ABLE AND MOSTLY WILLING: AN EMPIRICAL ANATOMY OF INFORMATION'S EFFECT ON VOTER-DRIVEN ACCOUNTABILITY IN SENEGAL*

ABHIT BHANDARI† HORACIO LARREGUY‡ JOHN MARSHALL§

JUNE 2020

Political accountability may be constrained by the reach and relevance of information campaigns in developing democracies and—upon receiving information—voters' ability and will to hold politicians accountable. To illuminate voter-level constraints and information relevance absent dissemination constraints, we conducted a field experiment around Senegal's 2017 parliamentary elections to examine the core theoretical steps linking receiving different types of incumbent performance information to electoral and non-electoral accountability. Voters immediately processed information as Bayesians, found temporally benchmarked local performance outcomes particularly informative, and updated their beliefs for at least a month. Learning that incumbents generally performed better than expected, voters durably requested greater politician contact after elections while incumbent vote choice increased among likely-voters and voters prioritizing local projects when appraising incumbents. In contrast, information about incumbent duties did not systematically influence beliefs or accountability. These findings suggest voters were able and mostly willing to use relevant information to hold politicians to account.

*We thank Fodé Sarr and his team of enumerators for invaluable research assistance, and Elimane Kane, Thierno Niang, and LEGS-Africa for partnering with us to implement this project. We thank Antonella Bandiera, Nilesh Fernando, Matthew Gichohi, Lakshmi Iyer, Kate Orkin, Julia Payson, Amanda Robinson, Arturas Rozenas, Cyrus Samii, Moses Shayo, Jay Shon, Alberto Simpser, and participants at talks at APSA, MPSA, Notre Dame, NYU, NYU CESS Experimental Political Science Conference, and WGAPE-NYU Abu Dhabi for excellent comments. This project received financial support from the Spencer Foundation, and was approved by the Columbia Institutional Review Board (IRB-AAAR3724) and the Harvard Committee on the Use of Human Subjects (IRB17-0880). Our pre-analysis plan was registered with the Social Science Registry, and is available at socialscienceregistry.org/trials/2324.

†Department of Political Science, Columbia University. Email: ab3901@columbia.edu.

‡Department of Government, Harvard University. Email: hlarreguy@fas.harvard.edu.

§Department of Political Science, Columbia University. Email: jm4401@columbia.edu.

1 Introduction

Informing voters about their incumbent's performance in office is thought to help citizens retain high-quality politicians (Fearon 1999) and hold politicians to account beyond the ballot box (Aker, Collier and Vicente 2017; Gottlieb 2016). In practice, however, recent studies identifying the effects of informational campaigns on electoral accountability (Banerjee et al. 2011; Chong et al. 2015; Cruz, Keefer and Labonne 2019; Dunning et al. 2019; Ferraz and Finan 2008; Humphreys and Weinstein 2012) and non-electoral political engagement (see Casey 2018; Lieberman, Posner and Tsai 2014) yield mixed findings.

Given the complex chain of conditions linking the provision of information to better governance (Dunning et al. 2019; Lieberman, Posner and Tsai 2014), it is hard to know where voter-politician accountability breaks down. For example, the limited effects of providing incumbent performance information on vote choice in the six-country Metaketa study (Dunning et al. 2019) could reflect difficulties of disseminating information, failures to provide sufficiently relevant information, voters' inability to internalize information, or voters' low willingness to electorally reward (punish) better(worse)-performing incumbents. Furthermore, while community empowerment interventions have received significant attention (Casey 2018), little is known about whether incumbent performance information can influence non-electoral accountability. A particularly important non-electoral means of communicating information, preferences, or requests is citizen contact with incumbent politicians *once in office* (Bussell 2019; Grossman, Humphreys and Sacramone-Lutz 2014).[1]

This article dissects voters' ability and will to use different types of incumbent performance information to hold legislative deputies to account. By personally distributing and explaining such information, we abstract from dissemination challenges to focus on three links between receiving incumbent performance information and voter engagement in electoral and non-electoral account ability. First, we illuminate voter internalization of information and its decay over time by ex-

[1]Such efforts could involve articulating programmatic demands or seeking pork.

amining the extent to which voters update their beliefs in a Bayesian manner, both immediately after receiving information and a month later. Second, we vary the information's content to understand what information voters regard as relevant. Specifically, we combine indicators of the current incumbent's national and local performance with: (i) information about deputy duties; and (ii) a temporal benchmark against previous incumbents' performance that helps voters to abstract from district-specific factors affecting every incumbent's performance. Third, we study whether persistent changes in beliefs translate into greater electoral support for, and greater post-election effort to request contact with, better-performing incumbents, and how this behavior varies with the information's relevance to voters.

Together with a local civil association, we designed a field experiment in Senegal around the 2017 parliamentary elections to examine these voter-level mechanisms underpinning political accountability among deputies seeking re-election for a second term. Across 450 rural villages from five of Senegal's 45 districts, we trained enumerators to personally distribute and explain informational leaflets to voters aged 20-38 in treated villages in the month preceding the election. Our factorial design varied whether respondents were informed about: (1) parliamentary deputies' duties; and (2) their current deputy's participation in legislative affairs and the projects and transfers received by their district, either with or without a comparison with their district's previous deputy. Our panel survey tracked voters' beliefs, vote intentions and ultimate choices, and post-election contact requests of incumbents immediately before and after treatment and again after the election.

