

# SMS texts on corruption help Ugandan voters hold elected councillors accountable at the polls

Buntaine, Jablonski, Nielson & Pickering 10.1073/pnas.1722306115

## Supporting Information (SI)

**Example Treatment Messages.** A list of example treatment and placebo messages sent to subjects is in Table S1. Illustrations of how these messages appeared on phones are in Figure S1.

**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.\* 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.

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).† 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, 12,566 people were reached at endline and were therefore included in the subsequent analysis that we present in this report. Figure S3 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.

**Validating Measures of Vote Choice.** As discussed in the main text, 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 S4. The almost perfectly linear relationship suggests there is limited bias in reported voting in favor of the announced winner of the election (chairs:  $\rho = 0.77$ ; councillors:  $\rho = 0.80$ ).

Second, we compare reported results within our sample to the official district and sub-county vote counts in Figure S5. Survey-reported voting for the incumbent party is strongly correlated with official results data,‡ although people in low-vote areas may have been somewhat more likely to report voting for the incumbent party than the official turnout data would suggest.

Third, we 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 S6, our data appear consistent with accurate reporting: 93% of respondents voting in the district council election recalled a basin or basin color in their polling station, and among these 77% were able to accurately name the color in an internally consistent manner.§

**Potential Biases in Reported Vote Choice.** As discussed in the main text, a further concern about our measures of vote choice is social desirability bias. If survey respondents believe that our enumerators or NGO partner, Twaweza, desires a particular response, then this belief might induce respondents to systematically over-report voting for well-performing incumbents and under-report voting for poorly performing incumbents. If this pattern were true, we might expect stronger treatment effects among respondents that trust Twaweza. As shown in Table S2, this dynamic does not appear to be occurring. Instead, there is evidence, among the good news group, that some treatments are stronger among those respondents who reported trusting the Auditor General (AG), which was the original source of budget information and an organization with which respondents had

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

† The 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.

‡ For chairs,  $\rho = 0.50$  ( $p = 0.09$ ). For councillors,  $\rho = 0.61$  ( $p < 0.001$ ).

§ The 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 (71%), water basins could be blue, brown, green, grey, orange, purple, red, white or yellow.

no social interaction. In contrast to what we would expect if there was severe social desirability bias, this finding suggests respondents were weighing the credibility of evidence.

Additionally, in Figure S6, we show that the ability of respondents to correctly name the color of the water basin in their polling booth is not associated with assignment to treatment, control, good news, or bad news. If we assume (as suggested by the literature on vote biases) that social desirability biases are especially likely to result in higher reported turnout (1), the fact any over-reporting of turnout is not related to treatment provides evidence that our treatments were not creating significant social desirability pressure with respect to vote choice. Additionally, we note that social desirability and strategic misrepresentation should have applied equally to chair and councillor elections. That we did not find significant evidence that our treatment affected reported voting for the chair elections suggests that response bias cannot explain our finding that the treatment significantly affected reported voting for councillors.

Another potential source of bias in reported vote choice could result from voters' fearing repercussions for responding in certain ways. To explore this possibility, we examine the effect of the information treatment among voters who expected a free and fair election at baseline. In our study, these voters reported that they considered it "highly unlikely" or "unlikely" that powerful people would learn how they vote and that they considered it "highly likely" or "likely" that vote counting would be fair. Because this subset of voters do not seem to perceive intimidation or the potential for tampering, they should have been less likely to alter their reported vote choice in response to a perceived threat of repercussions for answering our endline surveys in certain ways. As we display in Table S3, the main results reporting the information treatment holds in changing reported vote choice for councillor elections in the expected directions.

**Measuring Respondent Priors and Council Budget Performance.** To measure respondents' priors, we asked the following question at baseline: "If you compare your LC5's record of managing its budget expenditures and contracting to other districts in Uganda, how do you think it will compare? (a) much better, (b) better, (c) a little worse, (d) much worse, (e) don't know, (f) refused to answer." We plot the distribution of these priors in Figure S7.

We measure budget performance using official Ugandan Auditor General reports about the amount of irregular expenses as a percentage of the official council budget. Formally, irregular expenditures include those with unaccounted for funds (e.g., missing documentation); procurement issues (e.g., inflated or missing contracts); or payroll anomalies (e.g., "ghost" employees or overpayment of salaries). To make this performance information comparable to respondents' priors, we calculate how each district council's audit performance compares to all other district councils in Uganda. If a council's percent of irregular expenses is in the top quartile as compared to all of Uganda's districts, its performance is coded as "much worse." Similarly, councils in the third, second, and first quartiles are marked "a little worse," "better" and "much better," respectively.

We plot the distribution of the audit scores by subjects in Figure S8 and show the difference between priors and audit performance as compared to all of Uganda's districts in Figure S9. 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 ( $\mu = 0.32$ ,  $\sigma = 1.90$ ) indicates that most respondents had little ability to predict their council's performance. In total, 25% 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 Auditor General's data, 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 council did worse than the median and in the good-news subgroup when it did better. Our results do not appear sensitive to this coding decision. In Table S4 we show results conditional on whether respondents indicated "don't know" in their prior. There is no interaction between having unsure priors on budget performance and the treatment inconsistent with a homogeneous main effect. In Models 7/8, the treatment effect of bad news disappears in absolute value, however, because of the small number of subjects with unsure priors; it is not possible to clearly distinguish a difference caused by subgroup effects as compared to a difference caused by random variation.

When we examine the effects of the information treatment in groups not defined by respondents' priors, which was not the pre-registered analysis strategy, we find that only the effect of negative news is highly improbable assuming the null hypothesis and that the effect of positive news becomes less inconsistent with the null hypothesis (see Table S5). Thus, defining news groups by respondents' priors gives us some additional predictive power when considering good news based on priors versus positive news generally.

**Attrition and Balance.** In Figure S10, we display the counts of subjects assigned to treatment and control with different pre-treatment values of covariates. The figure confirms balance visually and provides an overview of the characteristics of our sample. We conducted a joint F-test for balance to determine whether all of the covariates are able predict treatment status, and we are unable to reject the null of no imbalance at  $p=0.16$ . Thus, our checks do not yield any cause for concern about random assignment. In Table S6 we show the means of all our pre-treatment covariates within treatment and control subsets, ordered by the severity of imbalance. In Table S7 we show the distribution of voters and incumbent party identification by

news grouping.

As outlined in the Metaketa I meta-pre-analysis plan, we conducted two tests for differential attrition between the baseline and endline surveys by treatment group. We first consider attrition rates with respect to the budget treatment, which is displayed in Table S8, with proportions computed by row. We fail to reject the null of no differential attrition with respect to the budget treatment (Chi-squared test,  $p = 0.18$ ). When we create a model of attrition with regressors as the interactions between the treatment status and the pre-specified covariates, we again find no indication of differential attrition with respect to treatment status and other observable subgroups (F-test,  $p = 0.74$ ). Since we do not find evidence of differential attrition, and as outlined in the Metaketa I meta-pre-analysis plan, we take no further steps to modify our analysis about the budget treatment to account for differential attrition. Figure S11 shows the balance between the treatment and control groups at the endline survey and characterizes our effective sample at the analysis stage.

We also consider differential attrition by the crossed budget treatment and density treatment indicators in the same way. Table S9 displays the count and proportion of subjects in different attrition categories with respect to the crossed treatment indicator. The results are not inconsistent with the null hypothesis of no differential attrition (Chi-squared test,  $p = 0.19$ ). When we model attrition with respect to treatment by covariate interactions, we fail to identify attrition by treatment status that differentially affects certain kinds of subjects (F-test,  $p = 0.95$ ). Since we fail to find an effect of density in the main specification and because estimating bounds of the treatment effect will only increase uncertainty, we do not take any further steps to analyze differential attrition as applied to the density treatment.

**Turnout and Estimates of Vote Choice.** One concern about our analytical strategy is that we estimate the effect of the informational treatments within the subset of voters who reported that they turned out, since we do not have vote choice data for subjects who did not turn out. If the treatment affected the composition of voters who reported voting, our estimates may be picking up a treatment effect on reported turnout, rather than a treatment effect on reported vote choice.

We address this ambiguity in three ways. First, we show in Table S10 that we see no main effect of the treatment on reported turnout, which confirms we are unlikely to have a different subset of voters in the treatment group as compared to the control group (e.g., with higher motivation) at the vote-choice stage. Second, when we conduct an F-Test that compares the baseline model of reported turnout predicted by treatment and covariates to a higher-order model that predicts reported turnout by a full set of treatment by covariate interactions, we see no increase in model fit that is distinguishable from adding random predictors (Good News:  $F = 0.984$ ,  $Df = 37$ ,  $p = 0.496$ ; Bad News:  $F = 0.898$ ,  $Df = 35$ ,  $p = 0.642$ ). This indicates that we do not have evidence that the treatment is affecting the kinds of eligible voters that are among the voting subset used to estimate treatment effects on reported vote choice. Third, since reported vote choice for the incumbent or incumbent party is our dependent variable, we estimate the effect of treatment on a reformulation of this outcome that codes non-turnout as a non-positive response category for vote for the incumbent. This formulation ensures that we have equivalent potential outcomes between the treatment and the control groups, since it includes all subjects randomized into treatment. As displayed in Table S11, we continue to recover the treatment effects reported in the main text under this formulation. Combined with the analysis showing no main effect of treatment on reported turnout, these analyses offer stronger confirmation that our treatment is operating by changing people's choices about who to vote into office, rather than by differentially mobilizing voters with different preferences for politicians.

**Vote Choice by Party.** One might be concerned that our results are idiosyncratic to particular parties, or biased in favor of one party or the other. For instance, if only NRM voters are eligible for good news and only opposition voters are eligible for bad news, one might be skeptical about generalizing conclusions about the implications for political accountability and democracy. We show the distribution of voter and incumbent types by news groups in Table S7, which rules out many of the concerns that results for our split good and bad news treatment will be affected strongly by the party identification of either the incumbent or of voters.

We further probe how party identification of voters and alignment with incumbents predicts how voters will respond to treatment. Table S12 shows how voters of different parties responded to budget treatment, with the NRM-aligned (ruling party) voters as the baseline condition. NRM-aligned voters respond to both good and bad news consistent with the average treatment effects reported in the main text. Independent and opposition voters do not report increasing their votes for the incumbent in response to good news on average, however. In terms of bad news, it appears that Independent voters punish incumbents like NRM voters, but that opposition party voters do not on average punish bad news by decreasing reported votes for the incumbent. Thus, the main effects appear to be primarily driven by NRM voters, not considering party alignment with incumbents. Opposition voters do not appear to be swayed in their vote choices by the information treatment.

In Table S13 we show how alignment in party identification with the incumbent politician of interest predicts how voters will respond to the budget treatment. Here the baseline condition is a voter who is not aligned with the incumbent. Interestingly, good news does not change the reported vote choice on average among unaligned voters, but does have an effect among aligned voters for councillor elections. We see marginal evidence that this pattern works in reverse for bad news, with unaligned voters responding most negatively to the treatment and the treatment effect diminishing for aligned voters. Overall, this supports that idea that voters respond to information that is consistent with their party alignments for councillor elections. This result does not hold for chair elections.

**Full Tables of Main Results Presented in Figure 1.** Table S14 and Table S15 report the full set of results of the main estimating equation using all treatment indicators and covariates as reported in Figure 1 in the main text. The only pre-treatment

covariate consistently associated with vote for the incumbent is gender; males appear to be less likely to vote for the incumbent in all specifications.

**Results using Pre-Registered Estimation Strategy.** For ease of interpretation and to correct for small imbalances in our realized randomization, in the main draft we deviated slightly from the pre-registered estimation strategy by not transforming the dependent variable. We prefer the specification in the main text because pre-treatment intent to vote for the incumbent is not perfectly balanced across treatment conditions due to the realized randomization draw (for example, the pre-treatment intent to vote for the current LC5 chairperson is imbalanced by 2%; see Table S6). Using the pre-registered strategy, this imbalance is incorporated into the differenced outcome and modeled as a function of treatment, rather than corrected using covariate adjustment as in our preferred specification in the main text. Since intention was imperfectly associated with reported vote choice, this is an inefficient specification. Since this small imbalance cannot be a function of treatment, but is rather a result of our realized random draw, in retrospect we consider it ill-advised to model it as a function of treatment. Instead, in the main specification the pre-treatment intention to vote for the incumbent in question is used as a covariate to adjust for imbalance in all specifications, including those specifications that do not use additional covariates. Recall that we did not find any overall problems with execution of the random assignment process per the randomization check in Section “Attrition and Balance.”

Nonetheless, and for the sake of transparency, Table S16 contains the results for vote choice using the pre-specified version of estimating Equation 1, which directly transforms the vote choice for the incumbent with the pre-treatment intention to vote for the incumbent, both with and without covariates as described above:

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

We cannot reject the null that the good news treatment did not affect reported vote choice for district chairs (Models 1-2) or that bad news did not negatively affect reported vote choice for district chairs (Models 5-6). Regarding district councillors, we find that good news boosted reported votes for the incumbent by approximately 3%, though this result is not highly inconsistent with the null hypothesis without covariates (Models 3-4). This is the only difference with the main specification reported in the main text and results from the small imbalance in intention being directly incorporated in the outcome when using a transformation. When we include covariates, reported vote choice in the previous election is able to account for much of the imbalance introduced by transforming the present vote choice outcome by a measure of pre-treatment intent. Consistent with the main results reported above, we also find that bad news decreased votes for the incumbent district councillor by a bit more than 3% (Models 7-8).

For our hypothesis on the density of treatment, we again pre-specified an estimating equation that uses an outcome vote-choice variable transformed by pre-treatment intention to vote for the incumbent, as follows:

$$y_{ij,t=1} - y_{ij,t=0} = \alpha + \tau_1 T_{ik}^+ + \tau_2 D_j + \tau_3 T_{ik}^+ D_j + \beta \mathbf{Z}_i + b_j + \epsilon_{jh} \quad [2]$$

Table S17 with the pre-specified estimating equation does not reveal any different results from those reported in the main text.

**Results on Vote Choice with Subsets that Define Incumbency in Different Ways.** We did not pre-register the exclusion of certain constituencies because we did not anticipate the scale of incumbent individuals who would switch parties, the presence of re-districting occurring between 2011-2016, or the number of uncontested elections (for further discussion, see this SI’s “Pre-registration” section). Table S18 shows the results of the budget treatment on voter choice using all of the data that we have available, dropping only uncontested elections, and with an interaction term for treatment with party switching. We see that the main results we report hold for elections without party switching. However, for the terms interacting budget treatment with candidates who switched parties, the signs flip and show magnitudes inconsistent with no interaction effect. This potentially indicates that voters primarily rewarded incumbent parties and not individual incumbent politicians who switched parties for good news and punished incumbent parties and not individual incumbent party switchers for bad news.

Party switching mainly occurred when incumbents lost primary elections, raising the question of which is the incumbent: the candidate or the party? Because of the theoretical ambiguity introduced by party switching and redistricting, and given that the goal of this project is to understand vote choice for incumbents, we present as our main results contested elections without party switching or redistricting. We made this decision after examining the data. The results reported in the main text are for the largest subset where ambiguity in incumbency is avoided and the subset that offers a test with the highest fidelity to the theoretical goals of the project. Nonetheless, we acknowledge that we did not anticipate the scale of party switching, redistricting, or uncontested elections in the design of the study and thus did not pre-register data sub-setting.

Therefore, to test the robustness of our findings to ambiguities in the definition of incumbency realized after we collected data, we re-estimate our results using three alternative codings of incumbency. First, we consider whether we can detect a treatment effect among the subset of constituencies where the individual incumbent is running again as member of the same party, which is arguably the cleanest test of our theory. However, our sample size and thus our statistical power is much reduced in this subset. Table S19 shows the results for the vote choice analysis estimated by the main estimating equation (1), with intention to vote for the incumbent used as a predictor variable. Because the sample size is much reduced for this specification, the estimates are less precise, though the estimated effect sizes are similar to the main results.

Next, we consider situations where individuals switched parties and ran for re-election. Many times these situations arose from incumbents losing in the primary elections. In these cases, it is not immediately clear whether information effects should



affect the incumbent party or the incumbent individual. Here we consider the implications of defining incumbency as the party previously elected to the seat, regardless of the individual running and the individual politician as the incumbent, and regardless of the party affiliation they ran under for the district elections. Observations not affected by party switching remain part of the sample for these analyses.

When we define incumbency as the party previously elected to the seat that was involved with party switching (see Table S20) the results are consistent with the main text, and *p*-values are smaller. When we define incumbency as held by the individual when he or she switches parties (see Table S21), results for the good news group decrease in their inconsistency with the null hypothesis and the results for bad news group attenuate slightly. Note that this is the same subset as used for Table S18 without the party switching interaction term included and includes all the available data from contested elections. We report these results for the sake of transparency, since we did not have clear expectations about how information would affect vote choice in situations where incumbency is difficult to define.

### **Pre-registration.**

**Timing.** We pre-registered an [initial plan](#) 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 plan](#) for analysis 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. Additionally, our team co-authored an earlier [meta-pre-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.** Our larger project included three treatment arms: (1) information on budget irregularities; (2) density of information on budget irregularities; and (3) information on the quality of public services. We collected data on vote choice and other attitudinal outcomes for two types of elections: the district (LC5) elections considered in the present manuscript and the subcounty (LC3) elections held several weeks later. All three treatment arms were deployed to subjects prior to the LC5 elections, with the public services treatment assigned randomly and fully independent of the budget irregularities and density treatment arms. Because the Auditor General did not conduct audits on all LC3s, we did not attempt to provide data to treated respondents on LC3s' budget management. As we note in the main manuscript, because the public services treatment crossed and does not interact with the budget treatment and because it applies to different elections for offices at the sub-county level that are additional to the ones considered in this manuscript, we are preparing a separate manuscript to report its results.

