340 | -0.061 (0.043) p = 0.163 |
| Village fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Covariates | No | No | No | No | No | No | No | No |
| Observations | 3,884 | 3,248 | 3,454 | 2,295 | 2,996 | 2,541 | 2,570 | 1,750 |
| Adjusted R ² | 0.283 | 0.342 | 0.369 | 0.289 | 0.255 | 0.363 | 0.331 | 0.240 |

Note: SEs clustered by politician and two-tailed p-values. Baseline category for moderator is any answer except trusting Twaweza for information a lot, including trusting a little, trusting not at all, and do not know responses.

Table S11: Heterogeneous effects of treatment by trust in Twaweza as an information source, with covariates

| | DV: Vote Choice for the Incumbent; Intention as Control | | | | | | | |
|---------------------------|---------------------------------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|
| | LC5 Chair | LC5 Councillor | LC3 Chair | LC3 Councillor | LC5 Chair | LC5 Councillor | LC3 Chair | LC3 Councillor |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | | | Good News | | | Bad News | | |
| Treatment | -0.003 (0.015) p = 0.865 | -0.031 (0.021) p = 0.131 | -0.018 (0.016) p = 0.262 | 0.023 (0.023) p = 0.312 | 0.010 (0.022) p = 0.657 | -0.006 (0.021) p = 0.760 | 0.023 (0.021) p = 0.276 | 0.027 (0.026) p = 0.301 |
| Treatment X Trust Twaweza | -0.021 (0.025) p = 0.416 | 0.015 (0.035) p = 0.665 | 0.032 (0.028) p = 0.249 | -0.062 (0.036) p = 0.085 | -0.009 (0.024) p = 0.730 | -0.010 (0.032) p = 0.753 | -0.033 (0.036) p = 0.361 | -0.058 (0.043) p = 0.185 |
| Village fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3,884 | 3,248 | 3,454 | 2,295 | 2,996 | 2,541 | 2,570 | 1,750 |
| Adjusted R ² | 0.289 | 0.346 | 0.370 | 0.294 | 0.265 | 0.365 | 0.335 | 0.241 |

Note: SEs clustered by politician and two-tailed p-values. Baseline category for moderator is any answer except trusting Twaweza for information a lot, including trusting a little, trusting not at all, and do not know responses.

9 Attribution for Public Service Outcomes

We argue in the main text that voters face significant difficulties in attributing public service outcomes to district (LC5) and subcounty (LC3) governments. In turn, this may have prevented the informational treatments from changing vote choices. However, two messages among the several sent to voters as part of treatment included information about responsibilities (see Table S2). Prior to the LC5 elections, voters in the roads condition were told that sub-county roads are managed by the LC3 and that primary education is the responsibility of the LC5. We were later corrected and learned that there are some shared responsibilities for primary education between the LC3 and LC5. Therefore, prior to the LC3 elections, only voters who selected to receive information about roads were informed that responsibility lay with a single level of government, in this case the LC3. In all other cases, treated subjects were informed about overlapping responsibility. In no case was information shared that attributed performance to a single chair or councillor office, leaving some ambiguity in attribution.

An observable implication of our interpretation of the main results is that treatment effects should be most likely to emerge among the subjects that selected into services with clearer attribution. To test for this possibility, we create a binary variable (Attribution) that is positive when the respondent expressed a preference for the service that had an attribution message linked to only one level of government. Since treatment effects might emerge when attribution is clearer, the interaction between treatment and the attribution indicator should be positive in the good news group and negative in the bad news group.

As displayed in Tables S12 and S13, we find no evidence that treatment effects are present in the subgroups who received messages attributing services to a single level of government. It is unlikely that two messages regarding responsibility for services were not enough to dislodge confusion about attribution among subjects. Furthermore, only sub-county roads are the responsibility of the LC3 government; other types of roads exist and are likely experienced regularly by subjects. Overall, the available evidence calls for more careful designs that can understand whether the combination of information about responsibility and programmatic performance affects vote choice.