Our findings first demonstrate that rural Senegalese citizens processed incumbent performance information in sophisticated ways. Immediately after receiving the information, voters favorably updated their beliefs in line with their relatively pessimistic prior beliefs and the fact that current incumbents mostly outperformed previous deputies. These changes in beliefs indicate that voters care principally about local outcomes (projects and transfers), rather than legislative efforts within parliament. Moreover, while information about deputy duties did not affect beliefs, temporally benchmarked information further improved voter appraisals of the incumbent and increased the

precision of such beliefs. We find similar—albeit somewhat smaller—effects of local performance indicators and temporal benchmarks on beliefs around a month after treatments were administered.

Immediately after receiving incumbent performance information, voters also sought electoral and non-electoral means of holding politicians to account. The average treated voter—who updated more favorably about current incumbents than challengers—became three percentage points more likely to intend to vote for the incumbent. Heterogeneity in such electoral rewards reflected the degree of voter belief updating and whether performance information was the most important factor determining vote choices. Treated voters also became significantly more likely to request a visit from, or an opportunity to express their views or demands to, winning incumbent deputies after the election.

While voters persistently updated their beliefs and demonstrated an initial willingness to hold politicians to account, electoral accountability ultimately only occurred among likely-voters who most valued performance on local outcomes. While our treatments did not affect self-reported vote choices *on average*, the treated respondents that cared most about incumbents lobbying for local development projects or had turned out at the last election did reward the incumbents overseeing more local projects and transfers. Consistent with substantial within-village diffusion of our information—by voters and political parties—to the more experienced voters most likely to respond to it, we further find greater incumbent vote shares at polling stations that received information revealing higher rates of local projects and transfers.

Non-electoral requests for incumbent contact after the election increased more uniformly, even a month after receiving treatment. The average respondent continued to make more requests of incumbents, who won in each race, especially in districts that received more projects and transfers. This increase in requests reflects not only relatively costless requests for the winning incumbents to call respondents or visit their village, but also citizens incurring the cost of sending SMS or voicemail messages to winning incumbents. This effect was also most pronounced among voters that received benchmarked information.

Our core finding that receiving relevant forms of incumbent performance information can in-

duce persistent voter belief updating and facilitate electoral and non-electoral accountability makes several main contributions. First, by unpacking the key links in the accountability chain *once voters receive relevant information*, we show that accountability failures are unlikely to reflect cognitive constraints—voters' inability to process information in a Bayesian manner (Gomez and Wilson 2006) or retain updated beliefs (Zaller 1992)—or voter unwillingness to hold politicians to account. In this regard, voters' sophisticated responses—even in a hyper-presidential context where almost half our respondents lacked any formal schooling—chime with Arias et al. (2018*b*), Humphreys and Weinstein (2012), and Kendall, Nannicini and Trebbi (2014). We advance this literature by showing that voters' initial belief updating persisted for the most relevant pieces of information, but only translated into electoral accountability among experienced voters and voters that care about the topics about which information was provided. This suggests that less effective information dissemination campaigns may instead reflect limited internalization (e.g. Dunning et al. 2019), a lack of relevant or credible information (e.g. Boas, Hidalgo and Melo 2019), or competing community or political responses to information campaigns (e.g. Banerjee et al. 2011; Cruz, Keefer and Labonne 2019).

Second, we illuminate the *types* of information that can facilitate "bottom-up" political accountability. We demonstrate that voters find temporal benchmarking against previous incumbents more relevant than information solely about current incumbents. This finding contrasts with studies that report no additional effect of combining cross-sectional comparisons alongside incumbent performance information in other developing contexts (Arias et al. 2018*a*; Campello and Zucco 2016). These contrasting findings suggest that future research, possibly comparing spatial and temporal benchmarks within the same experiment, is required to identify when different types of benchmark are most relevant. Furthermore, we find that Senegalese voters prioritize politicians bringing projects and higher-value transfers to their district. Conversely, greater involvement in parliamentary activities is—if anything—punished by voters (see also Adida et al. forthcoming; Humphreys and Weinstein 2012). However, we find little evidence that information about incumbent responsibilities influences voter appraisals on its own, or that it systematically substitutes or complements

the provision of performance information. This suggests that any accountability-enhancing effects of civic education programs (e.g. Gottlieb 2016) may operate through components of the program beyond information about incumbent responsibilities.

Finally, we show that incumbent performance information also influences a costly *non-electoral* means through which voters can seek political accountability, likely by altering expectations that politicians will be responsive to their constituents. This finding complements evidence that civic education and communication technologies can stimulate non-electoral political engagement (Aker, Collier and Vicente 2017; Gottlieb 2016; Grossman, Humphreys and Sacramone-Lutz 2014). Future research is required to establish whether voters' greater efforts to contact the best-performing legislators reflect particularistic desires for "pork" or—perhaps less likely—a desire for more programmatic policies.

2 Incumbent performance information and bottom-up political accountability

The canonical selection model of electoral accountability reflects the agency relationship between voters and politicians. In its simplest formulation, voters use performance indicators to identify high-quality incumbents, and then vote to retain them (Fearon 1999). This framework predicts that information which favorably updates voter beliefs about incumbent quality, relative to challenger quality, increases support for the incumbent, especially among voters for whom politician quality outweighs other factors entering their voting calculus. After elections, voters may similarly become more willing to engage in costly efforts to contact incumbents that they expect will be responsive and effective. Appendix section A formally summarizes these logics.

By providing information about incumbent performance—that incumbents could not have anticipated would be publicized when deciding how to act (c.f. Grossman and Michelitch 2018)—just before elections in which incumbents sought re-election, our design sidesteps strategic policy and candidacy choices. Furthermore, by directly providing information to voters, we also abstract from

potential failures in the process through which information is disseminated and consumed.