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 (2). The audit information, in contrast, provides information about the level of corruption and mismanagement among respondents' LC5 councils specifically.

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, such as in the case of primary education. 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.

For the arm featuring subjects' salient public service, we fail to reject the null hypothesis for vote choice, turnout or evaluations of candidate quality. The result is consistent across both offices at both the subcounty and district levels (3). While much research remains to be done, we suspect that voters may have been confused about which politician at which level of government should be most responsible for the quality of local public services. We plan to explore these possibilities in a separate manuscript. We note, however, that the public-services arm was substantively distinct from the budget arm and targeted different officials at a different level of government in different elections held weeks apart. Distinct results might naturally follow from these many differences.

**Hypotheses considered in the present manuscript.** In the present manuscript we consider only the following pre-registered hypotheses from our updated pre-analysis plan: H1 (vote choice) for the budget treatment; H7 (turnout) for the budget treatment; and H11 (density on vote choice) for the budget treatment. Note that these are the primary, self-reported *behavioral* outcomes that we measured and are the primary inferential targets of our project and the broader Metaketa initiative. Given these analytical priorities and the limited space available in this manuscript, we have chosen to focus on the main behavioral effect of the treatment. We do not report here pre-registered hypotheses on a series of mediators that might amplify the treatment and moderators that might produce heterogeneous treatment effects.

We report these pre-registered hypotheses on heterogeneous effects and intermediate outcomes for the budget treatment elsewhere (3). In particular, we report results from H2 (conditional effect of treatment on vote choice based on distance of the treatment information from priors), H3 (attitudes about candidate integrity), H4 (attitudes about candidate effort), H5 (conditional effect of treatment on vote choice based on uncertainty in priors), H6 (conditional effect of treatment on vote choice based on self-reported importance attached to information type), H8 (conditional effect of treatment on turnout based on alignment with the incumbent party), H9 (conditional effect of treatment on vote choice based on receipt of clientelistic benefits), and H10 (conditional effect of treatment on vote choice based on exposure to vote buying). We fail to find evidence that any of the potential mediators contributed to greater effects of the information treatment. We also fail to find evidence that any moderators affected vote choice.

**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.

**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.

There are also additional components of our pre-analysis plan that might appear ambiguous. We included in our pre-analysis plan’s endline survey a question that asked subjects to identify the color of the water basin in their polling station for the LC5 election. We did so with the intention of using subjects’ answers to this question to detect and correct for possible misreporting of voting. However, we did not explicitly mention our intent to use answers to this question to probe misreporting in our pre-analysis plan’s section on “covariate and outcome measurement,” though such an intent might be inferred.

Our pre-registered endline survey also included a question about whether respondents recalled seeing messages about public services, which was a separate (and fully crossed) treatment arm about water, roads, schools, and health centers. We intended to use responses to this question as a measure, albeit imperfect, of compliance suggesting receipt of information generally, including in the budget arm. However, we did not explicitly mention this intent in our pre-analysis plan. We refer to treated subjects who responded affirmatively to this survey question as “verified respondents” in the main text and label it as an “augmented analysis.”

Third, 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. Due to the directional nature of our hypotheses, we report in the main text’s Fig. 1 and Tables S14 and S15 one-tailed hypothesis tests, though we also report 95 percent confidence intervals in Fig. 1. Because we report *p*-values rather than categorical decisions at arbitrary thresholds regarding rejecting or failing to reject the null hypothesis, the reader interested in a non-directional hypothesis that we did not pre-register (e.g. good news affects vote choice) can easily double the reported *p*-values to adjust assessments of statistical significance.

Finally, 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, as outlined in the Metaketa I meta-pre-analysis plan, 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. An alternative measure that might be used is trust in the Auditor General, since the budget information used in the treatment is derived from audits conducted by this office. Adding or swapping this covariate for estimation has no effect on the magnitude or precision of the estimated treatment effects in the main text.

**Unanticipated developments not included in the pre-analysis plan.** As mentioned in the SI section on “Results on Vote Choice with Subsets that Define Incumbency in Different Ways,” 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 re-districted constituencies.

To probe the robustness of our findings, we re-estimate our results using four alternative codings of incumbency in this SI's section "Results on Vote Choice with Subsets that Define Incumbency in Different Ways." First, we consider incumbency in the most inclusive way that includes the party or the candidate that ran again in contested elections, even if the district was re-drawn. Second, we consider whether our treatment affects vote choice among the subset of constituencies where the individual incumbent is running again as member of the same party. Third, we consider constituencies where individuals switched parties and ran for re-election, which often happened as a result of the incumbent politician losing his or her party's primary. In these cases it is not clear whether information effects should affect the incumbent party or the incumbent individual. Here we consider the implications of defining incumbency as the party previously elected to the seat, regardless of the individual running. Fourth, we consider the individual politician as the incumbent, regardless of party affiliation in the district elections. As noted above, effects across the first three specifications are generally consistent with findings reported in the main text; effects are smaller in the fourth specification. Inconsistency with the null hypothesis varies across specifications but generally align with the findings reported in the main text.

Redistricting created multiple challenges for our analysis, particularly for the concept of incumbency. We sampled some re-districted constituencies because they already constituted part of the sample drawn by our NGO partner, Twaweza. We used our local partner's sample for several reasons. First, we used it in order to provide Twaweza with the possibility of using the same SMS service in the future to disseminate to citizens the information that they collect on primary schooling outcomes. In addition, using Twaweza's sample allowed us to disseminate the information they had already gathered on literacy and numeracy levels as part of our public services arm. However, citizens residing in districts that had been re-drawn between 2011 and 2016 complicated attribution. Should these citizens hold responsible for budget management and public services outcomes those officials defined prior to or after re-drawing of the district? Because there is no obvious answer to this question, it seems logical to exclude from analysis constituencies that were re-districted.

Only after receiving requests for more information about subjects' experiences with receiving SMS texts, we conducted a follow up survey with 100 randomly selected subjects who received the budget management treatment. Because of the timing of this effort, we were not able to include it in the pre-analysis plan. Key results of this survey are reported in penultimate paragraph of the main text's "Experimental treatment" section.

Our pre-analysis plan describes a plan to examine spillover between polling stations, based on the contiguity of polling stations. Despite significant effort, we have not been able to gain access to maps of polling stations and are not able to complete this pre-registered analysis.

**Simulation of a Hypothetical Scaled Treatment Intervention.** In this section we consider how a hypothetical scaled intervention might have impacted electoral outcomes. That is, if all eligible registered voters in Uganda were treated, how many councillor elections could be substantively altered, which we refer to here as "flippable"? We can represent this problem as follows:

$$v_j = S_j(\tau_j^+ p_j^+ - \tau_j^- p_j^-) \quad [3]$$

Where  $v_j$  represents the net intervention effect on incumbent  $j$ 's margin of victory.  $S_j$  represents the share of voters treated in  $j$ 's constituency.  $\tau_j^+$  represents the effect of good news on the probability of voting for  $j$ .  $\tau_j^-$  represents the effect of bad news on the probability of voting against  $j$ .  $p_j^+$  and  $p_j^-$  represent the proportion of voters in  $j$ 's constituency receiving good and bad news given their priors.

Unfortunately we lack the ability to estimate these parameters precisely for each election since we lack out-of-sample data on priors and voter characteristics; however, we can use our within-sample estimates to at least derive an out-of-sample average intervention effect for good and bad news areas. To estimate  $\tau^+$  and  $\tau^-$  we use the primary specification estimates from Table S14. This sample average treatment effect (SATE) is not an entirely unproblematic estimate of the population average treatment effect (PATE) since, while our experimental sample of villages is a random draw from the population, our within village recruitment relied on non-random self selection. While this self-selected sample is similar in many measurable characteristics to the population as a whole (as compared to representative Afrobarometer surveys, discussed in the main text), we cannot rule out the possibility that unobserved self-selection will result in an over- or under- estimate of the population average effect.

To estimate the share of voters eligible for good and bad news,  $p_j^+$  and  $p_j^-$ , we first calculate the within-sample probability a respondent receives good and bad news conditional on being in a district that performed better or worse than the median district in the audit. This gives us the following estimates:  $p_j^+ = 0.97$  if  $Q_j \geq \widehat{Q}$ ,  $p_j^+ = 0.025$  if  $Q_j < \widehat{Q}$ ,  $p_j^- = 0.91$  if  $Q_j < \widehat{Q}$ , and  $p_j^- = 0.09$  if  $Q_j \geq \widehat{Q}$ . Where  $Q_j$  equals the audit performance in  $j$ 's district and  $\widehat{Q}$  equals the median performance across all districts. We use these conditional probabilities to estimate  $p_j^+$  and  $p_j^-$  for each out of sample election.

The primary constraint on the share of voters treated in a scaled intervention,  $S_j$ , is likely to be access to cell phones. According to a 2015 Pew survey, approximately 65% of Ugandans have access to cell phones. Absent more fine-grained data, we use this as an estimate for each  $S_j$ .

Plugging these numbers into equation 3, and assuming a PATE range of  $\pm$  one standard deviation from the SATE, we estimate that a scaled intervention has the potential to increase vote margins on average by 0.6 to 2.7% when most voters are receiving good news (when  $Q_j \geq \hat{Q}$ ) and to decrease vote margins by 1.2 to 3.2% when most voters are receiving bad news (when  $Q_j < \hat{Q}$ ).

Using these estimates of  $v_j$ , we can also calculate the number of elections that could have flipped in 2016 under this hypothetical intervention. To estimate flipped elections, we first have to assume homogeneity in intervention effects across elections since we do not have a way to estimate election-specific out-of-sample intervention effects. This may be unreasonable, especially if treatment effects are stronger or weaker in marginal elections or if cell-phone access differs considerably in these areas. Using this homogeneity assumption, we calculate that between 16 and 39 electoral outcomes would have changed, or about 2 to 4% of all competitive and contested district councillor elections. We plot this range of effects graphically for good and bad news in Figures S12 and S13.

These estimates are in some ways conservative since a fully saturated treatment is more likely to be seen by voters, and our treatment-effect estimates might be downward biased by spillover. If we instead assume intervention effects in line with our complier average effects (CATE), we estimate between 3 to 7% of these elections are potentially flippable.

**Costs of the Intervention.** Here we briefly discuss the costs of the intervention. That is, the costs to recruit subjects and send messaging, excluding all costs related to the piloting and survey activities. This may provide reasonable bounds on the efficiency of accountability interventions via mobile phones compared to other dissemination methodologies.

In order to recruit 30,296 respondents we hired a team of 82 researchers (2 supervisors, 22 team leaders and 58 research assistants) who visited each of our sampled villages over 20 days. In total the costs for tablets, training, wages, car rental, fuel, per diem, airtime and stipends came to approximately \$96,000. These costs, however, were primarily associated with a village-level public-service audit that we conducted concurrently. Our research team estimated they spent approximately 30% of their time in the field on recruitment activities. Excluding fixed costs for training, tablets and stipends, this means that the recruitment cost approximately \$41,580. If we divide this by the number of respondents, this gives us a cost of \$1.37 per recruit; or \$2.59 if we account for attrition between recruitment and the baseline survey.

To send SMS messages, we relied on a commercial service called SMSOne. We sent 12 messages to each of the 8,063 respondents in the treatment group. These costs us 29.5 UGX, or \$0.009 USD to send, so our total treatment cost was about \$870; or about \$0.11 per respondent.

Our recruitment protocol was similar in logistical details to the group-meeting or door-to-door dissemination protocols used in other studies (4–6); and therefore our recruitment cost of \$2.59 represents a reasonable lower-bound on the counter-factual for the costs of delivering information in-person instead of via SMS messaging. A similarly structured door-to-door information treatment conducted in Kampala by Ferree, Dowd, Jung and Gibson (7) similarly estimated a cost of \$10,000 to implement a GOTV information campaign among 1,500 eligible voters in Kampala (i.e. with very few travel costs) and to survey them after the 2011 elections (authors' estimates). Thus we believe \$1.37–\$2.59 is a conservative estimate for an in-person information treatment with similar scope and content. By these numbers, our cell-phone information treatment was approximately 12–24 times cheaper; or 1.9 times cheaper if we add the costs of an in-person recruitment to the dissemination costs. Yet our treatment effects are comparable to in-person accountability interventions (6).

**Feasibility of the Intervention.** To consider the feasibility of civil-society organizations implementing a similar intervention using ICT, Table S22 includes several examples of NGOs using SMS campaigns in their programs.

**Updating of Respondent Prior Beliefs about Budget Performance.** We show the extent to which respondent beliefs about budget performance changed between baseline and endline in the treatment group as compared to the control group in Figure S14. For this figure, we code prior and posterior beliefs on a 1 to 5 scale, based on whether respondents indicated their LC5 was “much worse,” “a little worse,” “don’t know,” “better,” or “much better” at managing budgets than other districts, the same measure we took during the baseline survey. We then use a two sided t-test to compare changes between priors and posteriors between treatment and control by good- and bad-news groups. Consistent with what has been found in other electoral contexts (8), there is considerable instability in beliefs over time: the average change in beliefs between baseline and endline is  $-0.05$ , but with a standard deviation of 1.49. Only about 20% of respondents maintain consistent beliefs between baseline and endline. Despite this noisy environment, we see the subjects' beliefs become more positive in the good news group and more negative in the bad news groups when treated as compared to control. Among good-news eligible recipients, treatments increase the difference between priors and posteriors by 0.06 on average ( $p = 0.11$ ), and by 0.08 among verified recipients ( $p = 0.10$ ). Among bad news eligible recipients, treatment decreases the difference between priors and posteriors by 0.04 on average ( $p = 0.34$ ), and by 0.10 among verified recipients ( $p = 0.04$ ).

While the overall descriptives are suggestive of updating beliefs in the ways that are consistent with the news groups, a closer look at individual-level updating taking into account priors show only marginal evidence of updating. In Table S23, 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. 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. Directional updating is moving posterior beliefs in the direction of the treatment information or having correct prior and posterior beliefs. Although all specifications find point



estimates that indicate positive updating as a result of treatment, being treated with budget information is only inconsistent with the null hypothesis for loose updating. In our sample, subjects on average have posterior beliefs that are further away from treatment information than their prior beliefs. The positive finding for loose updating indicates that treatment may help voters not be driven toward even more incorrect beliefs about politicians' performance when treated, which might come about through campaign effects. Nonetheless, this evidence is marginal. This analysis was not pre-registered and is supplementary to probe mechanisms.

Updating beliefs is only one of several mechanisms through which information in our experiment may have affected voting. Information can also increase voters' certainty about politician performance or the salience of performance information in voting, neither of which we measured. Future research could fruitfully explore these mechanisms. Given the multi-part nature of our treatment, treatments may also have changed voters' beliefs along other unmeasured dimensions. Additionally, our measures of prior and posterior beliefs were both coarse and noisy, which may have prevented detection of updating. Future work might additionally look into the variance in beliefs as a function of treatment, since updating implies a distribution for beliefs rather than a point estimate.

**Multiple, Joint, and Pooled Hypothesis Testing.** In the main text, Table S14, and Table S15, we report standard errors (standard deviations of the randomization distribution under the sharp null) and one-sided  $p$ -values for individual specifications. While these values are accurate representations of uncertainty under the sharp null hypothesis for individual tests, they do not take into account that we have two tests for our hypotheses on good and bad news (chairs and councillors), nor do they characterize the joint uncertainty about the probability of estimating a positive effect for good news *and* a negative effect for bad news within a single office or across the two offices under consideration. Here we further explore the implications of multiple, joint, and pooled hypothesis tests. We did not pre-specify any particular tests of these kinds for our individual project, so these should be seen as augmenting our core analysis and probing its robustness.

For multiple tests, consider that our models produce a realized set of estimated treatment effects,

$$b^* = \{b_{\text{chair},+}^*, b_{\text{chair},-}^*, b_{\text{councillor},+}^*, b_{\text{councillor},-}^*\}$$

where  $+$  indexes the good news subgroup and  $-$  indexes the bad news subgroup. Each random assignment under the sharp null likewise produces four values  $\tilde{b}$  of estimated effects,

$$\tilde{b} = \{\tilde{b}_{\text{chair},+}, \tilde{b}_{\text{chair},-}, \tilde{b}_{\text{councillor},+}, \tilde{b}_{\text{councillor},-}\}$$

To investigate the implications of multiple tests within the good and bad news subgroups, we consider how often either of the offices would have shown an effect larger than the maximum (good news) or minimum (bad news) effect realized values in  $b^*$  relevant for the particular subgroup  $L^+$  or  $L^-$ . More formally, the good news multiple test considers the proportion of assignments under the sharp null where  $\max(\tilde{b}_+) > \max(b_+^*)$ . This test corrects for the higher likelihood of finding a significant result on at least one test when considering estimates from two tests. Likewise, the bad news multiple test considers the proportion of assignments under the sharp null where  $\min(\tilde{b}_-) < \min(b_-^*)$ . As expected, these test roughly double the  $p$ -value of obtaining at least one out of two tests that exceeds the maximum or minimum realized estimates under the sharp null.

In addition to multiple testing, our results can be interpreted as joint tests for the entire sample, rather than independent tests within the good and bad news subgroups. For this analysis, we consider the probability of observing results in the expected direction under the sharp null hypothesis that are more extreme than both of the realized values, either within single offices or across both offices. This provides evidence about how rare our results would be under the sharp null considering the effects that we observe in both directions. In particular, our joint test for chairs considers the probability that  $\tilde{b}_{\text{chair},+} > \max(b_{\text{chair}}^*) \cap \tilde{b}_{\text{chair},-} < \min(b_{\text{chair}}^*)$ . Likewise, for councillors we consider the probability that  $\tilde{b}_{\text{councillor},+} > \max(b_{\text{councillor}}^*) \cap \tilde{b}_{\text{councillor},-} < \min(b_{\text{councillor}}^*)$ . In addition to joint tests within offices, we also consider a joint test across all four outcomes. In this case, we consider the joint probability under the sharp null that the treatment effect for at least one of the offices will exceed the maximum observed value for both offices and that the minimum effect for at least one of the offices will exceed the minimum observed value for both offices. More formally, we consider the probability that  $\max(\tilde{b}_+) > \max(b_+^*) \cap \min(\tilde{b}_-) < \min(b_-^*)$ .