Table S12: Heterogeneous effects of treatment by public service subgroup with attribution information, without covariates

| | DV: Vote Choice for the Incumbent; Intention as Control | | | | | | | |
|-------------------------|---------------------------------------------------------|--------------------------------|--------------------------------|--------------------------------|-------------------------------|--------------------------------|--------------------------------|-------------------------------|
| | Good News | | | | Bad News | | | |
| | LC5 Chair (1) | LC5 Councillor (2) | LC3 Chair (3) | LC3 Councillor (4) | LC5 Chair (5) | LC5 Councillor (6) | LC3 Chair (7) | LC3 Councillor (8) |
| Treatment | -0.011 (0.017) p = 0.508 | -0.016 (0.016) p = 0.293 | -0.013 (0.015) p = 0.384 | 0.004 (0.020) p = 0.825 | 0.011 (0.023) p = 0.635 | -0.001 (0.020) p = 0.964 | 0.025 (0.018) p = 0.170 | 0.003 (0.024) p = 0.897 |
| Treatment X Attribution | 0.003 (0.038) p = 0.932 | -0.037 (0.028) p = 0.187 | 0.032 (0.037) p = 0.384 | -0.018 (0.047) p = 0.704 | 0.004 (0.028) p = 0.899 | -0.024 (0.030) p = 0.415 | -0.049 (0.038) p = 0.197 | 0.007 (0.053) p = 0.892 |
| Village fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Covariates | No | No | No | No | No | No | No | No |
| Observations | 3,884 | 3,248 | 3,454 | 2,295 | 2,996 | 2,541 | 2,570 | 1,750 |
| Adjusted R ² | 0.282 | 0.342 | 0.369 | 0.288 | 0.253 | 0.363 | 0.331 | 0.240 |

Note: SEs clustered by politician and two-tailed *p*-values

Table S13: Heterogeneous effects of treatment by public service subgroup with attribution information, with covariates

| | DV: Vote Choice for the Incumbent; Intention as Control | | | | | | | |
|-------------------------|---------------------------------------------------------|--------------------------------|--------------------------------|--------------------------------|-------------------------------|--------------------------------|--------------------------------|-------------------------------|
| | LC5 Chair (1) | LC5 Councillor (2) | LC3 Chair (3) | LC3 Councillor (4) | LC5 Chair (5) | LC5 Councillor (6) | LC3 Chair (7) | LC3 Councillor (8) |
| Treatment | -0.010 (0.017) p = 0.555 | -0.015 (0.016) p = 0.354 | -0.013 (0.015) p = 0.402 | 0.003 (0.021) p = 0.899 | 0.004 (0.023) p = 0.864 | -0.001 (0.020) p = 0.951 | 0.024 (0.019) p = 0.194 | 0.007 (0.025) p = 0.774 |
| Treatment X Attribution | 0.001 (0.038) p = 0.990 | -0.036 (0.028) p = 0.200 | 0.031 (0.037) p = 0.396 | -0.014 (0.047) p = 0.764 | 0.008 (0.028) p = 0.785 | -0.025 (0.030) p = 0.402 | -0.053 (0.040) p = 0.183 | 0.001 (0.053) p = 0.981 |
| Village fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3,884 | 3,248 | 3,454 | 2,295 | 2,996 | 2,541 | 2,570 | 1,750 |
| Adjusted R ² | 0.289 | 0.346 | 0.370 | 0.293 | 0.265 | 0.365 | 0.336 | 0.242 |

Note: SEs clustered by politician and two-tailed *p*-values

10 Coding Audit Performance and Measuring Respondent Priors

10.1 Coding Audit Scores

As described in Section 4, we fielded audits of roads, health centres, primary education and water access in all sample villages.¹¹ We used this audit information to construct a village performance index for each public service. Where audits measured multiple dimensions of a public service, we sum the z-scores of the the indices for each dimension using the following formula:

$$AuditScore = \sum_{n=1}^N \frac{I_n - \mu_I}{\sigma_I} \gamma_n \quad (2)$$

Where I_n indicates a village's score in dimension n of an audit, μ_I indicates the district mean of I , σ_I indicates the standard deviation of I within the village's district and γ_n indicates the weight given to dimension n in the index.

To make this audit information comparable to respondents' priors, we calculate how each village's audit performance compares to all other villages in a respondent's district. If a village's audit is in the top quartile as compared to all audited villages in a district, its performance is coded as "much better." Similarly, villages in the third, second, and first quartiles are marked "better," "a little worse" and "much worse," respectively. We plot the distribution of the audit scores by subjects in Figure S14.

10.2 Measuring Respondent Priors

At baseline we asked respondents to select the public service that was most important to them when deciding how to vote for local officials. Based upon their response, we then asked "If you compare your [LC5/LC3] performance in managing the quality of [name of public service] in your villages to other villages in your district how do you think it will compare? (1) much better, (2) better, (3) a little worse, (4) much worse, (8) don't know, (9) refused to answer." Answers to this question were used to construct respondent priors. We plot the distribution of respondent priors in Fig S13. The distribution of respondent choices can be seen in Fig S16 in SI section 11.