Whether receiving credible incumbent performance information causes voters to hold incumbents seeking re-election to account electorally and non-electorally thus rests upon: (i) voters' cognitive capacity to process and internalize novel information; (ii) information's relevance to voters; and (iii) voters' willingness to act on their updated beliefs about the incumbent's quality. The following subsections theorize key conditions under which each element of this anatomy of political accountability may hold.

2.1 Internalization of novel information

Political accountability relies on voters comprehending incumbent performance information and somewhat durably updating their beliefs about incumbent quality. Since the information that voters read, hear, and observe is often complex, they may only superficially understand it and therefore not meaningfully update their beliefs (Gomez and Wilson 2006). Moreover, because voters may reject novel information challenging their pre-existing beliefs or fail to internalize it over time (Zaller 1992), belief updating about incumbent quality may be too transient to influence voting behavior (Humphreys and Weinstein 2012).

Credible and comprehensible information is most likely to alter behavior when it differs from voters' prior beliefs about their incumbent's quality. Bayesian voters update the position and precision of their posterior beliefs most when the information provided is precise, the information deviates from their prior beliefs, and their prior beliefs are imprecise (e.g. Arias et al. 2018*b*; Kendall, Nannicini and Trebbi 2014).[2] Following prior research, we thus hypothesize that:

H1. *Incumbent performance information will increase (decrease) incumbent support and requests to the extent that such information causes voters to favorably (unfavorably) update about incumbent quality.*

[2]Voters may also update about challengers from incumbent performance, if their types are correlated.

2.2 Relevance of novel information

Even if novel incumbent performance information is credible and internalized, voters must also perceive it as relevant—that is, pertaining to incumbent quality—to influence political accountability. We study two aspects of relevance that could complement the provision of incumbent performance indicators: information about incumbent duties, and temporal performance benchmarks.

Information about incumbent duties could help voters to infer incumbent quality from performance signals in at least two ways. First, voters may only recognize performance information as relevant upon learning that politicians possess the capacity to feasibly influence such performance indicators (Gottlieb 2016). Second, specific information about an incumbent's duties may help voters to assign responsibility across multiple layers of government (Powell and Whitten 1993). While incumbent duties are often implicit when performance information is provided, or outlined alongside performance information (Gottlieb 2016), we explicitly separate between providing information about duties and performance to test whether:

H2. *Receiving information about incumbent duties, either alongside or without corresponding incumbent performance information, increases (decreases) incumbent support and requests among better(worse)-performing incumbents.*

Benchmarked incumbent performance information could increase the accuracy of voters' posterior beliefs through two main channels. First, receiving multiple performance signals helps voters to filter out common shocks influencing the performance of all agents in a given period or location (Aytaç 2018; Meyer and Vickers 1997). Second, benchmarks might enable voters to update about the absolute quality level of other politicians that resemble challengers, especially where benchmarks are from a different political party from the incumbent. Both channels facilitate more accurate and precise beliefs about *absolute* and *relative* incumbent, and possibly challenger, candidate quality—the key drivers of political accountability in our conceptual framework.

The relative utility of cross-sectional and inter-temporal benchmarks in a particular context

depends on the accuracy and uncertainty of voters' prior beliefs about time- and unit-specific shocks and their magnitude. Thus far, extant studies focusing on *spatial* benchmarks—that help filter out period-specific shocks that equally affect all incumbents holding office in different districts contemporaneously (e.g. changing national budgets)—generally find limited evidence that such benchmarks influence beliefs beyond providing information about only the incumbent's performance (Arias et al. 2018*a*; Aytaç 2018; Campello and Zucco 2016).[3] We focus on *temporal* benchmarks, which have yet to be tested experimentally. Temporal benchmarks help filter out the effects of time-invariant features of a district that affect all incumbents serving that district (e.g. geographical constraints or demographic political importance). We hypothesize that:

H3. *Relative to only providing incumbent performance information, temporal benchmarks increase (decrease) incumbent support and requests when: (i) incumbent performance is above (below) voters' prior belief; and/or (ii) the previous incumbent's performance was below (above) voters' prior belief.*

Appendix section B demonstrates formally that case (i) reflects benchmarked information facilitating more precise inferences, while case (ii) reflects benchmarked information updating beliefs about district-specific characteristics influencing performance.

2.3 Acting on internalized beliefs

Even if information meaningfully updates voters' beliefs, bottom-up political accountability requires that voters ultimately act on such beliefs. This likely requires that several conditions hold. First, voters must connect their beliefs about the incumbent to their available actions (Gomez and Wilson 2006). Second, in the case of voting, voters must attach significant weight to beliefs about incumbent quality in their voting calculus. Third, the process of providing information could set in motion other forces that override the influence of voter beliefs on vote choice, including voter coordination around particular candidates (Arias et al. 2019) or candidate campaign responses to

[3]The clearest electoral evidence comes from cross-national macroeconomic comparisons in advanced democracies (Aytaç 2018).

information revelations (Banerjee et al. 2011; Cruz, Keefer and Labonne 2019). If such equilibrium responses affect vote choices, changes in voter beliefs may not ultimately translate into changes in behavior. In sum, we anticipate that:

H4. *The magnitude of information's effects on incumbent support and efforts to contact incumbents after the election is greatest among voters that are civically-educated, value incumbent performance indicators, and are less susceptible to forces counteracting the information.*

3 Parliamentary accountability in Senegal

Senegal is one of Africa's oldest and strongest democracies. It has generally experienced robust multi-party political competition—including peaceful transitions in 2000 and 2012, following fair democratic elections—since 1981, and is known for its vibrant civil society and freedom of press and expression. However, voters are often poorly informed about legislative politics, and political accountability remains low.