Finally, because our hypotheses were about the effects of good and bad news generally, we consider the pooled effect of good news and bad news across both the chair and councillor offices. We also consider the joint probability of generating the pooled result in both directions under the sharp null hypothesis.

Table S24 shows the results of this process. For multiple testing, as expected, the probability that the results could be generated assuming no unit-level treatment effect increases within both the good and bad news subsets, though not to levels that call our main conclusions into question. For the pooled test, we likewise find that the average effect in the good and bad news subgroups across both offices are not inconsistent with what could have been observed if the sharp null hypothesis were true. This result can be visually observed by roughly averaging the treatment effects displayed in Figure 1. When we consider the joint probability of having observed both the positive effects of good news and the negative effects of bad news under the sharp null, our confidence in the main results increases dramatically. In some cases, there are no randomization draws assuming the sharp null out of 10,000 that would have produced values as or more extreme than the actual estimated treatment effects we observe. Thus, considering joint probabilities adds significant empirical support to our argument that the treatment helps voters hold some officials accountable for budget irregularities.

**Weighting and the fixed-effects estimator.** The fixed-effects estimator that we use for estimation can be substantively biased when strata defined by the fixed-effects have both (i) heterogeneity in the variance of treatment assignment; and (ii) heterogeneity in treatment effects (9). Recall that the budget treatment is assigned using a two-stage process. First, villages with at least 15 participating subjects are assigned to a density lottery within matched pairs within districts. One village in each matched pair is assigned to be treated at 80% density and the other at 20% density. Villages with fewer than 15 participating subjects are not assigned to the density lottery and are all treated at 50% density. Individual subjects are then assigned to treatment according to the assigned or fixed village-level densities. Although all subjects have a 50% chance of assignment to treatment with this process, villages not eligible for the density lottery have higher variance in treatment assignment than villages eligible for the density lottery. The main results we report are thus interpretable as the sample average treatment effect if we assume that treatment effects are constant across villages with different treatment densities. Note that even if this assumption is violated, the uncertainty about the resulting test statistics in the main text is still correct given the design because we derive it from randomization inference.

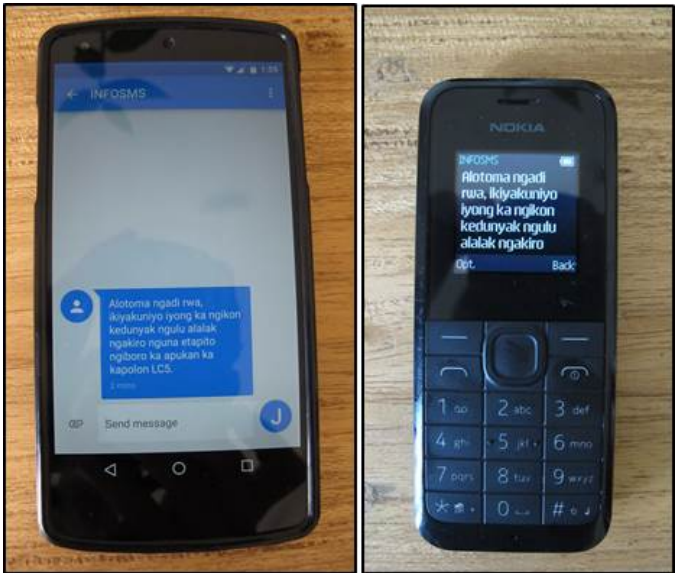
We also demonstrate how the results may be affected by accounting for heterogeneity in treatment effects that corresponds to heterogeneity in treatment assignment. The fixed-effects estimator up-weights strata with more variance in treatment assignment, such that the resulting estimate of effects no longer reflects the sample frequencies (10). One approach is to undo this implicit weighting by re-weighting all observations based on the inverse of the variance of their treatment assignment.<sup>†</sup> Intuitively, this is analogous to running separate estimations using the same specification for density eligible and ineligible villages and then taking a weighted average of the estimates by sample frequencies (this is not precisely correct but very close, since separate estimations will allow other parameters to vary as well).

When we take into account weights based on the inverse in the variance of treatment assignment in Table S25, the estimates are not significantly different than the original strategy, which has an embedded constant effects assumption. Nonetheless, the estimated effects of good news on reported votes for councillors attenuate slightly and are more consistent with the null hypothesis. In conclusion, taking into account fixed-effect weights causes us to be more cautious about the effect of good news of reported votes for councillors. This analysis using weights was not pre-registered and is an additional analysis to probe the robustness of our main results.

## References.

1. Selb P, Munzert S (2013) Voter overrepresentation, vote misreporting, and turnout bias in postelection surveys. *Electoral Studies* 32(1):186–196.
2. Grossman G, Michelitch K (2018) Information dissemination, competitive pressure, and politician performance between elections: A field experiment in Uganda. *American Political Science Review*.
3. Buntaine M, Bush S, Jablonski R, Nielson D, Pickering P (forthcoming) Informing voters in Uganda about budget management and public services by text-messaging in *Metaketa I: The Limits of Electoral Accountability*, eds. Dunning T, Grossman G, Humphreys M, Hyde S, McIntosh C. (Cambridge University Press, Cambridge), p. forthcoming.
4. Humphreys M, Weinstein JM (2013) Policing politicians: Citizen empowerment and political accountability in Uganda. Unpublished Manuscript.
5. Banerjee A, Kumar S, Pande R, Su F (2010) Do informed voters make better choices? experimental evidence from urban India. Unpublished Manuscript. Available at <http://www.povertyactionlab.org/node/2764> (last accessed March 22, 2017).
6. Dunning T, Grossman G, Humphreys M, Hyde SD, McIntosh C (forthcoming) *Metaketa I: The Limits of Electoral Accountability*. (Cambridge University Press, Cambridge).
7. Ferree K, Jung D, Dowd R, Gibson C (2015) Election ink and turnout in a fragile democracy. Unpublished Manuscript.
8. Achen CH (1975) Mass political attitudes and the survey response. *American Political Science Review* 69(4):1218–1231.
9. Gibbons CE, Serrato JCS, Urbancic MB (2018) Broken or fixed effects? *Journal of Econometric Methods*.
10. Humphreys M (2009) Bounds on least squares estimates of causal effects in the presence of heterogeneous assignment probabilities. Unpublished Manuscript. Available at <http://www.columbia.edu/~mh2245/papers1/monotonicity7.pdf> (last accessed March 13, 2018).

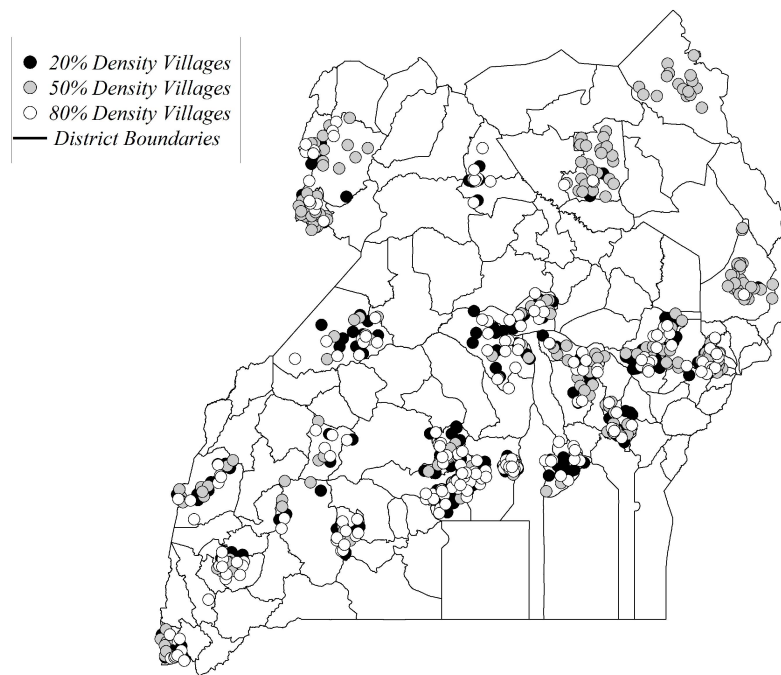
<sup>†</sup> We thank Macartan Humphreys for this suggestion.



Panel A. Smart Phone

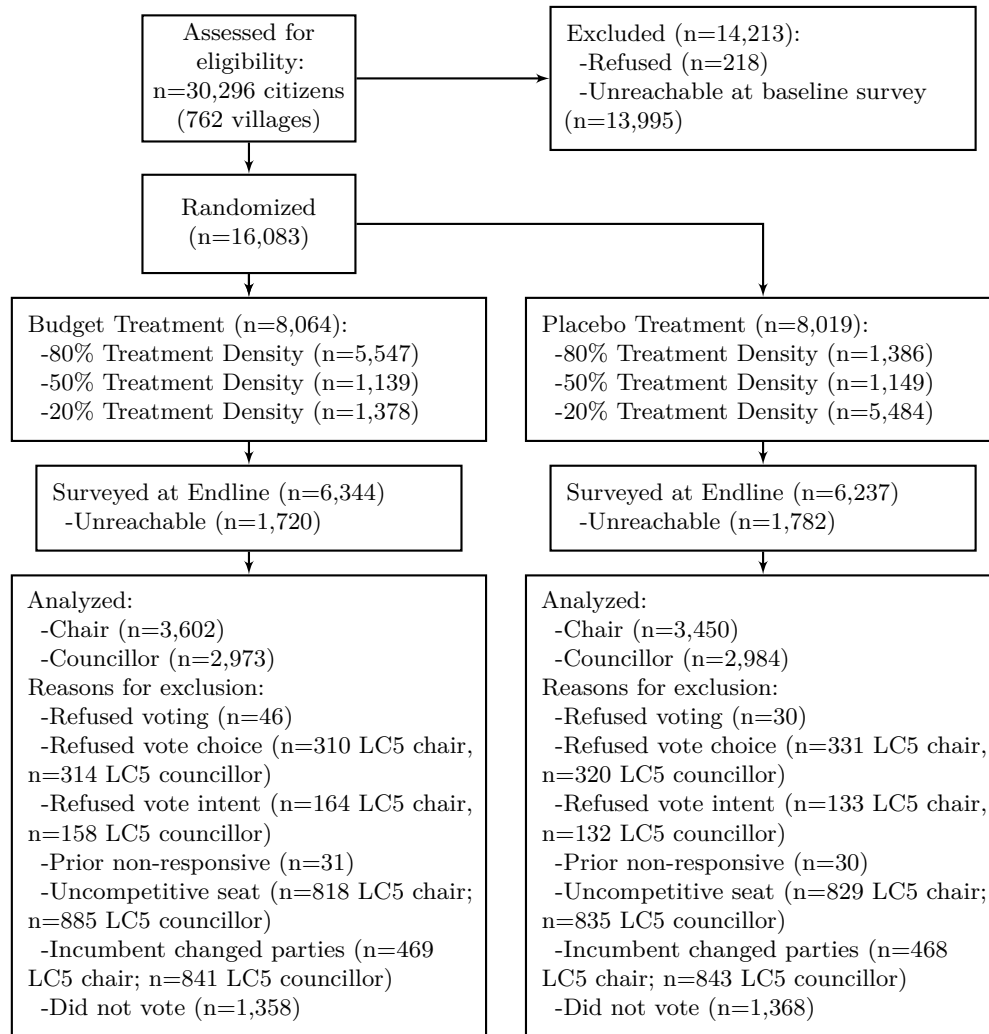
Panel B. Simple Phone

**Fig. S1. 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.

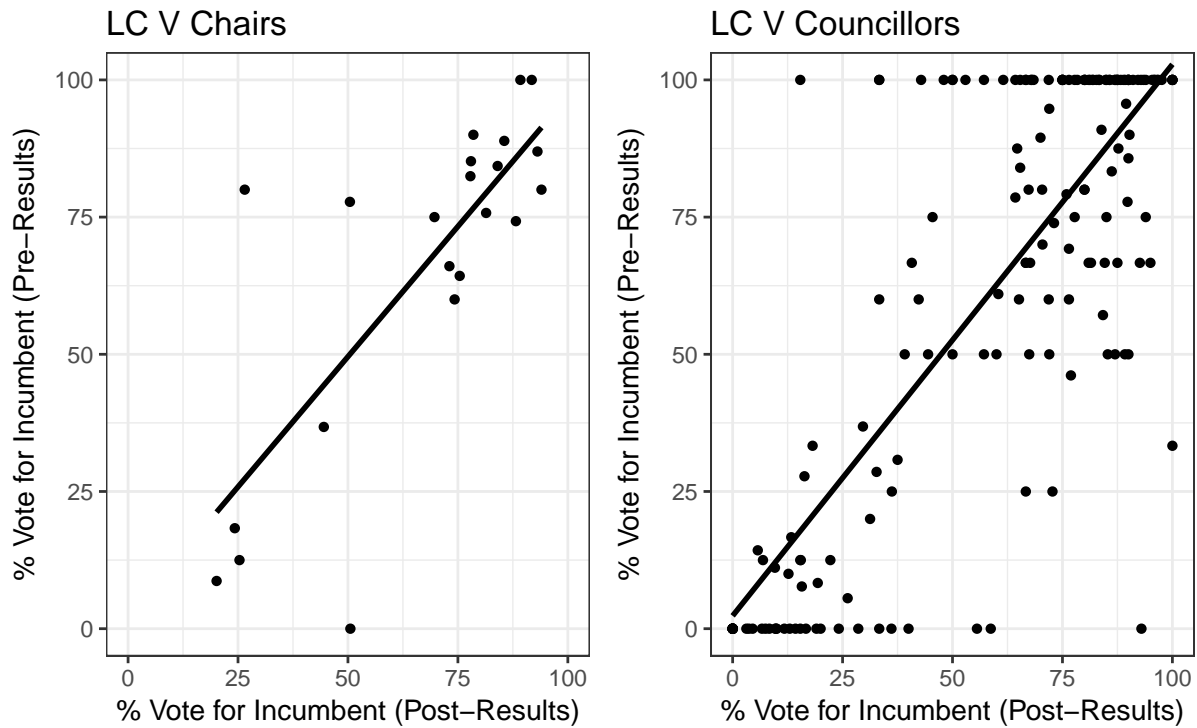


**Fig. S2. Map of sampled villages in Uganda.** Population-weighted sample of 30 villages within each of 27 randomly sampled districts. Black dots indicate villages with at least 20 subjects where 20% of subjects were treated. Grey dots indicate villages with fewer than 20 baseline respondents where 50% of villages were treated. White dots indicate villages with at least 20 subjects where 80% of subjects were treated.

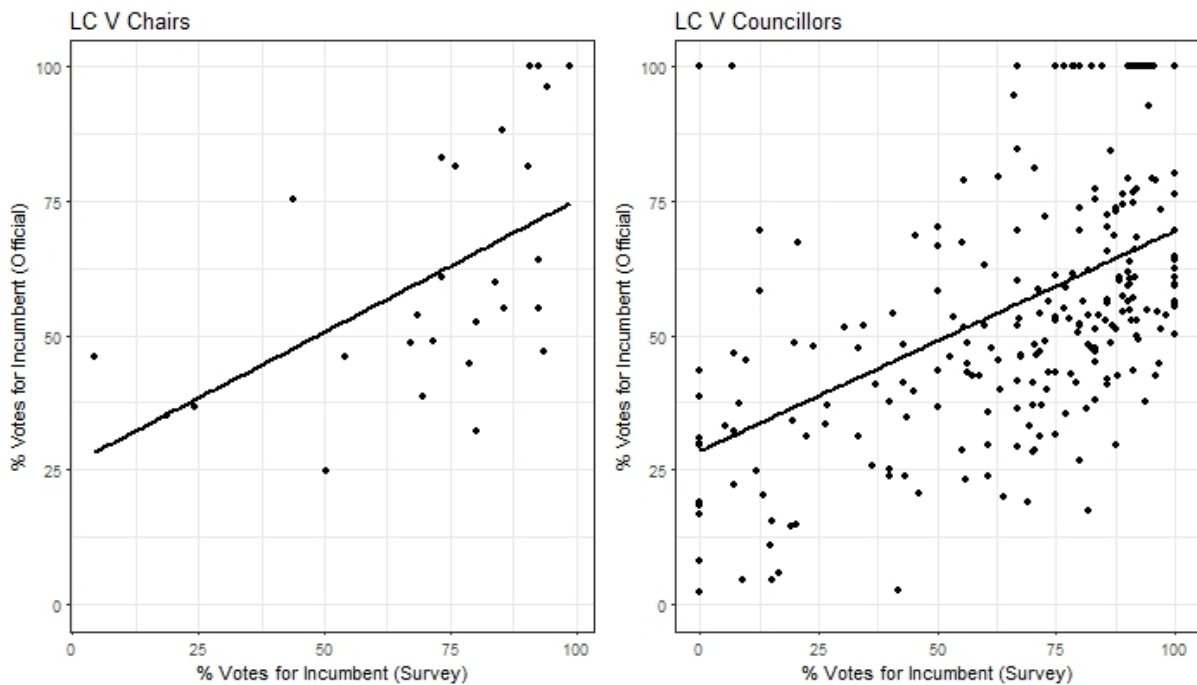




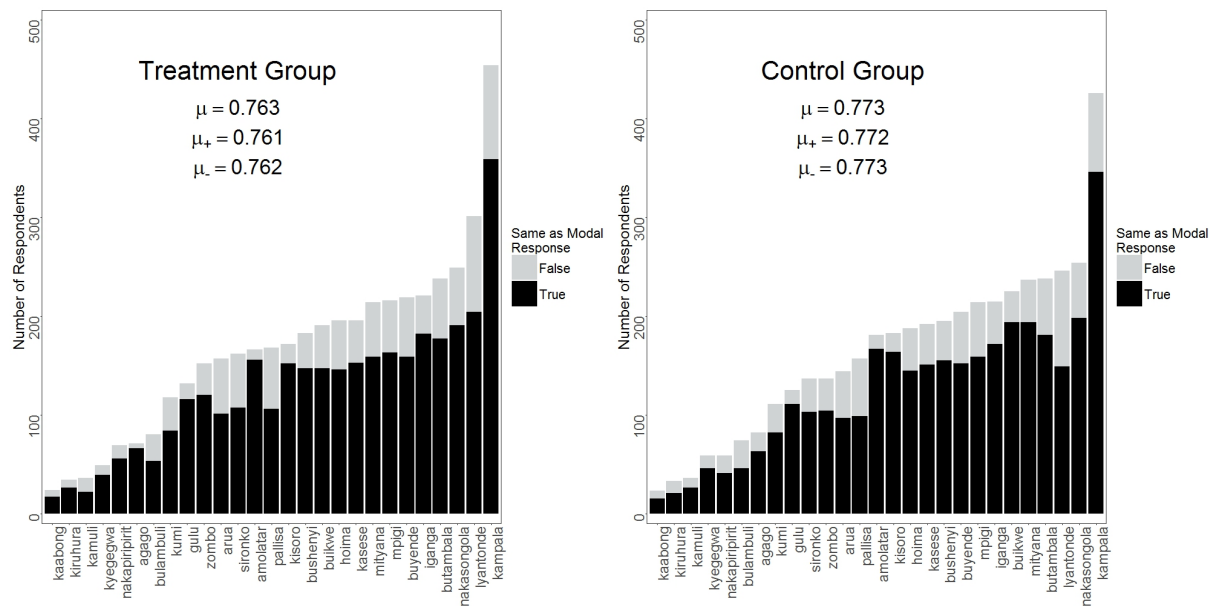
**Fig. S3. CONSORT diagram tracking study design.** This diagram shows the number of subjects from intake through analysis, including key steps that where subjects are excluded from the sample, including an inability to reach them for the baseline survey, attrition at the endline survey, village-level 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.