As noted in the main text, some respondents were treated with information on their second-choice public service since audit information on their first choice was unavailable due to enumerators being unable to conduct certain kinds of audits in some villages (for instance, because there was no local health centre or county road). We only collected priors on respondents' first choice public service, so we have missing priors for 21.8 percent of respondents who were assigned to a different public service as part of a random step-down procedure. For the analysis in the main text, we impute these priors using the mean prior for all other respondents in the village for whom we did have data on priors, as outlined in our pre-analysis plan. The only other exception is Figure 2, Panels A-B, which only use known priors for the test of heterogeneous treatment effects.

We show the difference between priors and audit performance as compared to all of Uganda's districts in Figure S15, but for ease of presentation do not use imputed priors in this figure. In these and other calculations, priors are coded on a 1-to-5 scale based upon whether respondents believed their council's "record of managing its budget expenditures and contracting" was "Much Worse," "A Little Worse," "Don't Know," "A Little Better," or "Much Better" than others. Audit scores are coded on the same scale based upon whether the share of irregularities in the budget fell in the best, third, second, or first quartile as compared to the distribution of irregularities across all districts, omitting the "Don't Know" category in the middle of the scale. The distribution of the difference between priors and audits is centered on zero with only a slight skew, indicating reasonable convergence. However, the considerable variance ($\mu = 0.001$, $\sigma = 1.70$) indicates that most respondents had little ability to predict their council's performance. In total, 18% ranked performance consistently with the audits (excluding "don't know" cases), with most over-estimating the performance of poorly performing councils and under-estimating the performance of well-performing councils.

To create good- and bad-news subgroups from these data, we compared respondents' priors to rankings derived from the audits, as discussed in the main text. When respondents indicated "don't know," we included them in analysis but assumed that their priors are uninformative by placing them in the bad-news subgroup when their village did worse than the median and in the good-news subgroup when it did better.

¹¹As described in section 4, there were some villages in which audits could not be completed for some public services.

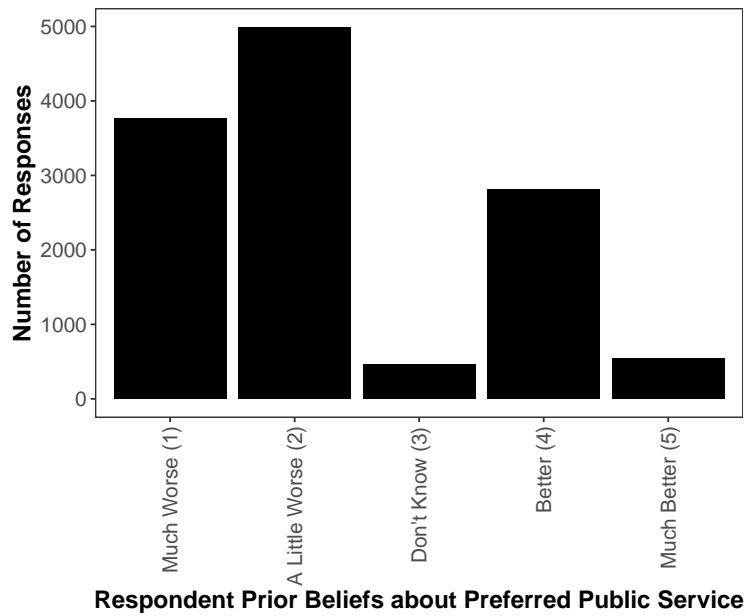


Figure S13: **Descriptive data for voters’ priors about preferred public services.** Figure shows distribution of responses to question, “If you compare your [LC5/LC3] performance in managing the quality of the [name of public service] in your villages to other villages in your district how do you think it will compare? (1) much better, (2) better, (3) a little worse, (4) much worse, (8) don’t know, (9) refused to answer” includes as part of a pre-treatment baseline survey. Respondents with imputed priors are excluded from this figure.