3.1 The Assemblée Nationale's role

The Assemblée Nationale (Parliament) plays a limited role in democratic representation in Senegal's hyper-presidential context (Beck 2012; Thomas and Sissokho 2005). Deputies are elected for five-year terms by a mixed system, where competing coalitions form a national list and submit lists for each of Senegal's 45 departments (which serve as parliamentary districts). In each department, the coalition winning most votes receives all seats allotted to the department. In 2017, 105 deputies were elected from 12 single and 33 multi-member departments and 60 seats were allocated in proportion to a coalition's national vote share. In the 2012 legislative elections, president Macky Sall's coalition—Benno Bokk Yakaar (BBY)—won 87 of 90 majoritarian departmental seats and approximately half the proportionally-allocated seats. Our study examines deputies elected from departmental majoritarian lists because of their stronger electoral ties to constituents.

The primary constitutional role of elected deputies is amending and voting on laws drafted by

government ministries. However, few laws are rejected by the Assemblée Nationale, and its role in checking executive power is often questioned by civil society. Deputies can also initiate laws themselves, although this is rare in practice (Thomas and Sissokho 2005).

Nonetheless, deputies can—and do—affect legislative decisions through their parliamentary duties. First, deputies can serve on the Assemblée's 11 parliamentary committees, through which they can make recommendations and amendments to ministerial bills before plenary debates. Second, deputies can submit questions to the government to defend and publicize their constituents' interests, which relevant ministers answer in open sessions. Third, although deputies do not receive specific funds for local development projects, they are widely believed—and themselves claim—to influence the allocation of local projects and government transfers by lobbying ministers. Indeed, one deputy described the biggest difference between good and bad deputies as their "capacity to lobby successfully."

3.2 Voter engagement with parliamentary elections and deputies

Voter turnout in Senegal reached 54% in the 2017 parliamentary elections, slightly below the sub-Saharan African mean. Nevertheless, Senegal's 2016 Afrobarometer round indicates that 87% of respondents viewed Senegal as a democracy, and 64% reported being satisfied with the functioning of Senegalese democracy. While direct citizen interaction with deputies is rare (only 9% of respondents in our sample had contacted a deputy within the last year) and voters are pessimistic about whether deputies listen to voters and respond to requests, interactions with party officials and brokers who report to deputies are relatively common.

Although election outcomes often reflect nationwide swings in coalition support, our baseline survey data indicates that many voters also seek to elect deputies that bring local development projects to their department. Figure 1 indicates that 46% of voters claim that a deputy's potential to lobby for projects and transfers benefiting their department is the most important factor driving their vote choice. Fewer voters regard national-level policy engagement as important. Moreover, when asked to choose between hypothetical deputies seeking to improve voters' welfare, 71% of