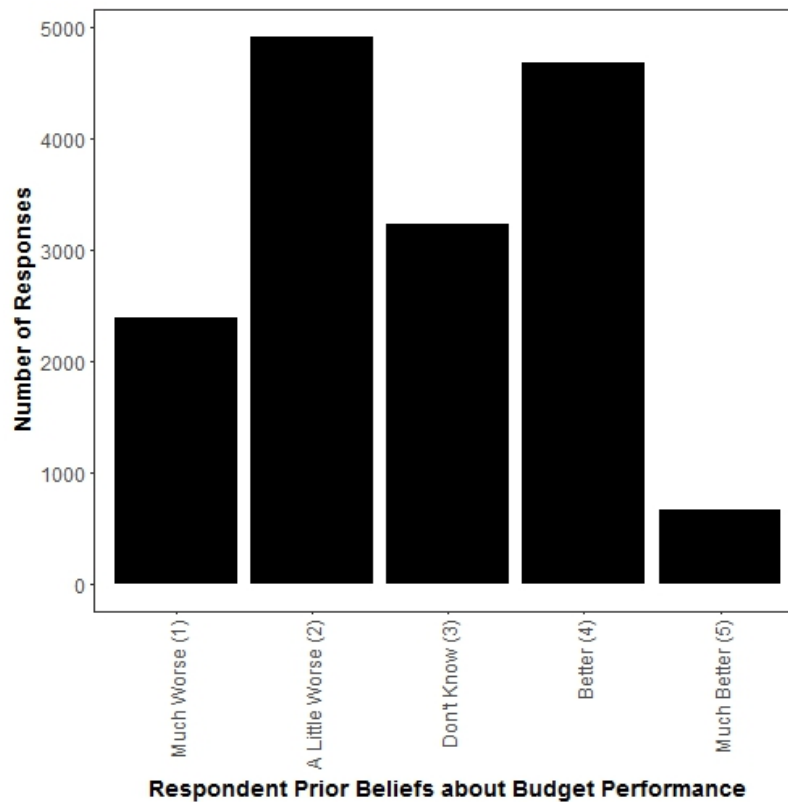
**Fig. S4. Comparison of incumbent vote share, pre- and post-election results being announced by politician.** The graphs show the proportion of voters in each district and sub-county that reported voting for the incumbent in our endline survey, comparing responses after the election results had been announced (approximately 90% of respondents in our survey) to responses before the election results had been announced (approximately 10% of respondents in our survey). Election results were announced approximately 24 hours after the election and our post-election call center began immediately when polls closed and lasted several days. We do not find evidence that announcement of the election results systematically affects reported vote choice. Note that pre-results reported vote choice proportions cluster at 0% and 100% because of low sample size.



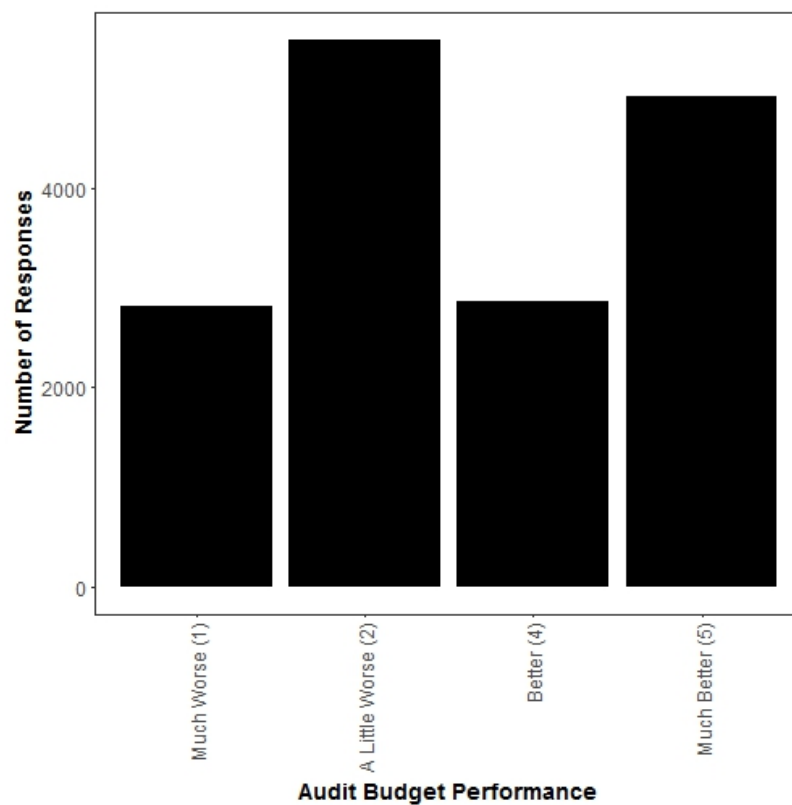
**Fig. S5. Comparing self-reported vote choice data from our study to official returns.** These graphs display the proportion of voters in our study's districts who reported voting for the incumbent party for both chair and councillor offices and compares this proportion with official results released by the Electoral Commission of Uganda. While sampling error is evident, these results show that our main reported vote choice measure is associated with official returns as would be expected for a valid measure.



**Fig. S6. Share of respondents in each district able to name the water basin color at their polling place.** Our data appear consistent with accurate reporting: 93% of voting respondents recalled a basin or basin color in their polling station, and among these 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.  $\mu_+$ ,  $\mu_-$ ,  $\mu_+$  indicate the means of correct responses for good news eligible, bad news eligible and all respondents in each group.

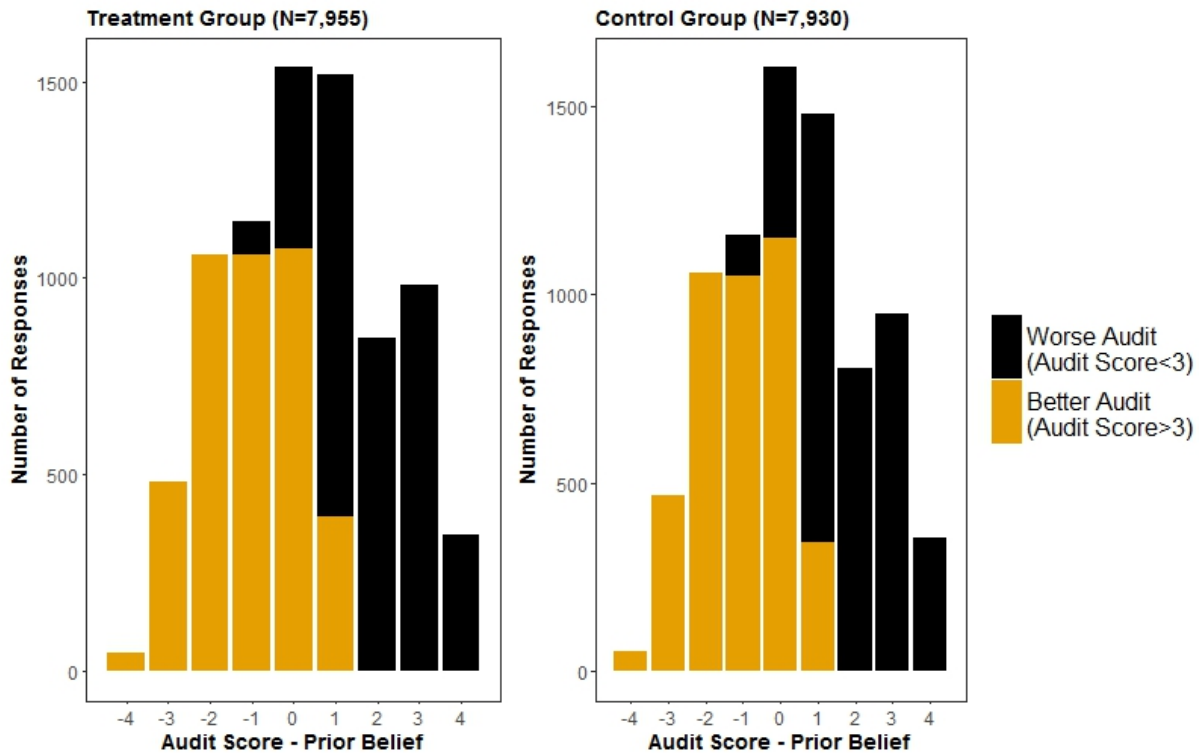


**Fig. S7. Descriptive data for voters' priors about budget management at the district level.** Figure shows distribution of responses to question, "If you compare your LC5's record of managing its budget expenditures and contracting to other districts in Uganda how do you think it will compare? much better; better; a little worse; much worse; don't know; refused to answer" as part of a pre-treatment baseline survey.

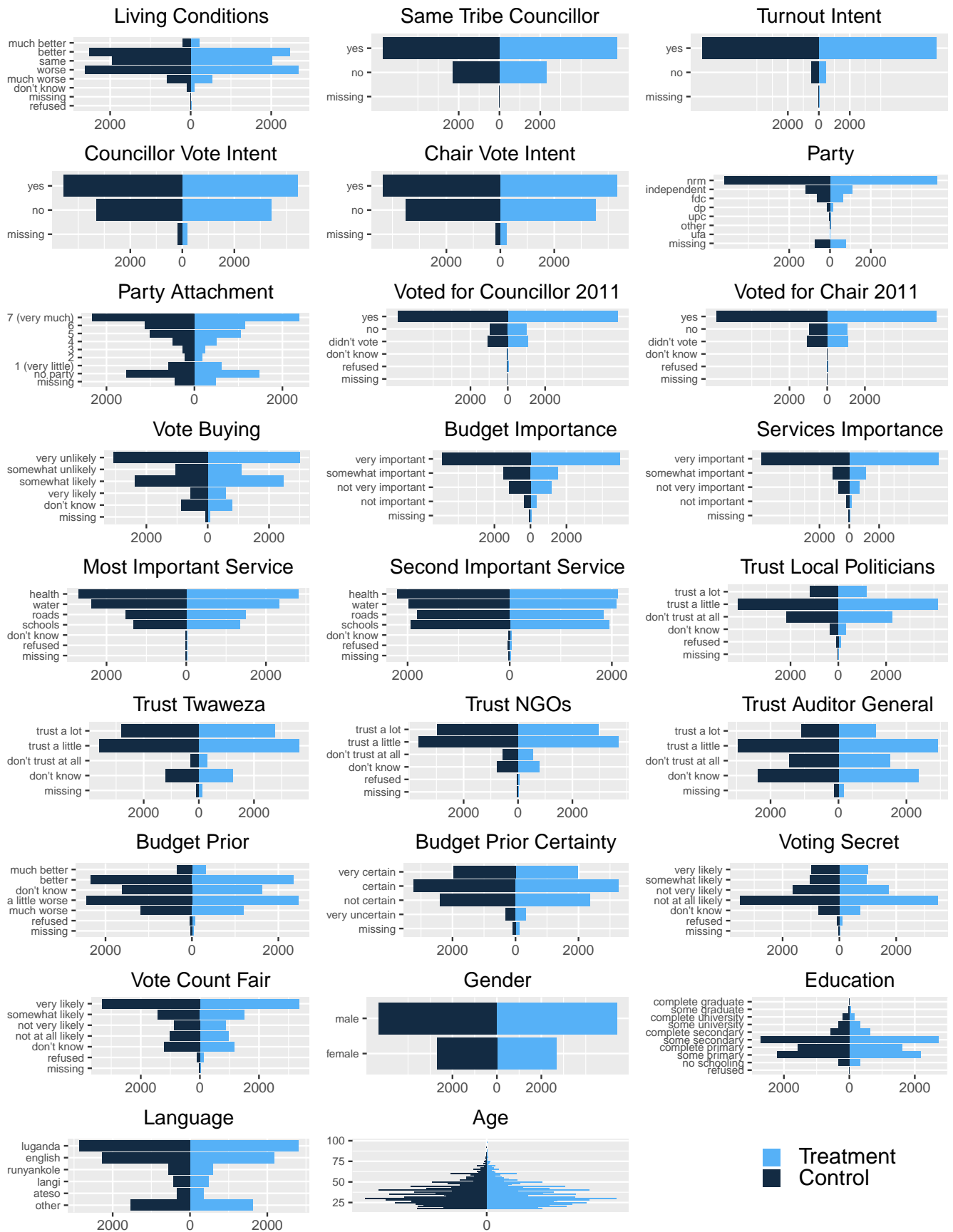


**Fig. S8. Distribution of audit performance scores for budget management disseminated as part of treatment.** Councils are coded as being "much worse," "a little worse," "better" or "much better" based upon whether they are in the first, second, third or fourth quartile of irregular expenditures relative to other district councils in Uganda. Note that we ended up with fewer "Much Worse" and "Better" districts due to random sampling variation.