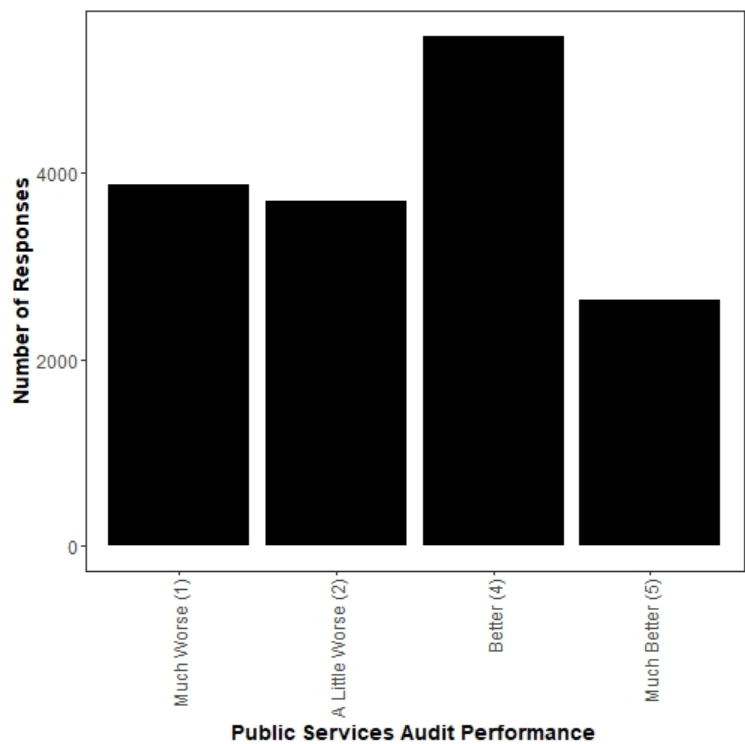


Figure S14: **Distribution of audit performance scores for respondents’ selected public service.** Councils are coded as being “much worse,” “a little worse,” “better” or “much better” based upon whether they are in the first, second, third or fourth quartile of their selected public service relative to other villages in their district.

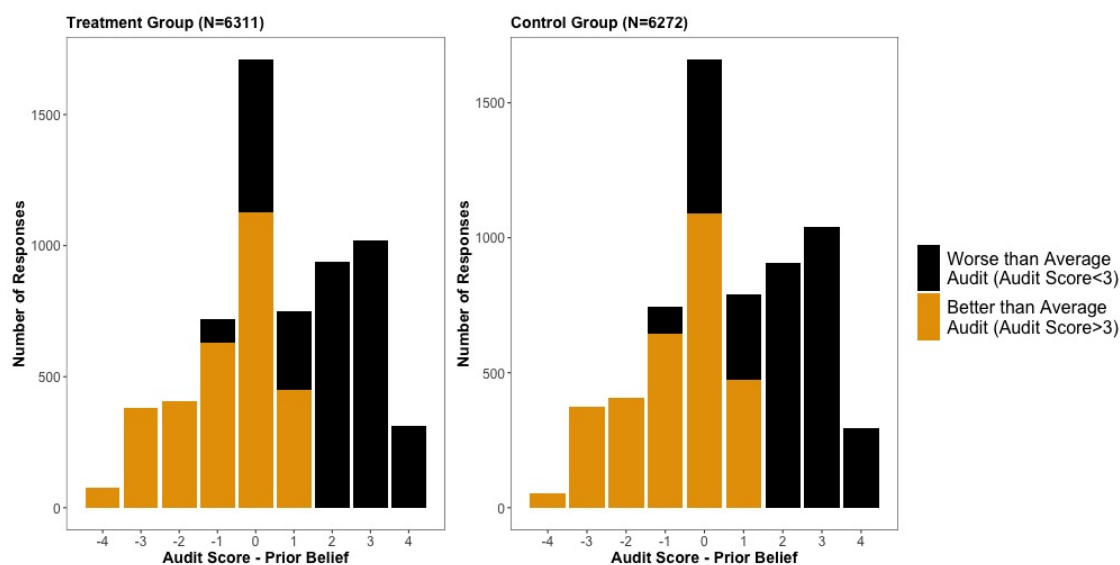


Figure S15: Difference between audits and respondents’ prior beliefs about public service quality. This figure shows the difference between respondent priors about their selected public service and the true audit scores. Priors are coded on 1 to 5 scale based upon whether respondents believed their selected public service was “Much Worse,” “A Little Worse,” “A Little Better,” or “Much Better” than others. We code “Don’t Know” as a 3. Audit scores are coded on a 1 to 5 scale based upon whether the quality of the respondents’ selected public service fell in the top, third, second or first quartile of villages in their district. This figure demonstrates that, while there is convergence between priors and audits on average, most respondents do not correctly rank their village’s performance. Dark bars indicate respondents who were in villages eligible to receive messages saying their village was “worse” than others. Light bars indicate respondents who were in villages eligible to receive messages saying their village was “better” than others. This visually illustrates the fact that most respondents over-rate villages doing “worse” than others and under-rate villages doing “better” than others. Respondents with imputed priors are excluded from this figure.