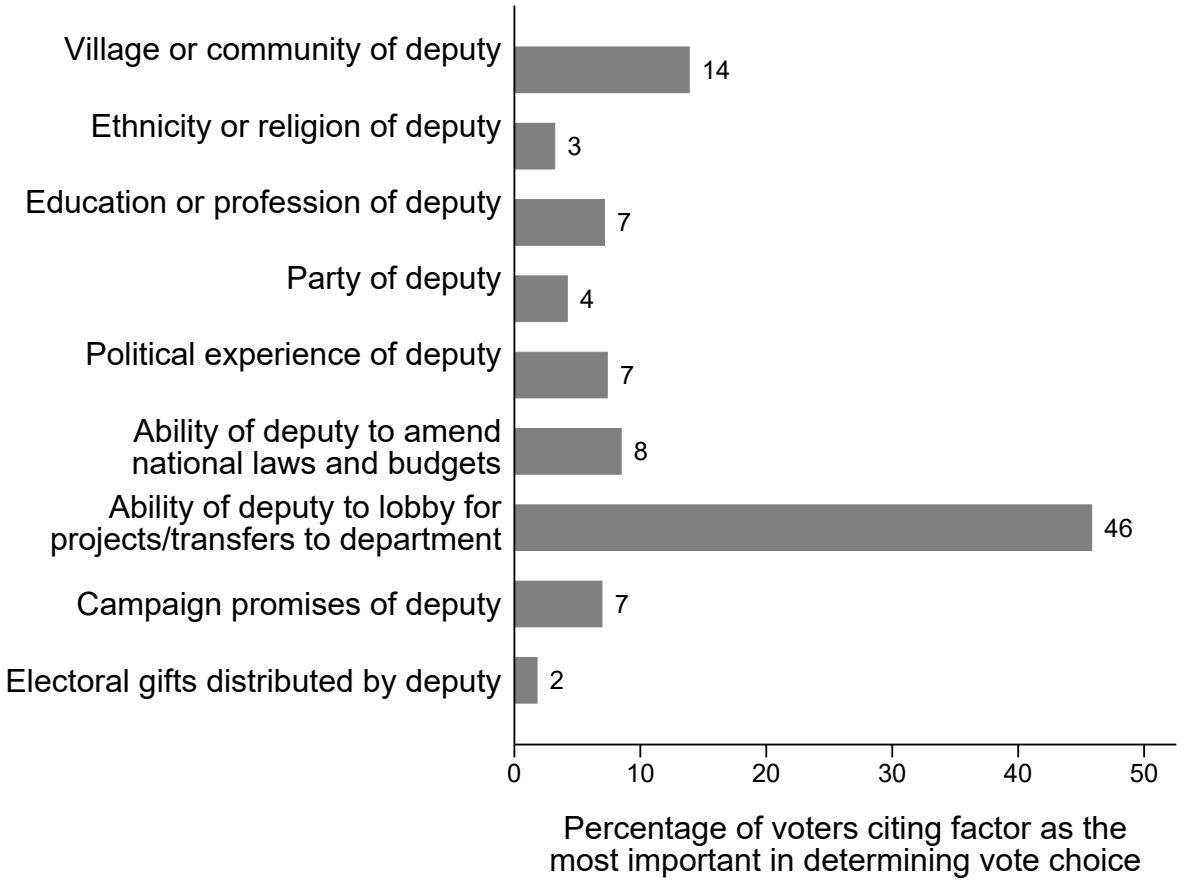

Figure 1: The most important factor driving individuals' vote choices

respondents favored locally-oriented politicians to nationally-oriented ones.

However, actually holding deputies to account has proved challenging for several reasons. First, voters often lack the information needed to identify the best-performing deputies. Only 35% of voters in our sample could name at least one of their parliamentary representatives, and only 61% could correctly identify the incumbent party in their department. Moreover, voters' prior beliefs are uncorrelated with the incumbent performance metrics that our treatment provides (see Figure 5 below). The paucity of reliable information partly reflects the limited penetration of mass media and election campaigns in rural communities. Second, attempts to hold deputies to account often compete against clientelistic incentives and coordinated group voting pushing vote choices in different directions. Political parties can heavily influence rural vote choice via village chiefs

and other local brokers as intermediaries (Koter 2013).

4 Research design

We designed an information dissemination campaign in partnership with a local civil association—LEGS-Africa, a transparency-oriented organization in Dakar—to personally deliver and explain incumbent performance information to voters prior to Senegal's 2017 parliamentary elections. We randomized core components of the information's content across villages, and used a panel study to track voter beliefs and actions before the information treatments were delivered, immediately after their delivery, and a month after the election. This design thus traces the key links between voters' receipt of information and whether and how voters hold incumbent deputies to account.

4.1 Sample selection

We conducted our study in the five departments shown in Figure 2: Fatick, Foundiougne, Kanel, Oussouye, and Ranérou Ferlo. In each department, the current deputies were from the BBY coalition—the president's ruling coalition—and the previous deputies were from the Sopi coalition—the previous president's ruling coalition and BBY's rival. Within these departments, we selected 450 rural villages containing 200-4,000 people for our sample. Appendix Table D1 shows that this sample is less educated and developed than the national average. Within each village, we aimed to survey nine registered voters aged 20-38 that had lived in the village prior to the age of primary school enrollment. Appendix section D provides further sampling information.

4.2 Information treatments

Our treatments entailed distributing and explaining scorecards detailing combinations of legislator duties, current incumbent performance, and previous incumbent legislator performance in the month preceding the election. Regarding legislator duties, we highlighted that legislators can: (1) serve on the 11 parliamentary committees; (2) participate in parliamentary debates; and (3) lobby

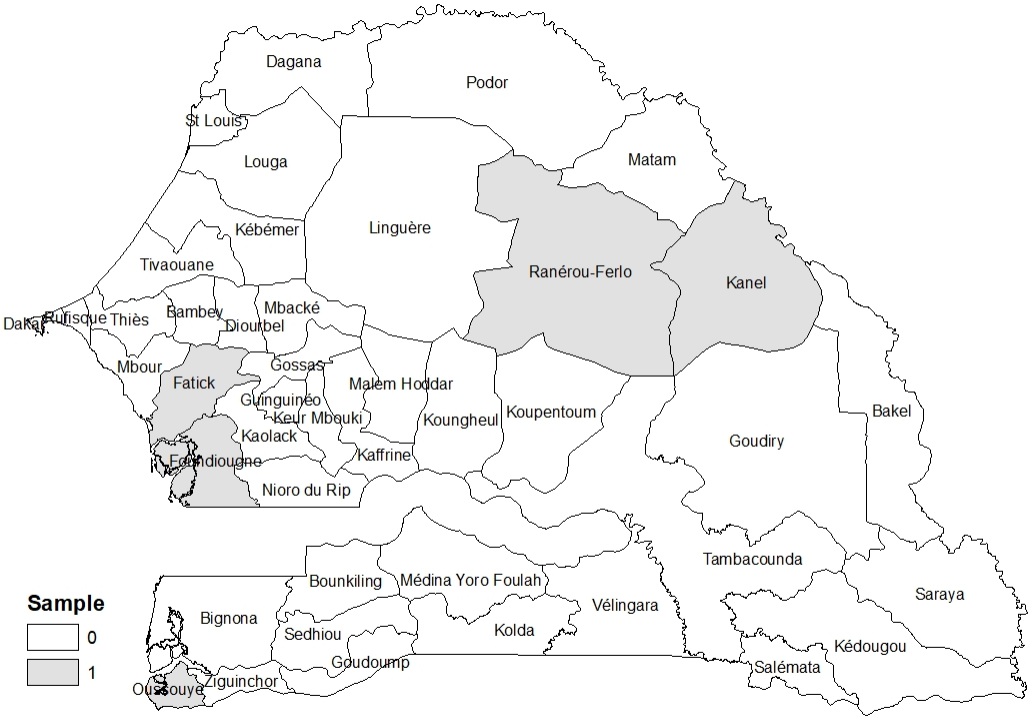

Figure 2: Sampled Senegalese departments

government ministers to allocate projects and transfers to their departments.