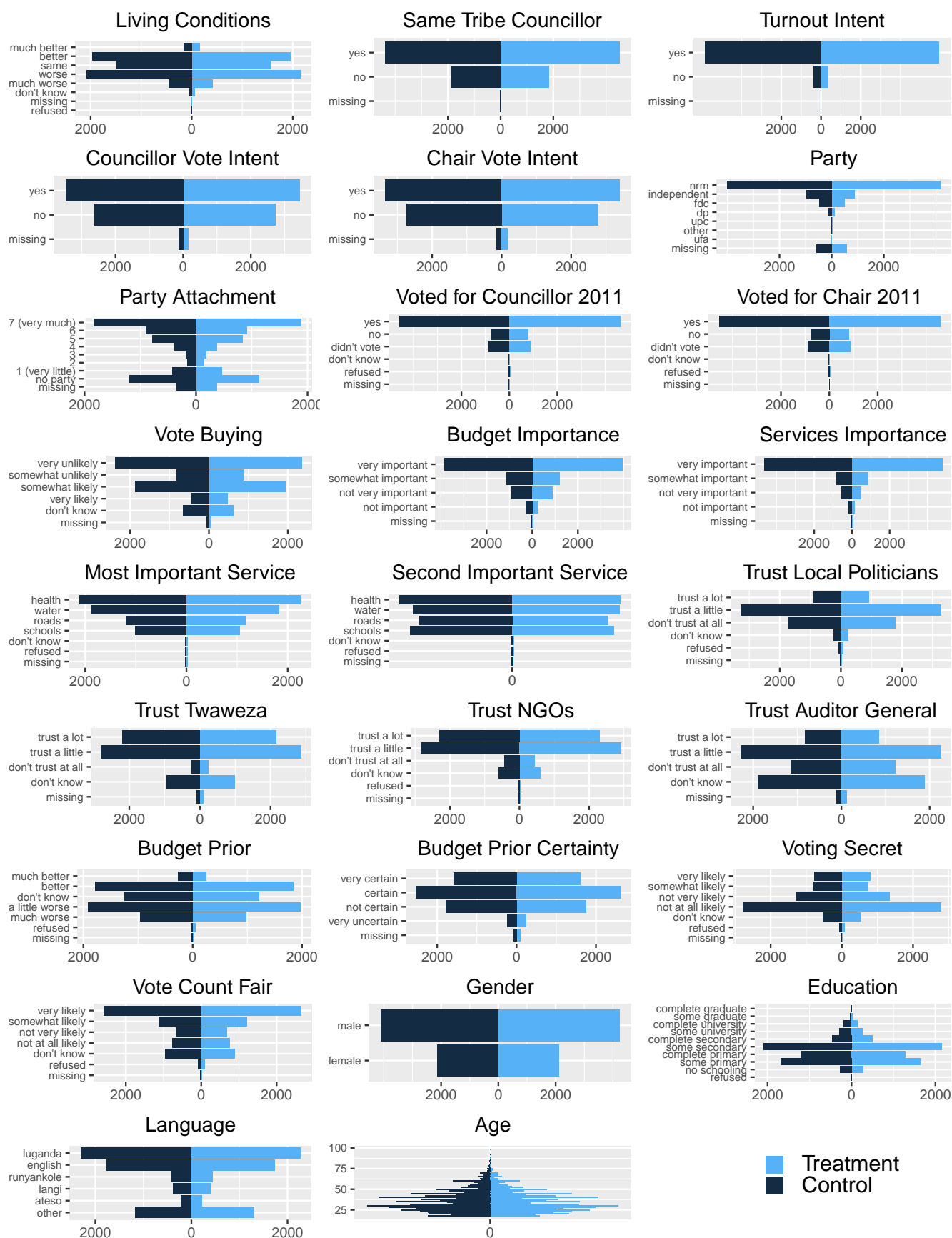




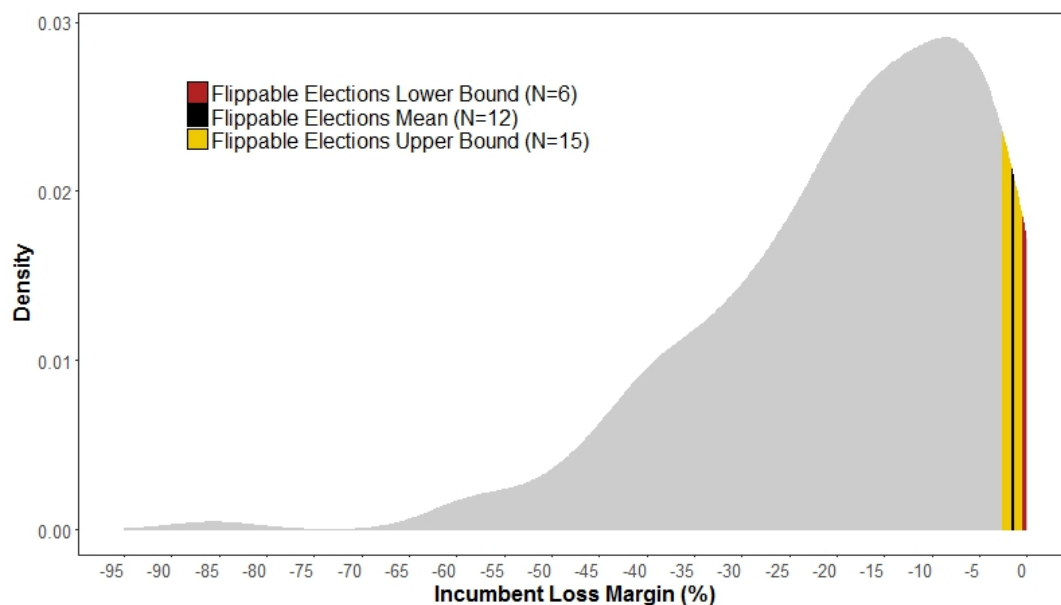
**Fig. S9. Difference between audit and voters' prior budget performance scores.** This figure shows the difference between respondent priors about budgets and true audit scores. Priors are coded on 1 to 5 scale based upon whether respondents believed their council's "record of managing its budget expenditures and contracting" was "Don't Know," "Much Worse," "A Little Worse," "A Little Better," or "Much Better" than others. Audit scores are coded on a 1 to 5 scale based upon whether the share of irregularities in the budget fell in the top, third, second or first quartile of Uganda's districts. This figure demonstrates that, while there is convergence between priors and audits on average, most respondents do not correctly rank their council's budget performance. Dark bars indicate respondents who were in districts eligible to receive messages saying their LC5 was "worse" than others. Light bars indicate respondents who were in districts eligible to receive messages saying their LC5 was "better" than others. This visually illustrates the fact that most respondents over-rate councils doing "worse" than others and under-rate councils doing "better" than others.



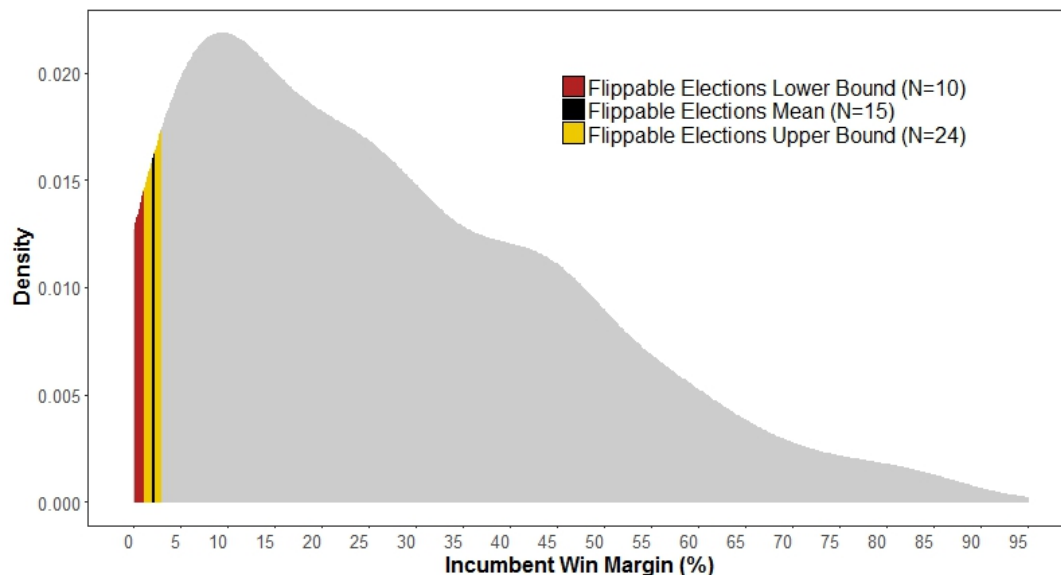
**Fig. S10. Balance of individual-level pre-treatment covariates at baseline.** This figure shows the distribution of each pre-treatment variable by treatment assignment for all subjects contacted in the baseline survey. Light blue indicates treatment-group subjects and dark blue indicates control-group subjects.



**Fig. S11. Balance of individual-level pre-treatment covariates among subjects reached during the endline survey, after attrition from baseline.** This figure shows the distribution of each pre-treatment variable by treatment assignment for all subjects contacted in the endline survey (and therefore included in subsequent analysis). Light blue indicate treatment group subjects and dark blue indicates control group subjects

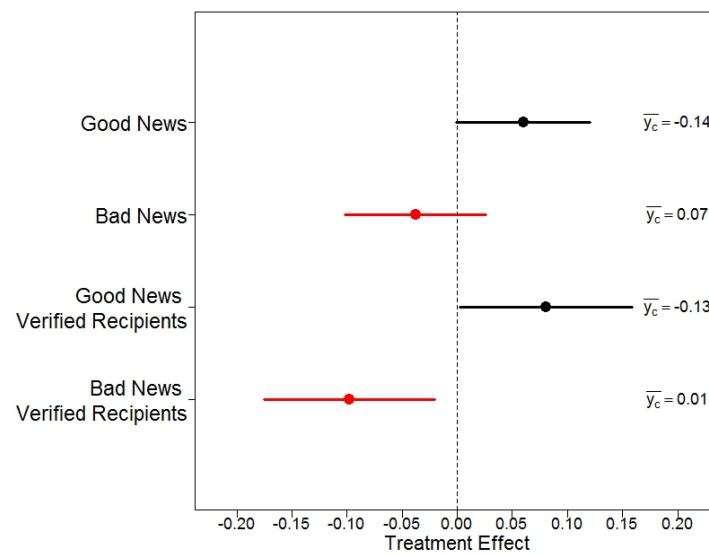


**Fig. S12. Proportion of Flippable Election Losses in Positive News Eligible Elections.** The plot shows all density of all competitive and contested district councillor election losses by the incumbent party's win margin. In green we show the density of potentially flippable elections assuming intervention effects associated with a PATE equivalent to SATE plus one standard deviation. In red we show the density of potentially flippable elections assuming intervention effects associated with a PATE equivalent to SATE minus one standard deviation.



**Fig. S13. Proportion of Flippable Election Wins in Negative News Eligible Elections.** The plot shows all density of all competitive and contested district council election wins by the incumbent party's win margin. In green we show the density of potentially flippable elections assuming intervention effects associated with a PATE equivalent to SATE plus one standard deviation. In red we show the density of potentially flippable elections assuming intervention effects associated with a PATE equivalent to SATE minus one standard deviation.





**Fig. S14. Effect of treatment on changes in respondent beliefs about budget performance.** At baseline and endline respondents were asked to rank councils based upon whether their budget performance was "much worse," "a little worse," "better" or "much better". For this figure, we code these performance priors and posteriors on a 1 to 4 scale. We then use a two-sided t-test to compare changes between priors and posteriors between treatment and control. Points indicate the difference in means between treatment and control. 90% confidence intervals are shown in the horizontal bars. Control group means are on the right.

## Supplementary Tables

**Table S1. Example SMS messages in treatment and placebo groups.**

	Treatment, Much Worse Budget	Treatment, Better Budget	Placebo
1	The Auditor General conducts yearly audits to record instances where LC5s could not satisfactorily explain how its money has been spent.	The Auditor General conducts yearly audits to record instances where LC5s could not satisfactorily explain how its money has been spent.	Research suggests that households which open a savings account are better able to save for school books and fees.
2	Unexplained spending is often an indicator of mismanagement, fraud or poor quality services.	Unexplained spending is often an indicator of mismanagement, fraud or poor quality services.	When youth learn to save and manage their money well, they often have higher savings and income later on in life.
3	Your LC5 did much worse than most other LC5s in the recent audit	Your LC5 did better than most other LC5s in the recent audit	
4	In your LC5, the auditor found issues with 120 million UGX from its budget of 19 billion UGX. This is much worse than in other districts.	In your LC5, the auditor found issues with 14 million UGX from its budget of 27 billion UGX. This is better than in other districts.	
5	This means that 6.3 out of 1000 UGX in your LC5 budget had issues. In most LC5s 2.2 out of 1000 UGX had issues. Your LC5 did much worse than average.	This means that 0.5 out of 1000 UGX in your LC5 budget had issues. In most LC5s 2.2 out of 1000 UGX had issues. Your LC5 did better than average.	
6	One reason your LC5 did much worse than average is that payments of 98 million UGX were made without proper documentation	One reason your LC5 did better than average is that payments on contracts were satisfactorily explained	
7	Another reason your LC5 did much worse than average is that a bid for borehole construction included unexplained expenditures	Another reason your LC5 did better than average is that all payroll spending was satisfactorily explained	

Table shows example budget messages sent to villages where the district council performed much worse (column 1) and where the district council performed better (column 2). Column 3 shows messages sent to subjects in the placebo condition. All messages were customized to individual districts based on the results of the Auditor General report on district budgets.

**Table S2. Conditional effects of budget treatment by trust in Twaweza and Auditor General.**

	DV: Vote Choice for the Incumbent							
	LC5 Chair	LC5 Councillor	LC5 Chair	LC5 Councillor	LC5 Chair	LC5 Councillor	LC5 Chair	LC5 Councillor
	Trust in Auditor General				Trust in Twaweza			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Budget Treatment	−0.047 (0.022) p = 0.036	−0.047 (0.039) p = 0.224	0.003 (0.038) p = 0.936	−0.040 (0.034) p = 0.245	0.093 (0.085) p = 0.275	0.183 (0.085) p = 0.031	−0.026 (0.079) p = 0.741	0.064 (0.061) p = 0.297
AG Trust: Don't Know	−0.011 (0.015)	−0.052 (0.038)	−0.014 (0.017)	−0.034 (0.039)				
AG Trust: Missing	−0.098 (0.099)	0.074 (0.142)	0.030 (0.059)	−0.043 (0.037)				
AG Trust: A Little	−0.026 (0.027)	−0.049 (0.038)	−0.002 (0.030)	0.011 (0.037)				
AG Trust: A Lot	0.002 (0.022)	0.008 (0.050)	0.026 (0.044)	0.014 (0.038)				
Twaweza Trust: Don't Know					0.063 (0.091)	0.045 (0.076)	0.083 (0.072)	0.043 (0.064)
Twaweza Trust: Missing					−0.335 (0.152)	0.069 (0.166)	0.022 (0.202)	−0.002 (0.056)
Twaweza Trust: A Little					0.026 (0.082)	0.103 (0.075)	0.059 (0.064)	0.022 (0.048)
Twaweza Trust: A Lot					0.045 (0.084)	0.077 (0.074)	0.063 (0.051)	0.031 (0.058)
Treatment *	0.054 (0.020)	0.094 (0.059)	0.008 (0.029)	0.026 (0.045)				
AG Trust: Don't Know	p = 0.007	p = 0.108	p = 0.796	p = 0.561				
Treatment *	0.154 (0.170)	0.036 (0.223)	−0.109 (0.082)	0.029 (0.070)				
AG Trust: Missing	p = 0.364	p = 0.873	p = 0.183	p = 0.679				
Treatment *	0.064 (0.037)	0.102 (0.052)	0.027 (0.046)	0.002 (0.042)				
AG Trust: A Little	p = 0.084	p = 0.050	p = 0.553	p = 0.963				
Treatment *	0.076 (0.032)	0.092 (0.065)	−0.029 (0.069)	−0.022 (0.044)				
AG Trust: A Lot	p = 0.019	p = 0.159	p = 0.669	p = 0.610				
Treatment *					−0.154 (0.087)	−0.137 (0.097)	0.015 (0.096)	−0.114 (0.075)
Twaweza Trust: Don't Know					p = 0.078	p = 0.156	p = 0.877	p = 0.129
Treatment *					0.614 (0.164)	−0.203 (0.218)	−0.039 (0.191)	−0.027 (0.117)
Twaweza Trust: Missing					p = 0.0002	p = 0.353	p = 0.840	p = 0.814
Treatment *					−0.084 (0.093)	−0.173 (0.093)	0.051 (0.080)	−0.080 (0.063)
Twaweza Trust: A Little					p = 0.370	p = 0.063	p = 0.520	p = 0.206
Treatment *					−0.086 (0.096)	−0.151 (0.088)	0.028 (0.081)	−0.125 (0.070)
Twaweza Trust: A Lot					p = 0.373	p = 0.088	p = 0.728	p = 0.073
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
News Type	Good	Good	Bad	Bad	Good	Good	Bad	Bad
Observations	3,921	3,074	3,131	2,883	3,921	3,074	3,131	2,883
Adjusted R <sup>2</sup>	0.205	0.325	0.344	0.396	0.206	0.323	0.344	0.396

Notes: Baseline category in models is no trust in the Auditor General or Twaweza. SEs clustered by politician; two-tailed tests; contested elections only without party switching by incumbent. We do not find consistent effects that responses to the treatment are conditional on trust in the source of the budget management information or the organization that disseminated the information as part of treatment.

**Table S3. Effect of treatment on reported vote choice among voters who expect free and fair district elections.**

	DV: Vote Choice for the Incumbent			
	LC5 Chair Good News		LC5 Councillor Bad News	
	(1)	(2)	(3)	(4)
Budget Treatment (RI)	−0.015 (0.028) p=0.715	0.043 (0.027) p=0.069	0.041 (0.025) p=0.947	−0.039 (0.026) p=0.071
LC5 Chair Intent	0.089 (0.023)		0.032 (0.032)	
LC5 Councillor Intent		0.063 (0.026)		0.010 (0.029)
Polling station fixed effects	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes
Observations	1,655	1,287	1,412	1,246
Adjusted R <sup>2</sup>	0.226	0.400	0.356	0.396

Notes: SEs on treatment derived from randomization distribution under the sharp null using exact clustered design; SEs on covariates clustered by politician; one-tailed tests matching directional hypotheses; contested elections only with incumbent who did not switch party.

**Table S4. Conditional effects of budget treatment by unsure budget prior**

	DV: Vote Choice for the Incumbent							
	LC5 Chair Good News		LC5 Councillor Bad News		LC5 Chair Bad News		LC5 Councillor Bad News	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Budget Treatment	0.003 (0.022) p = 0.905	0.003 (0.020) p = 0.895	0.029 (0.018) p = 0.098	0.029 (0.018) p = 0.103	0.014 (0.018) p = 0.432	0.013 (0.018) p = 0.480	−0.043 (0.015) p = 0.004	−0.043 (0.015) p = 0.006
Unsure Budget Prior	0.067 (0.022) p = 0.003	0.055 (0.028) p = 0.048	−0.020 (0.033) p = 0.547	−0.039 (0.033) p = 0.250	0.026 (0.023) p = 0.251	0.022 (0.023) p = 0.320	−0.006 (0.021) p = 0.762	−0.013 (0.021) p = 0.548
LC5 Chair Intent	0.083 (0.038) p = 0.028	0.076 (0.033) p = 0.022			0.036 (0.024) p = 0.128	0.032 (0.023) p = 0.169		
LC5 Councillor Intent			0.060 (0.016) p = 0.0002	0.051 (0.016) p = 0.002			0.002 (0.017) p = 0.914	0.005 (0.016) p = 0.761
Budget Treatment * Unsure Budget Prior	−0.007 (0.035) p = 0.855	−0.006 (0.032) p = 0.861	−0.011 (0.046) p = 0.811	−0.007 (0.047) p = 0.882	−0.015 (0.023) p = 0.513	−0.015 (0.022) p = 0.489	0.046 (0.031) p = 0.137	0.039 (0.032) p = 0.234
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Observations	3,921	3,921	3,074	3,074	3,131	3,131	2,883	2,883
Adjusted R <sup>2</sup>	0.201	0.206	0.317	0.324	0.341	0.344	0.392	0.396

Notes: SEs clustered by politician; one-tailed tests matching directional hypotheses; contested elections only. Data subset includes constituencies where the district incumbent is not running again, but excludes constituencies that were redistricted or where incumbents switched parties. Models include the following covariates: perception of living conditions, gender, education, age, trust in information from Twaweza, perception that powerful people will learn about vote choice, perception that vote counting will be fair, voted for incumbent in 2011 election.