11 Balance and Attrition

11.1 Balance

As displayed in Figure S16, we have examined the data for evidence that randomization was successful. We see appropriate balance on all pre-treatment covariates, indicating successful random assignment. To assess balance formally, we conducted a joint F-test for balance to determine whether all of the covariates displayed in Figure S16 are able predict treatment status. We are unable to reject the null of no imbalance at $p=0.95$. Thus, our checks do not yield cause for concern about random assignment.

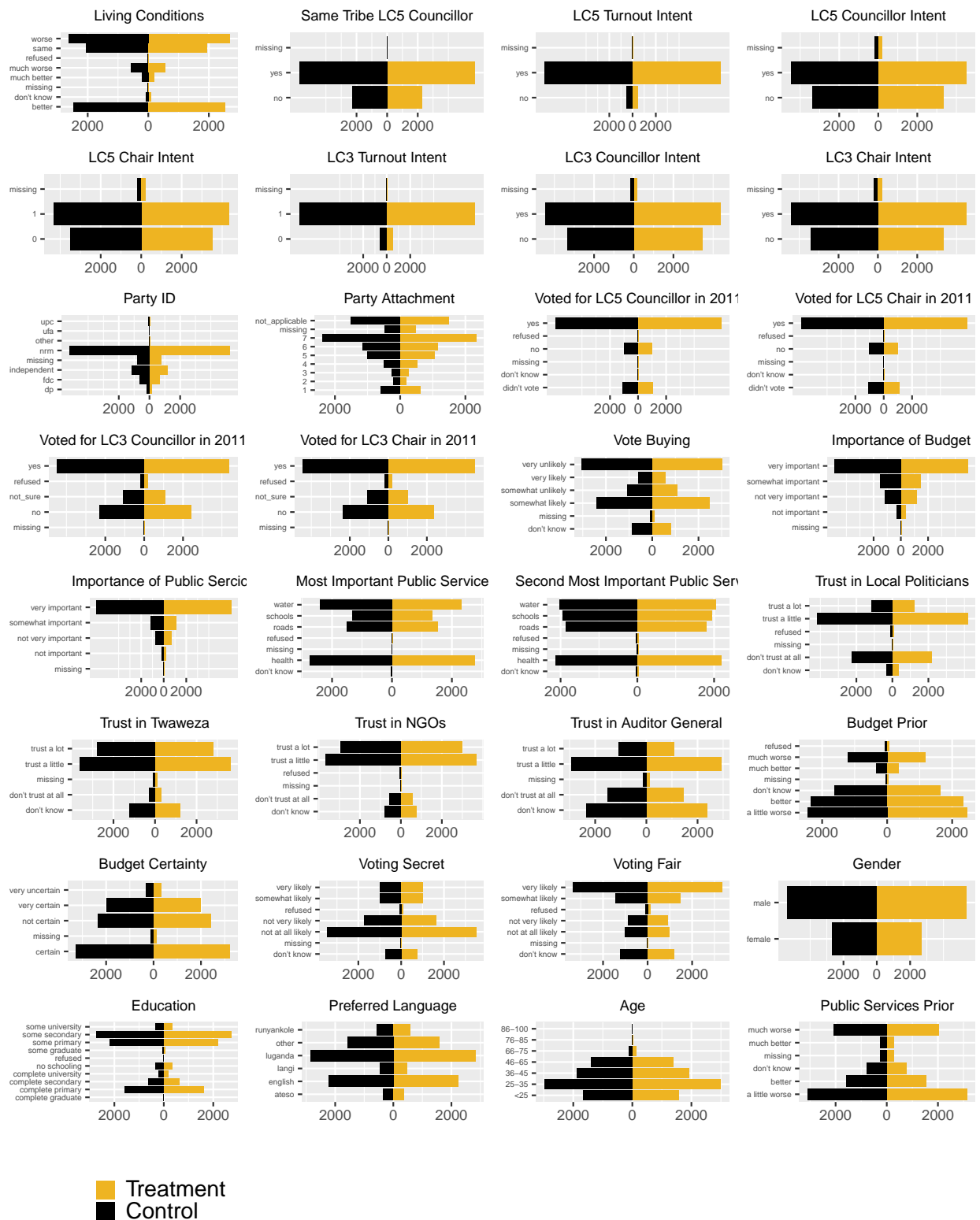


Figure S16: Balance of individual-level pre-treatment covariates at baseline. This figure shows the distribution of each pre-treatment variable by treatment assignment for all subjects contacted in the baseline survey. Orange indicates treatment-group subjects and black indicates control-group subjects.