Regarding incumbent legislator performance, we provided five nationally- and locally-oriented measures of performance in office over the five-year electoral cycle that relate to deputies' primary duties: (1) committee memberships; (2) positions of leadership within parliament; (3) the number of parliamentary debates participated in; (4) the number of local projects budgeted for their department in parliamentary documents; and (5) the number and (inflation-adjusted) per capita per year value of ministry transfers received by the department, decomposed by transfer category.[4] All deputy- and department-specific data was obtained from the Assemblée Nationale or ministries, and its accuracy, relevance, and impartiality was validated by the head of legislative services at the Assemblée Nationale, the librarians and archivists at the Assemblée Nationale, and several active

[4]Annual transfer data was available from 2010 to 2016, and normalized by 2013 population size. Transfers affecting multiple departments were distributed in proportion to each department's 2013 population.

Table 1: Treatment conditions

Whether duties information is provided:	**Type of performance information provided**: *None*	*Incumbent*	*Benchmark*
None	75 villages [pure control]	75 villages	75 villages
Duties	75 villages	75 villages	75 villages

and former deputies.

Based on the performance metrics just described, we used a 2×3 factorial design to randomly assign villages to one of the six experimental conditions in Table 1. Treatment conditions vary along two dimensions of content, and include a pure control group. First, the "duties" dimension informed voters of the three main functions (enumerated above) that deputies can perform. Second, the "performance" dimension varied whether voters received "incumbent" information relating to the/an incumbent representative's performance on the five measures described above or "benchmark" information additionally providing the same information pertaining to the performance of the/a department's previous incumbent representative. In the multi-member departments where two incumbent deputies sought re-election (Kanel) or more than one deputy held office during the previous legislative session (Fatick, Foundiougne, and Kanel), we maximized treatment homogeneity by randomly selecting one current deputy seeking re-election and one benchmark deputy per randomization block (defined below) to be reported on in every leaflet delivered within the block.

Figure 3 reports the distribution of the performance metrics provided, where each point represents a current incumbent-previous incumbent pairing. Points above the 45^{o} line represent cases where the current incumbent outperformed the previous incumbent. The current incumbents often outperformed preceding incumbents, especially with respect to debates, projects, and transfers. We accordingly anticipated that our performance information would increase voters' favorability towards current incumbents, on average, across departments. However, we also examine heterogeneity by performance level.

Each information treatment was distributed to voters through leaflets like the one in Figure 4.

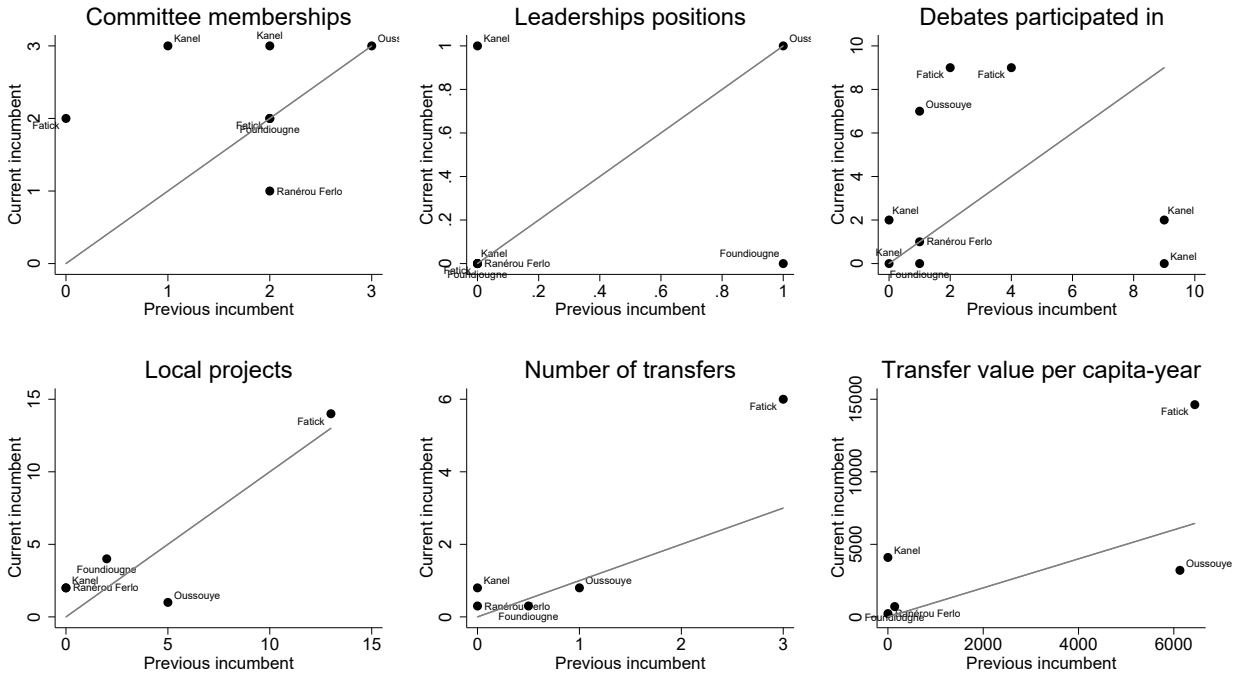

Figure 3: Distribution of treatment information across departments (45^{o} line in gray)

Note: Cases within departments where previous incumbents performed identically are not duplicated.