**Table S5. Treatment effects of information type unconditional on respondents' priors**

	DV: Vote Choice for the Incumbent							
	LC5 Chair		LC5 Councillor		LC5 Chair		LC5 Councillor	
	Positive News				Negative News			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Budget Treatment	0.004 (0.018) p = 0.414	0.003 (0.018) p = 0.437	0.019 (0.016) p = 0.129	0.019 (0.017) p = 0.130	0.005 (0.014) p = 0.650	0.005 (0.014) p = 0.643	−0.033 (0.014) p = 0.008	−0.034 (0.014) p = 0.007
LC5 Chair Intent	0.084 (0.041)	0.075 (0.035)			0.035 (0.021)	0.031 (0.022)		
LC5 Councillor Intent			0.060 (0.016)	0.049 (0.016)			0.003 (0.016)	0.005 (0.016)
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Observations	3,741	3,741	2,958	2,958	3,369	3,369	3,051	3,051
Adjusted R <sup>2</sup>	0.180	0.190	0.316	0.322	0.367	0.368	0.387	0.390

Notes: SEs clustered by politician; one-tailed tests matching directional hypotheses; contested elections only. The "negative news" group contains all subjects within districts which received messages indicating that their council was performing worse than others. The "positive news" group contains all subjects within districts which received messages indicating that their council was performing better than others. The similarity of these results to those in the main draft indicate that priors had only a small impact on treatment estimates.

Table S6. Balance on all pre-treatment covariates by budget treatment status (N=16,083).

Covariate	Treated (mean)	Control (mean)	Diff.
Councillor Vote Intent_no	0.4252 (0.0055)	0.411 (0.0055)	0.0142 (0.0078)
Second Important Service_water	0.2593 (0.0049)	0.2465 (0.0048)	0.0128 (0.0069)
Voting Secret_not very likely	0.2166 (0.0046)	0.2055 (0.0045)	0.0111 (0.0064)
Most Important Service_health	0.3488 (0.0053)	0.3379 (0.0053)	0.0109 (0.0075)
Language_other	0.2019 (0.0045)	0.1915 (0.0044)	0.0103 (0.0063)
Vote Buying_somewhat likely	0.3074 (0.0051)	0.2974 (0.0051)	0.01 (0.0072)
Living Conditions_same	0.2524 (0.0048)	0.2429 (0.0048)	0.0094 (0.0068)
Trust Local Politicians_don't trust at all	0.2801 (0.005)	0.2715 (0.005)	0.0087 (0.007)
Voted for Chair 2011_no	0.1312 (0.0038)	0.1226 (0.0037)	0.0086 (0.0052)
Vote Count Fair_somewhat likely	0.1855 (0.0043)	0.1772 (0.0043)	0.0083 (0.0061)
Services Importance_very important	0.7455 (0.0049)	0.7374 (0.0049)	0.0082 (0.0069)
Education_complete secondary	0.0799 (0.003)	0.073 (0.0029)	0.0069 (0.0042)
Trust Auditor General_don't trust at all	0.1875 (0.0043)	0.1807 (0.0043)	0.0068 (0.0061)
Voted for Councillor 2011_no	0.1246 (0.0037)	0.1185 (0.0036)	0.0062 (0.0052)
Education_complete primary	0.2018 (0.0045)	0.1957 (0.0044)	0.0061 (0.0063)
Party_nrm	0.6517 (0.0053)	0.646 (0.0053)	0.0057 (0.0075)
Vote Buying_somewhat unlikely	0.137 (0.0038)	0.1314 (0.0038)	0.0056 (0.0054)
Language_lang	0.0605 (0.0027)	0.0551 (0.0025)	0.0054 (0.0037)
Party Attachment_7 (very much)	0.2949 (0.0051)	0.2897 (0.0051)	0.0052 (0.0072)
Chair Vote Intent_missing	0.0269 (0.0018)	0.0219 (0.0016)	0.005 (0.0024)
Budget Importance_somewhat important	0.1907 (0.0044)	0.1861 (0.0043)	0.0047 (0.0062)
Language_runyankole	0.0737 (0.0029)	0.0691 (0.0028)	0.0046 (0.0041)
Party_missing	0.0991 (0.0033)	0.0949 (0.0033)	0.0042 (0.0047)
Vote Count Fair_very likely	0.418 (0.0055)	0.414 (0.0055)	0.004 (0.0078)
Party Attachment_5	0.1296 (0.0037)	0.1256 (0.0037)	0.004 (0.0053)
Living Conditions_worse	0.3326 (0.0052)	0.3286 (0.0052)	0.004 (0.0074)
Councillor Vote Intent_missing	0.0248 (0.0017)	0.021 (0.0016)	0.0039 (0.0024)
Party Attachment_missing	0.06 (0.0026)	0.0565 (0.0026)	0.0035 (0.0037)
Budget Prior Certainty_missing	0.0159 (0.0014)	0.0125 (0.0012)	0.0034 (0.0019)
Party Attachment_1 (very little)	0.0768 (0.003)	0.0735 (0.0029)	0.0033 (0.0042)
Trust NGOs_trust a little	0.4587 (0.0055)	0.4555 (0.0056)	0.0032 (0.0079)
Vote Buying_very likely	0.0743 (0.0029)	0.0712 (0.0029)	0.0031 (0.0041)
Trust Twaweza_trust a little	0.4521 (0.0055)	0.4494 (0.0056)	0.0027 (0.0078)
Party_fdc	0.0823 (0.0031)	0.0798 (0.003)	0.0025 (0.0043)
Voting Secret_refused	0.0131 (0.0013)	0.0107 (0.0012)	0.0024 (0.0017)
Trust Local Politicians_refused	0.014 (0.0013)	0.0116 (0.0012)	0.0024 (0.0018)
Trust Auditor General_missing	0.0193 (0.0015)	0.017 (0.0014)	0.0024 (0.0021)
Voting Secret_very likely	0.125 (0.0037)	0.1227 (0.0037)	0.0023 (0.0052)
Budget Prior_refused	0.0091 (0.0011)	0.0069 (0.0009)	0.0022 (0.0014)
Same Tribe Councillor_yes	0.714 (0.005)	0.7119 (0.0051)	0.0021 (0.0071)
Party_other	0.0048 (0.0008)	0.0027 (0.0006)	0.0021 (0.001)
Budget Prior Certainty_certain	0.4049 (0.0055)	0.4028 (0.0055)	0.0021 (0.0077)
Budget Prior_a little worse	0.3065 (0.0051)	0.3045 (0.0051)	0.002 (0.0073)
Voted for Councillor 2011_didn't vote	0.1355 (0.0038)	0.1337 (0.0038)	0.0019 (0.0054)
Trust Twaweza_missing	0.0139 (0.0013)	0.0121 (0.0012)	0.0018 (0.0018)
Budget Importance_very important	0.6146 (0.0054)	0.6128 (0.0054)	0.0018 (0.0077)
Language_ateso	0.0453 (0.0023)	0.0435 (0.0023)	0.0017 (0.0032)
Most Important Service_schools	0.1669 (0.0042)	0.1652 (0.0041)	0.0017 (0.0059)
Voted for Chair 2011_refused	0.0061 (0.0009)	0.0045 (0.0007)	0.0016 (0.0011)
Voted for Councillor 2011_refused	0.0056 (0.0008)	0.0041 (0.0007)	0.0015 (0.0011)
Education_some secondary	0.3393 (0.0053)	0.3378 (0.0053)	0.0015 (0.0075)
Chair Vote Intent_no	0.4371 (0.0055)	0.4358 (0.0055)	0.0013 (0.0078)
Trust Twaweza_don't trust at all	0.0384 (0.0021)	0.0372 (0.0021)	0.0013 (0.003)
Budget Prior_much worse	0.1489 (0.004)	0.1478 (0.004)	0.0012 (0.0056)
Vote Count Fair_refused	0.0156 (0.0014)	0.0145 (0.0013)	0.0012 (0.0019)
Trust Twaweza_don't know	0.1525 (0.004)	0.1514 (0.004)	0.0011 (0.0057)
Trust Local Politicians_missing	0.0036 (0.0007)	0.0026 (0.0006)	0.001 (0.0009)
Party_upc	0.0061 (0.0009)	0.0051 (0.0008)	0.001 (0.0012)

**Table S6. Balance on all pre-treatment covariates by budget treatment status (N=16,083).**

<b>Covariate</b>	<b>Treated (mean)</b>	<b>Control (mean)</b>	<b>Diff.</b>
Budget Prior Certainty_very uncertain	0.0402 (0.0022)	0.0393 (0.0022)	0.0009 (0.0031)
Second Important Service_roads	0.2269 (0.0047)	0.2261 (0.0047)	0.0008 (0.0066)
Trust NGOs_don't know	0.0971 (0.0033)	0.0963 (0.0033)	0.0008 (0.0047)
Voted for Chair 2011_didn't vote	0.1369 (0.0038)	0.1362 (0.0038)	0.0007 (0.0054)
Most Important Service_don't know	0.0036 (0.0007)	0.0029 (0.0006)	0.0007 (0.0009)
Services Importance_somewhat important	0.1394 (0.0039)	0.1387 (0.0039)	0.0007 (0.0055)
Vote Buying_missing	0.0094 (0.0011)	0.0087 (0.001)	0.0007 (0.0015)
Living Conditions_refused	0.0025 (0.0006)	0.0019 (0.0005)	0.0006 (0.0007)
Education_some graduate	0.005 (0.0008)	0.0044 (0.0007)	0.0006 (0.0011)
Turnout Intent_yes	0.9391 (0.0027)	0.9385 (0.0027)	0.0006 (0.0038)
Trust NGOs_refused	0.0058 (0.0008)	0.0052 (0.0008)	0.0006 (0.0012)
Party Attachment_6	0.142 (0.0039)	0.1414 (0.0039)	0.0006 (0.0055)
Trust Local Politicians_trust a lot	0.1479 (0.004)	0.1474 (0.004)	0.0005 (0.0056)
Trust NGOs_missing	0.0043 (0.0007)	0.0039 (0.0007)	0.0005 (0.001)
Budget Prior Certainty_very certain	0.2468 (0.0048)	0.2464 (0.0048)	0.0004 (0.0068)
Living Conditions_much better	0.0264 (0.0018)	0.0261 (0.0018)	0.0004 (0.0025)
Second Important Service_don't know	0.0042 (0.0007)	0.0039 (0.0007)	0.0004 (0.001)
Party_ufa	0.0002 (0.0002)	0 (0)	0.0002 (0.0002)
Same Tribe Councillor_missing	0.0016 (0.0004)	0.0014 (0.0004)	0.0002 (0.0006)
Budget Prior_missing	0.0045 (0.0007)	0.0042 (0.0007)	0.0002 (0.001)
Living Conditions_don't know	0.0114 (0.0012)	0.0112 (0.0012)	0.0002 (0.0017)
Party Attachment_4	0.0625 (0.0027)	0.0624 (0.0027)	0.0001 (0.0038)
Voted for Chair 2011_missing	0.002 (0.0005)	0.002 (0.0005)	0 (0.0007)
Budget Prior_better	0.2913 (0.0051)	0.2913 (0.0051)	0 (0.0072)
Voting Secret_missing	0.0037 (0.0007)	0.0037 (0.0007)	0 (0.001)
Turnout Intent_missing	0.0033 (0.0006)	0.0035 (0.0007)	-0.0001 (0.0009)
Voted for Councillor 2011_missing	0.0022 (0.0005)	0.0026 (0.0006)	-0.0004 (0.0008)
Second Important Service_missing	0.0026 (0.0006)	0.003 (0.0006)	-0.0004 (0.0008)
Most Important Service_refused	0.0031 (0.0006)	0.0035 (0.0007)	-0.0004 (0.0009)
Trust Auditor General_trust a lot	0.1368 (0.0038)	0.1372 (0.0038)	-0.0004 (0.0054)
Turnout Intent_no	0.0575 (0.0026)	0.058 (0.0026)	-0.0004 (0.0037)
Education_complete graduate	0.0001 (0.0001)	0.0006 (0.0003)	-0.0005 (0.0003)
Second Important Service_refused	0.0037 (0.0007)	0.0042 (0.0007)	-0.0005 (0.001)
Education_no schooling	0.0424 (0.0022)	0.043 (0.0023)	-0.0006 (0.0032)
Voted for Chair 2011_don't know	0.0029 (0.0006)	0.0035 (0.0007)	-0.0006 (0.0009)
Second Important Service_schools	0.2413 (0.0048)	0.2421 (0.0048)	-0.0007 (0.0068)
Vote Count Fair_missing	0.0037 (0.0007)	0.0046 (0.0008)	-0.0009 (0.001)
Education_refused	0.0001 (0.0001)	0.0011 (0.0004)	-0.001 (0.0004)
Living Conditions_missing	0.001 (0.0004)	0.002 (0.0005)	-0.001 (0.0006)
Budget Importance_missing	0.0072 (0.0009)	0.0084 (0.001)	-0.0012 (0.0014)
Trust NGOs_don't trust at all	0.0681 (0.0028)	0.0693 (0.0028)	-0.0013 (0.004)
Voted for Councillor 2011_don't know	0.0026 (0.0006)	0.0039 (0.0007)	-0.0013 (0.0009)
Services Importance_missing	0.0072 (0.0009)	0.0085 (0.001)	-0.0013 (0.0014)
Vote Count Fair_not very likely	0.1085 (0.0035)	0.11 (0.0035)	-0.0015 (0.0049)
Budget Prior_don't know	0.2005 (0.0045)	0.202 (0.0045)	-0.0015 (0.0063)
Party_dp	0.0186 (0.0015)	0.0206 (0.0016)	-0.002 (0.0022)
Trust Auditor General_don't know	0.2938 (0.0051)	0.2959 (0.0051)	-0.0021 (0.0072)
Budget Importance_not very important	0.146 (0.0039)	0.1481 (0.004)	-0.0022 (0.0056)
Voting Secret_don't know	0.0909 (0.0032)	0.0932 (0.0032)	-0.0023 (0.0046)
Same Tribe Councillor_no	0.2844 (0.005)	0.2867 (0.005)	-0.0023 (0.0071)
Education_some university	0.0401 (0.0022)	0.0426 (0.0023)	-0.0026 (0.0031)
Party Attachment_3	0.0295 (0.0019)	0.0325 (0.002)	-0.003 (0.0027)
Budget Importance_not important	0.0415 (0.0022)	0.0446 (0.0023)	-0.0031 (0.0032)
Services Importance_not important	0.0222 (0.0016)	0.0257 (0.0018)	-0.0035 (0.0024)
Trust NGOs_trust a lot	0.3659 (0.0054)	0.3697 (0.0054)	-0.0038 (0.0076)
Budget Prior_much better	0.0392 (0.0022)	0.0433 (0.0023)	-0.0041 (0.0031)
Services Importance_not very important	0.0857 (0.0031)	0.0898 (0.0032)	-0.0041 (0.0045)
Party Attachment_2	0.0222 (0.0016)	0.0263 (0.0018)	-0.0041 (0.0024)
Education_some primary	0.2705 (0.0049)	0.2747 (0.005)	-0.0043 (0.007)

**Table S6. Balance on all pre-treatment covariates by budget treatment status (N=16,083).**

Covariate	Treated (mean)	Control (mean)	Diff.
Vote Count Fair_not at all likely	0.1224 (0.0036)	0.1271 (0.0037)	-0.0047 (0.0052)
Trust Local Politicians_don't know	0.0381 (0.0021)	0.0429 (0.0023)	-0.0048 (0.0031)
Most Important Service_roads	0.1856 (0.0043)	0.1905 (0.0044)	-0.0049 (0.0062)
Gender_female	0.3313 (0.0052)	0.3365 (0.0053)	-0.0051 (0.0074)
Voting Secret_not at all likely	0.4302 (0.0055)	0.4356 (0.0055)	-0.0054 (0.0078)
Living Conditions_much worse	0.0672 (0.0028)	0.073 (0.0029)	-0.0057 (0.004)
Education_complete university	0.021 (0.0016)	0.0271 (0.0018)	-0.0061 (0.0024)
Chair Vote Intent_yes	0.536 (0.0056)	0.5422 (0.0056)	-0.0062 (0.0079)
Vote Count Fair_don't know	0.1462 (0.0039)	0.1526 (0.004)	-0.0064 (0.0056)
Trust Auditor General_trust a little	0.3626 (0.0054)	0.3692 (0.0054)	-0.0066 (0.0076)
Budget Prior Certainty_not certain	0.2923 (0.0051)	0.299 (0.0051)	-0.0068 (0.0072)
Trust Twaweza_trust a lot	0.343 (0.0053)	0.3499 (0.0053)	-0.0069 (0.0075)
Trust Local Politicians_trust a little	0.5162 (0.0056)	0.524 (0.0056)	-0.0078 (0.0079)
Living Conditions_better	0.3065 (0.0051)	0.3144 (0.0052)	-0.0078 (0.0073)
Voted for Councillor 2011_yes	0.7294 (0.0049)	0.7372 (0.0049)	-0.0078 (0.007)
Most Important Service_water	0.2894 (0.0051)	0.2973 (0.0051)	-0.0079 (0.0072)
Voting Secret_somewhat likely	0.1204 (0.0036)	0.1286 (0.0037)	-0.0082 (0.0052)
Vote Buying_don't know	0.0985 (0.0033)	0.1077 (0.0035)	-0.0093 (0.0048)
Party Attachment_no party	0.1825 (0.0043)	0.1922 (0.0044)	-0.0096 (0.0062)
Language_luganda	0.3478 (0.0053)	0.3577 (0.0054)	-0.0098 (0.0075)
Vote Buying_very unlikely	0.3734 (0.0054)	0.3835 (0.0054)	-0.0101 (0.0076)
Voted for Chair 2011_yes	0.721 (0.005)	0.7313 (0.005)	-0.0103 (0.007)
Language_english	0.2708 (0.0049)	0.2831 (0.005)	-0.0122 (0.0071)
Second Important Service_health	0.2619 (0.0049)	0.2742 (0.005)	-0.0123 (0.007)
Party_independent	0.1372 (0.0038)	0.1509 (0.004)	-0.0137 (0.0055)
Age	35.2681 (0.1364)	35.4751 (0.1367)	-0.0169 (0.5226)
Councillor Vote Intent_yes	0.55 (0.0055)	0.568 (0.0055)	-0.0181 (0.0078)

Notes: Results are sorted by most positively imbalanced to most negatively imbalanced. This table shows the difference between treatment and control groups in the means of all pre-treatment variables. Note that almost all variables have multiple discrete response categories (as displayed in Figure S10-11). This table takes each response category for each variable as a binary outcome for the purpose of exposition. In parentheses we show sampling-based standard errors for each mean and the difference in means, assuming the observed values are true values for each of the treatment and control group and drawn from a binomial distribution. These standard errors show likely variation in subject characteristics had the experimental sample of subjects been redrawn from the same population, rather than variation introduced by randomization draws within our sample. In order to help the reader assess the importance of each variable, we have ranked the variables by the size of the absolute difference between treatment and control groups. The following codes indicate the time that we collected the covariate: r=Recruitment; b=Baseline.