11.2 Attrition

We used a Chi-Square test to determine whether there are indications of differential attrition by treatment condition for the public services arm. We do not find any indications of differential attrition by treatment status (LC5: $p=0.44$; LC3: $p=0.44$). We also specify two models of attrition where the regressors are either all the variables displayed in Figure S16 interacted with treatment or all of the same variables not interacted with treatment status. We then compare the fit of the models using an F-test to determine whether there is evidence of differential attrition within subgroups. We fail to reject the null hypothesis of differential attrition both for the LC5 endline ($p = 0.76$) and the LC3 endline ($p = 0.24$). Since we do not find evidence of differential attrition, we take no further steps to modify our analysis to account for differential attrition.

Table S14: Attrition with respect to the public services treatment

| | <i>No Attrition</i> | <i>Attrition</i> |
|-----------|---------------------|------------------|
| | LC5 Endline | |
| Control | 6307 (0.785) | 1729 (0.215) |
| Treatment | 6274 (0.780) | 1773 (0.220) |
| | LC3 Endline | |
| Control | 6036 (0.751) | 2000 (0.249) |
| Treatment | 6001 (0.746) | 2046 (0.254) |

Notes: This table shows the rates of attrition by treatment assignment (row-wise proportions are in parentheses). The lack of meaningful difference between groups suggests that treatment assignment did not differentially affect attrition.

12 Pre-Registration

Timing. We pre-registered at <http://egap.org> an initial for analysis on November 19, 2015 prior to any research activities and an updated plan for analysis on February 17, 2016 following the field-based recruitment drive and the baseline survey by call center, but prior to the random assignment of treatments. We also filed an addendum on April 21, 2016 after the collection of outcome data for an analysis about the conditional effects of treatment by exposure to election irregularities, prior to collecting data on election irregularities or completing analysis on the hypotheses listed. None of the analyses proposed in the addendum are considered in the present manuscript and are instead presented in a separate manuscript. These pre-registration files are available at <https://osf.io/t4qjx/>. Additionally, our team co-authored an earlier meta-analysis plan filed on March 9, 2015 for a larger initiative of field experiments on information and accountability in elections, which outlines general procedures for analysis and the scope of the overall initiative. Below we refer to the updated pre-analysis plan filed on February 17, 2016.

Scope of the present manuscript. This manuscript focuses on the impact of information about a public services audit. In addition to this public services treatment, the larger project included a separate and fully crossed treatment that provided information about budget irregularities in LC5 councils. Because the public services treatment crossed and does not interact with the budget treatment, the results of the budget treatment are published elsewhere.

It is also important to note that the public-services treatment had a different logic than the budget treatment reported here. Information about public services was intended to inform subjects about how very local services related to roads, education, health and water differed across Ugandan communities. In contrast, information about budget irregularities was intended to inform subjects about how district offices were spending their budgets. The attribution to public officials also differs across the treatment arms. Local public services are the responsibility of multiple layers of government, and are affected by many things extraneous to current officials' performance Grossman and Michelitch (2018).

The longer and more complex attribution chain involved in local public-service provision than in budget management may suggest that information on public services is likely to have less impact on vote choice for subcounty and/or district officials where competences are shared. Also different from the budget management information, where treated subjects all received information about their district officials' budget mismanagement, was the fact that treated subjects in the public-services arm were provided with information about the quality of the particular public service (primary education, water access, health clinics, or local roads) that they had identified as most important to them in the baseline survey.

Deviations from the pre-analysis plan and justifications. As noted in the SI section "Results using Pre-Registered Estimation Strategy" above and in the main text, the results in the main text do not strictly match our pre-registered analysis. Rather, the findings reported in the text reflect adjustments in light of unanticipated elements in the data. We nevertheless worked diligently to ensure that the analyses are as consistent as reasonably possible with pre-specification given the unforeseen data challenges and features of the setting unknown at the time we pre-registered.

For the main text, Figure 2 shows the results of our analysis of heterogeneous treatment effects derived by pooling across the offices, rather than individual by each office. Pooling was not part of our pre-specified analysis strategy and was used to show overall results in a compact format. The results disaggregated by office are displayed in SI Figures S9 - S12.