The leaflets were professionally designed in partnership with LEGS-Africa. Each leaflet variant showed the LEGS-Africa logo alongside a statement that the organization is non-partisan at the top, while data sources and (redacted) contact information were provided at the bottom. The example in Figure 4 depicts the duties and benchmark treatment variant—the maximum amount of information that was provided.[5] The three paragraphs below the LEGS-Africa logo were provided to all participants receiving a "duties" variant. The current incumbent performance information on the left of the leaflet was provided to participants receiving the "incumbent" variant, while the performance information on the left *and right* was provided to participants receiving the "benchmark" variant. The leaflet was piloted to ensure comprehensibility.

The leaflet was delivered and explained *in person*, on behalf of LEGS-Africa, to respondents during our baseline survey. Enumerators gave each voter several minutes to read the leaflet in French and then spent several minutes explaining the meaning of each component in the re-

[5]Appendix Figures E1-E4 show our other leaflets.

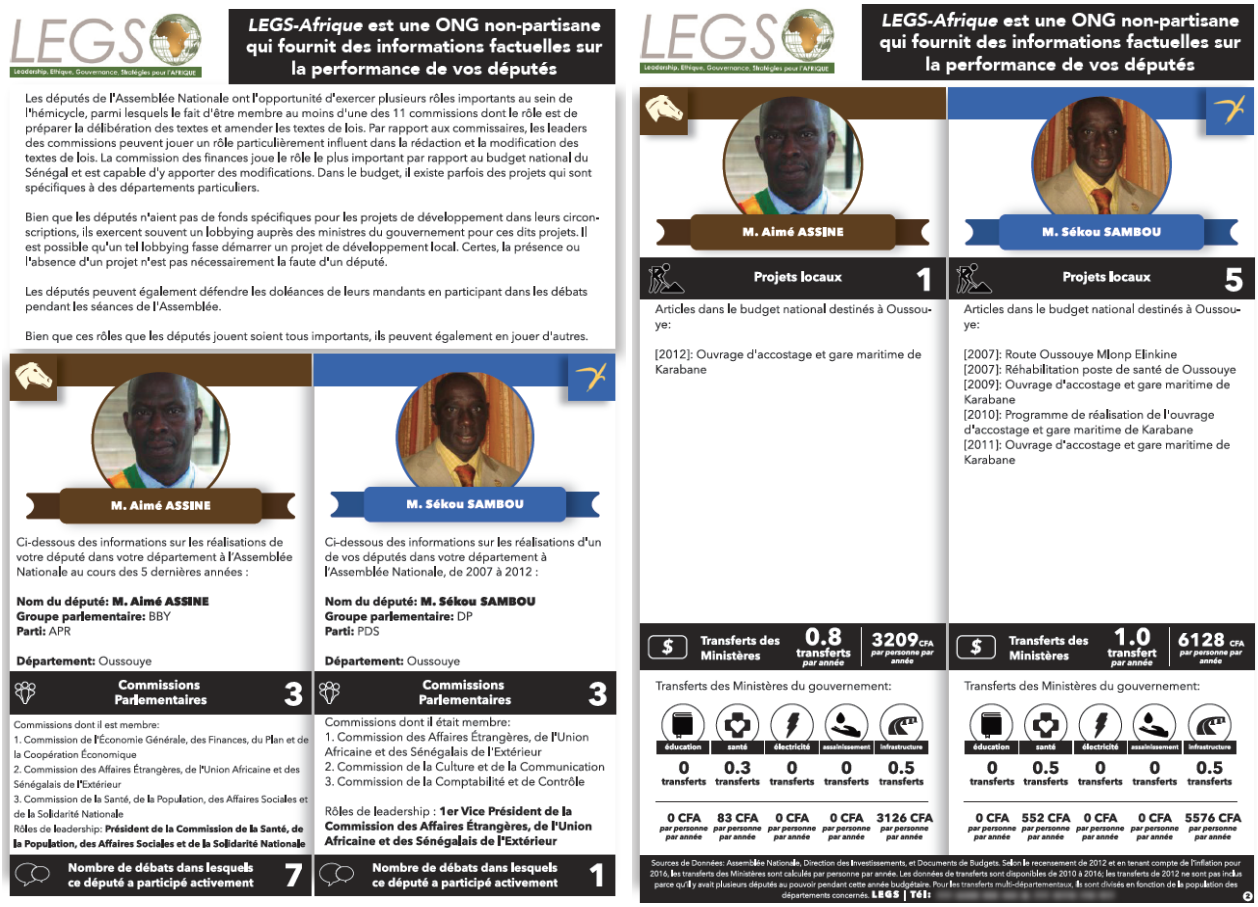

Figure 4: Example of "duties + benchmark" treatment in Oussouye

spondent's local language. Our training ensured that enumerators—mostly university graduates—themselves understood and could clearly explain the leaflets' content in both languages. On average, treatment delivery took around five minutes. Suggesting that the treatment was regarded as credible, 82% of treated respondents reported that the leaflet came from an NGO.

Our intervention is thus heavier-handed than most prior information dissemination campaigns. Previous campaigns have posted fliers, sent SMS messages, created newspaper articles, or arranged dissemination meetings or video viewings (e.g. Banerjee et al. 2011; Chong et al. 2015; Dunning et al. 2019). Unlike studies providing *access* to information, we ensure that voters *received and understood* the information to focus on belief updating and voter behavior absent dissemination constraints. Several prior studies conduct similarly intensive interventions (e.g. Boas, Hidalgo and

Melo 2019; Humphreys and Weinstein 2012).