**Table S7. Distribution of voter party identification, politician party identification, and alignment by news group for main estimation.**

LCV Chair		
	Good News	Bad News
Party ID voter: NRM	2432 (68%)	2290 (77%)
Party ID voter: Independent	638 (18%)	378 (13%)
Party ID voter: FDC	360 (10%)	219 (7%)
Party ID voter: DP	114 (3%)	44 (1%)
Party ID voter: Other	34 (1%)	27 (1%)
Party ID Incumbent: NRM	2994 (76%)	2338 (75%)
Party ID Incumbent: Independent	879 (22%)	526 (17%)
Party ID Incumbent: FDC	48 (1%)	267 (9%)
Voter-Incumbent Alignment	2225 (62%)	1921 (65%)
Total Responses	3921	3131
LCV Councillor		
	Good News	Bad News
Party ID voter: NRM	1970 (70%)	2170 (80%)
Party ID voter: Independent	442 (16%)	295 (11%)
Party ID voter: FDC	292 (10%)	192 (7%)
Party ID voter: DP	91 (3%)	38 (1%)
Party ID voter: Other	27 (1%)	23 (1%)
Party ID Incumbent: NRM	2316 (75%)	2230 (77%)
Party ID Incumbent: Independent	317 (10%)	312 (11%)
Party ID Incumbent: FDC	238 (8%)	254 (9%)
Party ID Incumbent: DP	187 (6%)	56 (2%)
Party ID Incumbent: Other	16 (1%)	31 (1%)
Voter-Incumbent Alignment	1760 (55%)	1857 (63%)
Total Responses	3074	2883

Notes: This table shows how the distribution of good and bad news is related to respondents' party alignment. The results indicate that good and bad news are similarly distributed across parties. Party ID data is missing for some voters.

**Table S8. Attrition with respect to the budget treatment**

	<i>No Attrition</i>	<i>Attrition</i>
Control	6237 (0.778)	1782 (0.222)
Treatment	6344 (0.787)	1720 (0.213)

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.

**Table S9. Attrition with respect to the budget treatment and density**

	<i>No Attrition</i>	<i>Attrition</i>
Control, Low-Density	4263 (0.777)	1220 (0.223)
Control, High-Density	1101 (0.794)	285 (0.206)
Treated, Low-Density	1077 (0.782)	300 (0.218)
Treated, High-Density	4400 (0.793)	1147 (0.207)

Notes: This table shows the rates of attrition by treatment assignment and density (row-wise proportions are in parentheses). The minimal difference between treatment groups of each density suggests that treatment assignment did not differentially affect attrition within high or low density groups.

**Table S10. Treatment effects on reported turnout**

	DV: Turnout for LC5 Election			
	Good News		Bad News	
	(1)	(2)	(3)	(4)
Budget Treatment (RI)	−0.002 (0.012) p = 0.583	−0.005 (0.012) p = 0.650	−0.0001 (0.012) p = 0.499	−0.001 (0.012) p = 0.482
Turnout Intent	0.186 (0.042)	0.152 (0.037)	0.135 (0.041)	0.102 (0.037)
Polling station fixed effects	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes
Observations	6,468	6,468	5,836	5,836
Adjusted R <sup>2</sup>	0.075	0.096	0.132	0.142

Notes: SEs derived from the randomization distribution under the sharp null for treatment and clustered by politician for all other covariates; one-tailed tests for treatment matching directional hypotheses. The outcome variable equals one if the respondent indicated they voted in the LC5 election and zero if the respondent indicated that they did not vote.

**Table S11. Treatment effect on reported vote choice for the incumbent, including non-turnout as non-positive outcome**

	DV: Vote Choice for the Incumbent Including Turnout			
	LC5 Chair Good News		LC5 Chair Bad News	
	(1)	(2)	(3)	(4)
Budget Treatment (RI)	0.003 (0.016) p = 0.419	0.024 (0.017) p = 0.084	−0.001 (0.017) p = 0.484	−0.035 (0.017) p = 0.020
LC5 Chair Intent	0.081 (0.028)		0.031 (0.026)	
LC5 Councillor Intent		0.069 (0.016)		0.029 (0.017)
Polling station fixed effects	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes
Observations	5,070	3,911	3,950	3,656
Adjusted R <sup>2</sup>	0.159	0.234	0.235	0.294

Notes: SEs on treatment derived from the randomization distribution under the sharp null using exact clustered design; one-tailed tests matching directional hypotheses; contested elections only. These coefficients are estimated similarly to Table S14 except respondents who failed to turn out are not excluded from the sample, but rather coded as having not voted for the incumbent.



**Table S12. Conditional effects of budget treatment based on the party identification of the voter**

	DV: Vote Choice for the Incumbent			
	LC5 Chair	LC5 Councillor	LC5 Chair	LC5 Councillor
	Good News		Bad News	
	(1)	(2)	(3)	(4)
Budget Treatment	0.0002 (0.020) p = 0.994	0.066 (0.018) p = 0.0002	0.003 (0.021) p = 0.898	−0.043 (0.017) p = 0.015
Opposition Voter	−0.129 (0.048) p = 0.008	−0.017 (0.047) p = 0.715	−0.073 (0.082) p = 0.375	−0.058 (0.051) p = 0.254
Independent Voter	−0.048 (0.051) p = 0.349	−0.050 (0.034) p = 0.142	−0.050 (0.040) p = 0.217	−0.031 (0.042) p = 0.450
Budget Treatment * Opposition Voter	0.047 (0.082) p = 0.566	−0.124 (0.050) p = 0.014	0.059 (0.072) p = 0.410	0.059 (0.055) p = 0.287
Budget Treatment * Independent Voter	−0.010 (0.030) p = 0.730	−0.077 (0.041) p = 0.059	0.021 (0.066) p = 0.757	0.005 (0.051) p = 0.925
Village fixed effects	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes
Observations	3,578	2,822	2,958	2,718
Adjusted R <sup>2</sup>	0.223	0.333	0.352	0.401

Notes: SEs clustered by politician; two-tailed tests; contested elections only without party switching by incumbent; baseline condition is NRM-aligned (ruling party) voter. Party alignment is measured via a question in the baseline, pre-treatment survey asking "what party do you align with most closely?"

**Table S13. Conditional effects of budget treatment based on the alignment of the voter and the incumbent**

	DV: Vote Choice for the Incumbent			
	LC5 Chair Good News	LC5 Councillor	LC5 Chair Bad News	LC5 Councillor Bad News
	(1)	(2)	(3)	(4)
Budget Treatment	0.006 (0.032) p = 0.863	−0.0005 (0.024) p = 0.985	0.029 (0.037) p = 0.428	−0.056 (0.030) p = 0.063
Aligned	0.149 (0.030) p = 0.00000	0.099 (0.030) p = 0.002	0.115 (0.041) p = 0.005	0.073 (0.034) p = 0.031
Alignment Unknown	0.105 (0.055) p = 0.055	0.063 (0.049) p = 0.203	0.093 (0.053) p = 0.081	0.040 (0.049) p = 0.412
Budget Treatment * Aligned	−0.006 (0.038) p = 0.882	0.060 (0.032) p = 0.059	−0.030 (0.048) p = 0.530	0.032 (0.037) p = 0.381
Budget Treatment * Alignment Unknown	−0.022 (0.061) p = 0.718	−0.052 (0.061) p = 0.397	−0.022 (0.075) p = 0.770	0.006 (0.079) p = 0.940
Village fixed effects	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes
Observations	3,921	3,074	3,131	2,883
Adjusted R <sup>2</sup>	0.221	0.335	0.350	0.400

Notes: SEs clustered by politician; two-tailed tests; contested elections only without party switching by incumbent; baseline condition is not aligned incumbent and voter. Incumbent alignment is measured via a question in the baseline, pre-treatment survey asking "Generally speaking, what party do you identify with most?" If the party alignment of the voter matches the alignment of the incumbent, then the respondent is coded as Aligned and Not Aligned otherwise. Respondents who did not provide a party alignment are coded as "Alignment Unknown."

**Table S14. Results on vote choice for primary specification (main effects)**

	DV: Vote Choice for the Incumbent							
	LC5 Chair		LC5 Councillor		LC5 Chair		LC5 Councillor	
	Good News				Bad News			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Budget Treatment (RI)	0.002 (0.017) p=0.454	0.002 (0.017) p=0.451	0.027 (0.018) p=0.065	0.028 (0.018) p=0.061	0.011 (0.016) p=0.754	0.009 (0.016) p=0.722	-0.034 (0.016) p=0.015	-0.035 (0.016) p=0.013
LC5 Chair Intent	0.081 (0.038)	0.074 (0.033)			0.036 (0.024)	0.032 (0.023)		
LC5 Councillor Intent			0.060 (0.016)	0.052 (0.016)			0.001 (0.017)	0.004 (0.016)
Living Conditions: Don't Know		0.100 (0.089)		0.067 (0.132)		0.018 (0.053)		0.007 (0.059)
Living Conditions: Missing		-0.160 (0.235)		0.014 (0.172)		-0.130 (0.219)		0.016 (0.130)
Living Conditions: Much Better		0.094 (0.025)		0.091 (0.044)		-0.051 (0.040)		0.023 (0.043)
Living Conditions: Much Worse		-0.031 (0.026)		0.005 (0.040)		0.001 (0.033)		-0.009 (0.026)
Living Conditions: Refused		-0.112 (0.200)		0.272 (0.388)		0.328 (0.111)		-0.033 (0.156)
Living Conditions: Same		0.002 (0.018)		-0.031 (0.025)		-0.012 (0.022)		-0.021 (0.024)
Living Conditions: Worse		-0.009 (0.023)		-0.019 (0.021)		-0.029 (0.012)		0.015 (0.018)
Trust in Twaweza: Don't Know		-0.021 (0.053)		-0.026 (0.062)		0.088 (0.055)		-0.003 (0.051)
Trust in Twaweza: Missing		-0.190 (0.198)		-0.029 (0.118)		-0.004 (0.116)		-0.003 (0.070)
Trust in Twaweza: A Little		-0.021 (0.039)		0.013 (0.055)		0.083 (0.046)		-0.008 (0.037)
Trust in Twaweza: A Lot		-0.004 (0.039)		-0.001 (0.060)		0.075 (0.040)		-0.021 (0.042)
Voting Not Secret: Missing		0.176 (0.279)		-0.157 (0.158)		0.011 (0.033)		-0.127 (0.042)
Voting Not Secret: Not At All Likely		-0.002 (0.033)		-0.025 (0.039)		-0.031 (0.026)		-0.054 (0.032)
Voting Not Secret: Not Very Likely		0.011 (0.049)		0.001 (0.051)		-0.012 (0.027)		-0.059 (0.029)
Voting Not Secret: Refused		0.087 (0.194)		-0.108 (0.132)		-0.086 (0.065)		0.101 (0.058)

Voting Not Secret: Somewhat Likely		−0.032 (0.032)	−0.062 (0.050)	−0.067 (0.027)	−0.063 (0.032)			
Voting Not Secret: Very Likely		−0.002 (0.052)	−0.014 (0.043)	−0.011 (0.033)	−0.027 (0.033)			
Fair Vote Counting: Missing		−0.192 (0.204)	−0.072 (0.238)	−0.111 (0.118)	−0.363 (0.182)			
Fair Vote Counting: Not At All Likely		−0.061 (0.025)	−0.044 (0.040)	−0.034 (0.043)	0.039 (0.037)			
Fair Vote Counting: Not Very Likely		−0.037 (0.024)	−0.068 (0.040)	−0.037 (0.039)	−0.018 (0.032)			
Fair Vote Counting: Refused		0.037 (0.076)	−0.002 (0.081)	−0.017 (0.094)	0.092 (0.077)			
Fair Vote Counting: Somewhat Likely		−0.031 (0.016)	−0.065 (0.040)	0.031 (0.032)	0.009 (0.027)			
Fair Vote Counting: Very Likely		−0.026 (0.032)	0.009 (0.038)	0.028 (0.026)	0.002 (0.028)			
Male		−0.055 (0.033)	−0.059 (0.020)	−0.034 (0.018)	−0.041 (0.021)			
Education: Some Primary		−0.051 (0.031)	0.001 (0.042)	−0.009 (0.051)	−0.090 (0.030)			
Education: Completed Primary		−0.035 (0.037)	−0.034 (0.039)	0.007 (0.054)	−0.034 (0.030)			
Education: Some Secondary		−0.081 (0.031)	−0.052 (0.040)	−0.008 (0.049)	−0.074 (0.032)			
Education: Completed Secondary		−0.042 (0.046)	−0.027 (0.054)	−0.014 (0.057)	−0.064 (0.041)			
Education: Some University		−0.126 (0.062)	−0.078 (0.061)	−0.070 (0.063)	−0.130 (0.055)			
Education: Completed University		−0.087 (0.088)	−0.061 (0.070)	0.006 (0.061)	−0.194 (0.072)			
Education: Some Post-Graduate		−0.057 (0.096)	−0.136 (0.061)	0.096 (0.198)	0.057 (0.077)			
Education: Completed Masters		−0.252 (0.050)	−0.210 (0.040)	0.192 (0.148)	0.016 (0.044)			
Education: Refused		−0.755 (0.044)		0.176 (0.044)	0.362 (0.046)			
Age		0.002 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)			
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,921	3,921	3,074	3,074	3,131	3,131	2,883	2,883
Adjusted R <sup>2</sup>	0.199	0.205	0.317	0.323	0.341	0.345	0.392	0.396

---

Notes: SEs derived from the randomization distribution under the sharp null for treatment and clustered by politician for all other covariates; one-tailed tests on treatment matching directional hypotheses; contested elections where incumbent did not switch parties only. This table shows coefficient estimates from which we derive Figure 1 panels A and C in the main draft.