Ambiguities and errors in the pre-analysis plan. We note one editing omission on p. 8 of our updated pre-analysis plan where we enumerate the list of outcomes that we will consider related to vote choice. We state:

We will use two surveys to evaluate the effects of information on vote choice (see appendix). We will use a post-election survey conducted by a call center to evaluate the effects of information on vote choice, voting motivations, and voter perceptions. We will use these surveys to measure (1) votes for sub-county and district council chairpersons; (2) perceptions and knowledge of the performance of sub-county and district chairpersons and councillors, (3) vote buying and motivations for voting, (4) engagement with elected officials, and (5) voter turnout.

The first enumerated item in this list should read "votes for sub-county and district council chairpersons and councillors." We note that this sentence refers to both the budget and public-services treatment arms and is clearly inconsistent with every other part of the pre-analysis plan that lists an interest in understanding vote choice for both chair and councillor offices at the district (LC5) and sub-county (LC3) levels. We regret this editing error. Nevertheless, the initial sentence of the paragraph refers directly and expressly to the survey items in which vote choices for both

chairs and councillors are probed independently. The survey and prior references to vote choice for both chairs and councillors best represent the intent of the design and pre-specification.

We included in our pre-analysis plan's section on "power analysis" our intent to employ one-tailed hypothesis tests to estimate treatment effects in the calculation of statistical power for our experiment. We did so because our hypotheses are directional. That is, we pre-registered the hypotheses that bad news will decrease votes and good news will increase votes for the incumbent. To improve clarity, we should have also included in the pre-registered plan's "Estimation Strategy" our planned use of one-tailed hypothesis tests, though it would be reasonable to conclude that the estimation strategy would follow plans in the power analysis.

Our list of covariates to include in our specifications is ambiguous to which measure of trust in institutions would be included. The original intent of this measure is trust in the information source, which in our case was our partner organization Twaweza. Thus, this is the measure that we chose to include.

Unanticipated developments not included in the pre-analysis plan. We did not pre-register sub-setting the data with regard to contested elections, redistricting, or party switching. Because we considered it self-evident that we can analyze the impact of information on vote choice only if voters face an actual choice between candidates, we did not include in our pre-analysis plan that we would exclude constituencies in which the office was not contested. In cases of sub-setting involving the incumbent switching parties or involving re-districted constituencies, we did not become aware of their scope until well into the analytical stage of our study. We present in the main text treatment effects on subjects in areas where incumbency was the most unambiguous and therefore to our minds best matched pre-specified theoretical goals – excluding uncontested elections, elections with party-switching incumbents, and redistricted constituencies that might change a voters' incumbent.

13 Ethics

The study involved the supply of information to voters in anticipation of upcoming elections. Voters in Uganda, most of whom are economically disadvantaged by global standards, should be classified as relatively "low-power" subjects. Thus, recruiting them as participants requires special sensitivity to ethical concerns. Moreover, researcher involvement in political processes such as elections likewise requires consideration of potential harms to subjects and to the society more broadly.

In designing the study, the researchers considered each concern carefully and built in safeguards to obviate and mitigate possible harms. These design adjustments are in harmony with the American Political Science Association's "Principles and Guidance for Human Subjects Research" promulgated on 20 April 2020 (Council, 2020).

First, before participation in the study, all subjects offered individual fully informed consent to provide their phone numbers and receive SMS information about the upcoming elections. Enumerators in the field encountered enthusiasm for the offered political information, and the vast majority of people contacted consented to participate. Additionally, we gave respondents the opportunity to opt out of the treatment at any time by sending a "STOP" message via SMS. Researchers studying ethics have found that obtaining fully informed consent from subjects substantially and significantly ameliorates both potential subjects' and researchers' concerns about the study's ethics (Desposato, 2018).

Second, the information provided to subjects as part of the study was truthful. The information was drawn from official, credible government sources; from systematic audits performed by our partner organization, Twaweza; and from audits undertaken by our enumeration team using strict, consistent protocols. Thus, no deception was used in the study.

Third, before fielding the experiment, researchers worried that involvement in the political processes of voting and elections might possibly entail broader effects on the governance of locales where the study was conducted. Given this concern, subject recruitment and the information treatments were dispersed in such a way so that treatment density was deliberately kept very low in any given jurisdiction. Thus, if the information had meaningful effects on voting, those effects would be diffused within the local elections to the degree that they would have minimal probability of affecting the aggregate electoral outcome.

This project was approved by: the UCSB Human Subjects Committee (#15-0690); IRBs at BYU (#15381), William and Mary (2015-09-10-10589), and Temple (via IAA); the LSE Research Ethics Committee; the Uganda Mildmay Research Ethics Committee (0309-2015); the Uganda National Council for Science and Technology (SS 3943); and the Ugandan Office of the President (ADM 154/212/03).