**Table S15. Results on vote choice for primary specification (verified recipient effects)**

	DV: Vote Choice for the Incumbent							
	LC5 Chair		LC5 Councillor		LC5 Chair		LC5 Councillor	
	Good News				Bad News			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Budget Treatment (RI)	0.004 (0.022) p=0.424	0.006 (0.022) p=0.391	0.048 (0.023) p=0.021	0.051 (0.024) p=0.017	-0.007 (0.019) p=0.349	-0.009 (0.019) p=0.313	-0.055 (0.020) p=0.003	-0.056 (0.020) p=0.003
LC5 Chair Intent	0.092 (0.045)	0.086 (0.037)			0.033 (0.023)	0.031 (0.023)		
LC5 Councillor Intent			0.076 (0.023)	0.066 (0.023)			-0.006 (0.021)	-0.003 (0.021)
Living Conditions: Don't Know		-0.069 (0.123)		-0.006 (0.241)		-0.061 (0.072)		0.018 (0.073)
Living Conditions: Missing		-0.284 (0.218)		-0.023 (0.201)		-0.174 (0.194)		0.035 (0.063)
Living Conditions: Much Better		0.038 (0.077)		0.071 (0.068)		-0.061 (0.056)		0.023 (0.046)
Living Conditions: Much Worse		-0.022 (0.039)		0.003 (0.055)		0.025 (0.034)		-0.011 (0.032)
Living Conditions: Refused		-0.142 (0.222)		0.312 (0.498)				0.070 (0.036)
Living Conditions: Same		-0.016 (0.031)		-0.031 (0.033)		-0.035 (0.025)		-0.038 (0.026)
Living Conditions: Worse		-0.014 (0.014)		-0.019 (0.027)		-0.037 (0.018)		0.007 (0.026)
Trust in Twaweza: Don't Know		-0.009 (0.081)		0.042 (0.088)		0.068 (0.062)		-0.014 (0.046)
Trust in Twaweza: Missing		-0.218 (0.208)		-0.017 (0.167)		-0.145 (0.174)		0.022 (0.069)
Trust in Twaweza: A Little		-0.024 (0.062)		0.046 (0.084)		0.072 (0.052)		-0.020 (0.040)
Trust in Twaweza: A Lot		-0.021 (0.066)		0.061 (0.092)		0.052 (0.048)		-0.016 (0.042)
Voting Not Secret: Missing		0.468 (0.049)		-0.333 (0.059)		-0.065 (0.048)		-0.057 (0.071)
Voting Not Secret: Not At All Likely		-0.011 (0.033)		-0.055 (0.053)		-0.046 (0.038)		-0.060 (0.037)
Voting Not Secret: Not Very Likely		0.011 (0.042)		-0.022 (0.065)		-0.021 (0.034)		-0.072 (0.034)
Voting Not Secret: Refused		-0.014 (0.216)		-0.221 (0.174)		-0.166 (0.097)		0.133 (0.088)



Voting Not Secret: Somewhat Likely	−0.035 (0.052)	−0.053 (0.076)	−0.087 (0.036)	−0.072 (0.040)
Voting Not Secret: Very Likely	0.020 (0.054)	0.018 (0.058)	−0.043 (0.042)	−0.008 (0.038)
Fair Vote Counting: Missing	−0.260 (0.034)		−0.074 (0.120)	−0.280 (0.152)
Fair Vote Counting: Not At All Likely	−0.066 (0.032)	−0.027 (0.058)	−0.015 (0.046)	0.045 (0.046)
Fair Vote Counting: Not Very Likely	−0.075 (0.031)	−0.080 (0.054)	−0.080 (0.054)	−0.025 (0.045)
Fair Vote Counting: Refused	0.026 (0.078)	−0.039 (0.104)	−0.007 (0.132)	0.091 (0.101)
Fair Vote Counting: Somewhat Likely	−0.044 (0.019)	−0.055 (0.058)	0.003 (0.039)	−0.002 (0.041)
Fair Vote Counting: Very Likely	−0.043 (0.038)	0.005 (0.058)	0.005 (0.036)	0.001 (0.046)
Male	−0.051 (0.030)	−0.043 (0.025)	−0.016 (0.017)	−0.038 (0.024)
Education: Some Primary	−0.080 (0.043)	−0.025 (0.056)	0.039 (0.073)	−0.022 (0.052)
Education: Completed Primary	−0.053 (0.047)	−0.064 (0.056)	0.037 (0.072)	0.010 (0.053)
Education: Some Secondary	−0.116 (0.039)	−0.083 (0.059)	0.016 (0.070)	−0.029 (0.053)
Education: Completed Secondary	−0.066 (0.071)	−0.054 (0.077)	0.005 (0.067)	0.015 (0.056)
Education: Some University	−0.115 (0.073)	−0.164 (0.078)	−0.043 (0.093)	−0.076 (0.084)
Education: Completed University	−0.090 (0.115)	−0.085 (0.086)	0.041 (0.079)	−0.144 (0.088)
Education: Some Post-Graduate	−0.245 (0.186)	−0.281 (0.092)	0.052 (0.304)	0.053 (0.126)
Education: Completed Masters	−0.249 (0.043)	−0.302 (0.062)	−0.0002 (0.065)	0.062 (0.079)
Education: Refused	−0.806 (0.055)		0.216 (0.070)	0.422 (0.067)
Age	0.002 (0.001)	0.001 (0.001)	0.002 (0.001)	0.001 (0.001)

Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,476	2,476	1,925	1,925	2,205	2,205	2,018	2,018

Adjusted R <sup>2</sup>	0.201	0.206	0.317	0.320	0.327	0.331	0.416	0.417
-------------------------	-------	-------	-------	-------	-------	-------	-------	-------

Notes: SEs derived from the randomization distribution under the sharp null for treatment and clustered by politician for all other covariates; one-tailed tests on treatment matching directional hypotheses; contested elections where incumbent did not switch parties only. This table shows coefficient estimates from which we derive Figure 1 panels B and D in the main draft.

**Table S16. Effect of budget treatment on votes for district (LC5) incumbents, estimated as pre-specified**

	DV: Vote Choice for the Incumbent							
	LC5 Chair		LC5 Councillor		LC5 Chair		LC5 Councillor	
	Good News				Bad News			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Budget Treatment (RI)	0.005 (0.024) p = 0.421	0.003 (0.017) p = 0.439	0.006 (0.027) p = 0.411	0.028 (0.018) p = 0.061	−0.019 (0.024) p = 0.216	0.009 (0.016) p = 0.714	−0.042 (0.025) p = 0.051	−0.036 (0.016) p = 0.012
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Observations	3,921	3,921	3,074	3,074	3,131	3,131	2,883	2,883
Adjusted R <sup>2</sup>	0.092	0.564	0.181	0.635	0.265	0.688	0.255	0.691

Notes: SEs derived from the randomization distribution under the sharp null for treatment and clustered by politician for all other covariates; one-tailed tests on treatment matching directional hypotheses; contested elections where incumbent did not switch parties only. Data subset includes constituencies where the district incumbent is not running again, but excludes constituencies that were redistricted or where incumbents switched parties. Models 1,3,5, and 7 are without covariates, while models 2,4,6, and 8 include the following covariates: perception of living conditions, gender, education, age, trust in information from Twaweza, perception that powerful people will learn about vote choice, perception that vote counting will be fair, voted for incumbent in 2011 election.

**Table S17. Effect of budget treatment and treatment density on vote choice for district (LCV) incumbents, estimated as pre-specified**

	DV: Vote Choice for the Incumbent			
	LC5 Chair		LC5 Councillor	
	Good News		Bad News	
	(1)	(2)	(3)	(4)
Control, High Density (RI)	0.004 (0.029) p = 0.455	−0.028 (0.050) p = 0.718	−0.035 (0.029) p = 0.114	−0.044 (0.055) p = 0.207
Treated, Low Density (RI)	0.023 (0.026) p = 0.187	0.012 (0.030) p = 0.349	−0.028 (0.025) p = 0.138	−0.054 (0.028) p = 0.028
Treated, High Density (RI)	−0.010 (0.021) p = 0.684	−0.020 (0.045) p = 0.669	−0.030 (0.021) p = 0.078	−0.091 (0.050) p = 0.035
Observations	3,281	2,585	2,633	2,391
Adjusted R <sup>2</sup>	0.570	0.596	0.668	0.624

Notes: SEs derived from RI; one-tailed tests matching directional hypotheses; contested elections only. The outcome variable is the difference between the vote for the incumbent and the intention to vote for the incumbent. Data subset includes constituencies where the district incumbent is not running again, but excludes constituencies that were redistricted or where incumbents switched parties. Models include the following covariates: perception of living conditions, gender, education, age, trust in information from Twaweza, perception that powerful people will learn about vote choice, perception that vote counting will be fair, voted for incumbent in 2011 election.

**Table S18. Conditional effects of budget treatment on votes for incumbents based on party switching**

	DV: Vote Choice for the Incumbent							
	LC5 Chair		LC5 Councillor		LC5 Chair		LC5 Councillor	
	Good News				Bad News			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Budget Treatment	0.002 (0.018) p = 0.456	0.003 (0.017) p = 0.439	0.025 (0.016) p = 0.059	0.027 (0.016) p = 0.054	0.011 (0.016) p = 0.744	0.009 (0.017) p = 0.714	−0.031 (0.014) p = 0.013	−0.032 (0.014) p = 0.009
LC5 Chair Intent	0.072 (0.035)	0.066 (0.030)			0.035 (0.022)	0.032 (0.022)		
Budget Treatment X Chair Party Switch	0.027 (0.016)	0.032 (0.017)			0.016 (0.016)	0.030 (0.018)		
LC5 Councillor Intent			0.055 (0.015)	0.047 (0.015)			0.006 (0.016)	0.008 (0.015)
Budget Treatment X Councillor Party Switch			−0.098 (0.040)	−0.091 (0.043)			0.058 (0.036)	0.057 (0.036)
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Observations	4,398	4,398	3,545	3,545	3,333	3,333	3,687	3,687
Adjusted R <sup>2</sup>	0.243	0.248	0.350	0.355	0.342	0.346	0.482	0.484

Notes: SEs clustered by politician; one-tailed tests matching directional hypotheses; contested elections only. In this table, we include all voters that reported vote choice, including voters in elections with party switching or redistricting. Note that for elections for councillors, party switching incumbents have the opposite response to good and bad news as reported in the main subset, suggesting that voters attribute performance information to parties on average.

**Table S19. Effects of budget treatment on votes for individual incumbents who ran for re-election as members of the same party**

	DV: Vote Choice for the Incumbent							
	LC5 Chair		LC5 Councillor		LC5 Chair		LC5 Councillor	
	Good News				Bad News			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Budget Treatment	0.0004 (0.019) p = 0.491	0.001 (0.018) p = 0.469	0.035 (0.026) p = 0.090	0.041 (0.028) p = 0.071	0.040 (0.027) p = 0.933	0.040 (0.027) p = 0.934	−0.026 (0.019) p = 0.080	−0.022 (0.019) p = 0.130
LC5 Chair Intent	0.089 (0.042)	0.083 (0.037)			0.074 (0.019)	0.077 (0.019)		
LC5 Councillor Intent			0.083 (0.025)	0.081 (0.025)			−0.013 (0.024)	−0.009 (0.022)
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Observations	3,520	3,520	1,331	1,331	1,248	1,248	1,461	1,461
Adjusted R <sup>2</sup>	0.204	0.210	0.356	0.368	0.361	0.363	0.340	0.342

Notes: SEs clustered by politician; one-tailed tests matching directional hypotheses; contested elections only. In this table, only individual incumbent politicians who are confirmed as running for re-election as a member of the same party are considered. We exclude elections where the incumbent ran again in 2016, but as a member of a different party from the one she/he ran under in 2011.

**Table S20. Effects of budget treatment on votes for district incumbents, with incumbency defined as the party previously elected to the seat**

	DV: Vote Choice for the Incumbent							
	LC5 Chair		LC5 Councillor		LC5 Chair		LC5 Councillor	
	Good News				Bad News			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Budget Treatment	0.001 (0.016) p = 0.467	0.0004 (0.016) p = 0.490	0.029 (0.016) p = 0.030	0.030 (0.016) p = 0.031	0.008 (0.016) p = 0.691	0.007 (0.016) p = 0.683	−0.033 (0.013) p = 0.006	−0.034 (0.013) p = 0.004
LC5 Chair Intent	0.072 (0.034)	0.064 (0.030)			0.034 (0.023)	0.031 (0.022)		
LC5 Councillor Intent			0.054 (0.014)	0.045 (0.015)			0.0005 (0.016)	0.002 (0.016)
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Observations	4,398	4,398	3,517	3,517	3,333	3,333	3,621	3,621
Adjusted R <sup>2</sup>	0.180	0.186	0.320	0.325	0.416	0.418	0.389	0.392

Notes: SEs clustered by politician; one-tailed tests matching directional hypotheses; contested elections only. In this table a vote for the incumbent is coded if a respondent indicated voting for the party previously elected to the council seat, regardless of whether the candidate running under that party is the same person that won the 2011 election. This would be the case, for instance, if an incumbent lost a primary or retired.

**Table S21. Effects of budget treatment on votes for district incumbents, with incumbency defined by individual if the individual is running for re-election**

	DV: Vote Choice for the Incumbent							
	LC5 Chair		LC5 Councillor		LC5 Chair		LC5 Councillor	
	Good News				Bad News			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Budget Treatment	0.005 (0.016) p = 0.379	0.006 (0.016) p = 0.353	0.013 (0.015) p = 0.188	0.016 (0.015) p = 0.154	0.012 (0.016) p = 0.777	0.011 (0.016) p = 0.764	−0.018 (0.013) p = 0.076	−0.020 (0.013) p = 0.059
LC5 Chair Intent	0.072 (0.035)	0.066 (0.031)			0.035 (0.022)	0.032 (0.022)		
LC5 Councillor Intent			0.056 (0.015)	0.047 (0.015)			0.006 (0.016)	0.008 (0.015)
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Observations	4,398	4,398	3,545	3,545	3,333	3,333	3,687	3,687
Adjusted R <sup>2</sup>	0.243	0.248	0.350	0.355	0.343	0.346	0.482	0.483

Notes: SEs clustered by politician; one-tailed tests matching directional hypotheses; contested elections only. In this table a vote for the incumbent is coded if the respondent indicated voting for the individual who won in 2011, regardless of whether that individual is contesting the council seat under the same party banner. This occurs, for instance, when incumbents lose a primary and run as independents.

**Table S22. Example NGO Campaigns Using SMS.**

Organization, Country	Brief Description	Further Reading
Sisi Ni Amani, Kenya	Sent SMS messages to around 60,000 subscribers targeted by individual characteristics (e.g., gender, region) to encourage peaceful electoral participation	Neelam Verjee, "Will Kenya's elections transform the text message from deadly weapon to peace offering?", <i>Quartz</i> , March 2, 2013.
Kubatana, Zimbabwe	Seeks to provide credible, nonpartisan information about politics to the public. Uses a variety of ICT (e.g., WhatsApp, Facebook, etc.) but the biggest number of users come from SMS messages.	<a href="http://kubatana.net/">http://kubatana.net/</a>
Advocates Coalition for Development and Environment (ACODE), Uganda	Sent information from their Local Government Councils' Scorecard Initiative (which measures local politician performance) via SMS messages to citizens.	<a href="http://www.acode-u.org/LGCSCI.html">http://www.acode-u.org/LGCSCI.html</a>
UNICEF U-Report, Uganda (and elsewhere)	Millions of individuals, known as "U-Reporters," have signed up for this free service, and are periodically asked questions about politics through text message.	<a href="https://www.centreforpublicimpact.org/case-study/unicef-ureport">https://www.centreforpublicimpact.org/case-study/unicef-ureport</a>
African Elections Project, Africa	Allows users to send information of voter suppression, violence, etc. via SMS and also sends election results, voter information, etc. to users via SMS.	<a href="http://www.africanelections.org/#">http://www.africanelections.org/#</a>

Notes: This table shows a sample of NGO campaigns that have utilized SMS messages as part of a political information campaign.

**Table S23. Manipulation checks for updating in beliefs about budget performance of LC5**

	(OLS)	(OLS FE)	(OLS FE Clustered)
Perfect Updating	0.007 (0.007) p=0.180	0.004 (0.008) p=0.328	0.004 (0.008) p=0.325
Updating	0.011 (0.009) p=0.104	0.009 (0.010) p=0.183	0.009 (0.010) p=0.161
Loose Updating	0.015 (0.009) p=0.041	0.020 (0.010) p=0.024	0.020 (0.008) p=0.006
Loose Updating (Eligible)	0.016 (0.009) p=0.042	0.021 (0.011) p=0.025	0.021 (0.008) p=0.004
Directional Updating	0.012 (0.009) p=0.086	0.010 (0.010) p=0.164	0.010 (0.010) p=0.158

Notes: This table shows increased rates of different kinds of updating in the treatment group as compared to the control group. Perfect updating is having correct posterior beliefs in relation to the treatment information. 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. Directional updating is moving posterior beliefs in the direction of the treatment information or having correct prior and posterior beliefs. Model specifications are: (1) updating indicator by treatment OLS; (2) updating indicator by treatment OLS with village fixed effects; and (3) updating indicator by treatment OLS with village fixed effects and district-level clustering.



**Table S24. Multiple, joint, and pooled hypothesis tests using randomization inference**

Test	Main Subset		Verified Recipients	
	No Covariates	Covariates	No Covariates	Covariates
$\max(\tilde{b}_+) > \max(b_+^*)$	0.1105	0.1031	0.0329	0.0268
$\min(\tilde{b}_-) < \min(b_-^*)$	0.0301	0.0248	0.0037	0.0035
$\tilde{b}_{\text{chair},+} > \max(b_{\text{chair}}^*) \cap \tilde{b}_{\text{chair},-} < \min(b_{\text{chair}}^*)$	0.3433	0.3269	0.1476	0.1210
$\tilde{b}_{\text{councillor},+} > \max(b_{\text{councillor}}^*) \cap \tilde{b}_{\text{councillor},-} < \min(b_{\text{councillor}}^*)$	0.0007	0.0006	0.0000	0.0000
$\max(\tilde{b}_+) > \max(b_+^*) \cap \min(\tilde{b}_-) < \min(b_-^*)$	0.0031	0.0024	0.0000	0.0000
$\tilde{b}_{\text{pooled},+} > b_{\text{pooled},+}^*$	0.1737	0.1742		
$\tilde{b}_{\text{pooled},-} < b_{\text{pooled},-}^*$	0.1848	0.1575		
$\tilde{b}_{\text{pooled},+} > b_{\text{pooled},+}^* \cap \tilde{b}_{\text{pooled},-} < b_{\text{pooled},-}^*$	0.0341	0.0303		

Notes: All values in this table are  $p$ -values from the noted statistical test, computing using 10,000 draws from the exact randomization procedure used to assign treatment. The verified recipients analysis is not reported for the pooled tests because the goal of those tests is to examine the results among the broadest possible pool of subjects.

**Table S25. Main results (Figure 1) with fixed-effects weights applied**

	LC5 Chair Good News		LC5 Councillor Good News		LC5 Chair Bad News		LC5 Councillor Bad News	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DV: Vote Choice for the Incumbent, Main Subset								
Budget Treatment (RI)	0.002 (0.017) p=0.460	0.001 (0.017) p=0.467	0.022 (0.018) p=0.107	0.023 (0.018) p=0.101	0.006 (0.016) p=0.648	0.005 (0.016) p=0.615	-0.036 (0.016) p=0.011	-0.037 (0.016) p=0.010
DV: Vote Choice for the Incumbent, Verified Recipients								
Budget Treatment (RI)	0.005 (0.022) p=0.413	0.006 (0.022) p=0.388	0.044 (0.023) p=0.031	0.047 (0.024) p=0.024	-0.012 (0.019) p=0.262	-0.014 (0.019) p=0.224	-0.056 (0.020) p=0.002	-0.056 (0.020) p=0.002
Fixed effect weights	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	No	Yes	No	Yes

Notes: SEs reported are the standard deviation of the randomization distribution under the sharp null hypothesis. One-tailed tests matching directional hypotheses; contested elections only with no party switching or redistricting. Sample sizes are equivalent to those reported in Figure 1.