References

- Afrobarometer. 2015. "Afrobarometer Data Round VI, Uganda." Available at <http://afrobarometer.org/countries/uganda-0> (Accessed March 2017).
- Bainomugisha, Arthur, L. Muyomba-Tamale, Wilson W. Muhwezi, Kiran Cunningham, Eugene G. Ssemakula, George Bogere, Russell Rhoads and Jonas Mbabazi. 2015. "Local Government Councils Scorecard Assessment 2014/15." ACODE Policy Research Series, No. 70, 2015. http://www.acode-u.org/Files/Publications/PRS_70.pdf (Accessed December 2016).
- Carlson, Elizabeth. 2015. "Ethnic Voting and Accountability in Africa: A Choice Experiment in Uganda." *World Politics* 67(2):353–385.
- Council, American Political Science Association. 2020. "Principles and Guidance for Human Subjects Research." URL: <https://connect.apsanet.org/hsr/principles-and-guidance/>
- Desposato, Scott. 2018. "Subjects and scholars' views on the ethics of political science field experiments." *Perspectives on Politics* 16(3):739–750.

- Electoral Commission. 2016. The Republic of Uganda 2015/2016 General Elections Report. Technical report Electoral Commission.
- Gandhi, Jennifer and Ellen Lust-Okar. 2009. "Elections under authoritarianism." *Annual review of political science* 12:403–422.
- Gerber, Alan S, James G Gimpel, Donald P Green and Daron R Shaw. 2011. "How large and long-lasting are the persuasive effects of televised campaign ads? Results from a randomized field experiment." *American Political Science Review* 105(1):135–150.
- Green, Elliott. 2015. "Decentralization and Development in Contemporary Uganda." *Regional & Federal Studies* 25(5):491–508.
- Grossman, Guy and Kristin Michelitch. 2018. "Information dissemination, competitive pressure, and politician performance between elections: A field experiment in Uganda." *American Political Science Review* 112(2):280–301.
- Humphreys, Macartan and Jeremy M. Weinstein. 2013. "Policing Politicians: Citizen Empowerment and Political Accountability in Uganda." Unpublished Manuscript.
- Manyak, Terrell G and Isaac Wasswa Katono. 2011. "Impact of multiparty politics on local government in Uganda." *African Conflict and Peace Building Review* 1(1):8–38.
- Martin, Lucy and Pia J Raffler. 2021a. "Fault Lines: The Effects of Bureaucratic Power on Electoral Accountability." *American Journal of Political Science* 65(1):210–224.
- Martin, Lucy and Pia J Raffler. 2021b. "Fault Lines: The Effects of Bureaucratic Power on Electoral Accountability." *American Journal of Political Science* 65(1):210–224.
- Natamba, Edward F., Lillian Muyomba-Tamale, Eugene Ssemakula, Enock Nimpammya and Immaculate Asiimirwe. 2010. "Local Government Councils Performance and the Quality of Service Delivery in Uganda: Ntungamo District Council Score-Card 2008/9." Unpublished Manuscript.
- of Uganda, Parliament. 1997. "Republic of Uganda: Local Governments Act." <https://ulii.org/ug/legislation/consolidated-act/243>.
- Office of the Auditor General. 2014. "Annual Report of the Auditor General for the Year Ended 30th June 2014 Local Authorities." Volume 3. The Republic of Uganda.
- Raffler, Pia. 2018. Does Political Oversight of the Bureaucracy Increase Accountability? Field Experimental Evidence from an Electoral Autocracy. Technical report Working paper, Yale University.
- Tripp, Aili Mari. 2010. *Museveni's Uganda: paradoxes of power in a hybrid regime*. Vol. 6 Lynne Rienner Publishers Boulder, CO.
- Tumushabe, Godber, Elijah Dickens Mushemeza, Lillian Muyomba Tamale, Daniel Lukwago and Eugene Ssemakula. 2010. "Monitoring and Assessing the Performance of Local Government District Councils in Uganda: Background, Methodology and Score Card." ACODE Policy Research Series, No. 31. http://www.acode-u.org/Files/Publications/PRS_31.pdf (last accessed July 30, 2020).
- Twaweza. 2017. Looking for Data to Publish: Use Data from Africa's Largest Learning Assessment. Technical report Twaweza.
URL: <https://twaweza.org/go/uwezo-data-booklet>
- Van Breukelen, Gerard JP. 2006. "ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies." *Journal of clinical epidemiology* 59(9):920–925.