4.3 Information provision randomization

Leaflet treatment conditions were block-randomized at the village level to mitigate contamination arising from within-village spillovers. Specifically, we constructed 75 blocks, and assigned each experimental condition to one of six similar villages from within the same department.[6] In multi-member departments, we used complete randomization to assign an incumbent-previous incumbent pair to each block.[7]

4.4 Data collection

We designed a two-wave panel survey and collected polling station-level electoral returns. The baseline survey was conducted in person between July 4 and July 29, and our treatments were administered after enumerators collected respondents' characteristics, baseline beliefs, previous behaviors, and intentions. The shorter post-election survey was conducted by telephone between August 4 and August 26. We also mapped each village to its associated polling station.

4.4.1 Measurement of primary outcomes

Our primary classes of outcomes focus on voter beliefs and whether voters ultimately engaged in electoral accountability and requests to contact winning incumbents after the election.

First, we measured voters' beliefs about how well incumbents have done overall since they were elected in 2012, how they compare with the previous incumbent, and how the current incumbent seeking re-election would do if re-elected on five-point scales from "very bad" (1) to "very good" (5).[8] For each variable, we also elicited the strength of voters' assessment on a ten-point scale ranging from "not at all certain" (1) to "completely certain" (10).[9] These beliefs, and their as-

[6] After stratifying by department, village similarity was determined by Mahalanobis distance across 24 pre-treatment covariates.

[7] Appendix Table E1 reports the distribution of configurations by deputy.

[8] "Don't know" responses are coded at the mid-level of the scale.

[9] "Don't know" responses are coded as the lowest level of certainty.

sociated certainty, were elicited at baseline before and after information treatments were delivered for both treated and control respondents. The first two questions were repeated at endline.

Second, we elicited voting behavior: vote intention before and after treatment in the baseline survey, and self-reported turnout and vote choice at endline. We use indicators for respondents stating that they would or did vote for the incumbent. We address self-reporting concerns at endline with a robustness check that only counts votes as valid where the respondent correctly recalled features of the ballot and its party-specific color. We also elicited certainty about intended vote choice on a ten-point scale. Furthermore, we use the electoral returns to calculate incumbent party vote share at the polling station corresponding to each village in our sample. Although fewer than 2% of voters were treated within polling stations, information could spread within our tight-knit set of rural villages to affect this behavioral outcome.

Third, we measure costly efforts to contact winning incumbents after the election. At baseline, we offered respondents the opportunity to request a visit from, or sign up to be contacted to express their views to, any party or candidate if they were subsequently elected.[10] These behavioral measures capture non-electoral means of seeking accountability, akin to Aker, Collier and Vicente (2017), Bussell (2019), and Grossman, Humphreys and Sacramone-Lutz (2014). The post-election endline survey again offered respondents the opportunity to request a visit from and sign up to be contacted by the winning candidate—the incumbent in each department. At endline we also created a hotline where respondents could send text messages (costing around US$0.04, or 5% of rural per capita daily expenses) or leave voicemails (US$0.18, or 21% of rural per capita daily expenses) requesting to be contacted by the winning candidate. We measure this by linking telephone numbers to the respondent. These types of opportunity to engage directly with elected politicians are rare in rural Senegal, as in other African contexts; see Appendix section C.

Given the large number of outcomes—which engender concerns about multiple comparisons and noise in specific variables—we combine related individual-level outcomes using indexes. Separately within baseline and endline panel waves, we created inverse-covariance weighted (ICW)

[10]The voter's name and village were shared with the party by LEGS-Africa.

indexes to summarize two groups of items: incumbent evaluation, i.e. all attitudinal and voting outcomes; and behavioral indicators of requests to contact incumbents.[11] By standardizing all indexes with respect to the control group, effect magnitudes represent standard deviation changes in control group outcomes.

4.4.2 Compliance

We encountered two minor forms of data missingness. First, we could not access 7 villages. However, since villages were surveyed identically by enumerators and not informed of treatment status in advance, the opportunity to conduct surveys was unaffected by treatment assignment. Second, 4% of respondents attrited between baseline and endline surveys, but not differentially so across treatment conditions. Appendix section F provides additional details, and reports balance tests supporting the randomization's integrity.

4.5 Estimation

Following our pre-analysis plan, the following fully-saturated OLS regression specification estimates the average treatment effect of different informational components of the leaflet:[12]

$$\begin{aligned} Y_{iv} &= \alpha Y_{iv}^{baseline} + \beta_1 duties_v + \beta_2 incumbent_v + \beta_3 benchmark_v \\ &\quad + \beta_4 \left(incumbent_v \times duties_v\right) + \beta_5 \left(benchmark_v \times duties_v\right) + \gamma_b + \delta_e + \varepsilon_{iv}, \end{aligned} \tag{1}$$

where Y_{iv} is an outcome for respondent i in village v, γ_b are randomization block fixed effects, and δ_e are enumerator fixed effects. Wherever possible, the outcome's pre-treatment baseline counterpart $Y_{iv}^{baseline}$ is included to increase estimation efficiency. For polling station-level outcomes, we replace the iv subscript with a p subscript. To recover the village-level average treatment effect, all survey-based regressions are weighted by the inverse of the number of respondents in the corresponding baseline or endline survey. Standards errors are clustered by village. One-sided t tests

[11] The ICW approach accounts for correlation among items; see Appendix section G.

[12] Appendix section H explains minor deviations from the pre-analysis plan.

are applied to pre-specified directional hypotheses. Two-sided t tests are applied to hypotheses that were not pre-specified or were pre-specified without a hypothesized direction, and—denoted by distinct symbols—to estimates in the opposite direction to our pre-specified hypothesis.

We test additional hypotheses underpinning the accountability process by further estimating heterogeneity in treatment effects by the content of the information provided, voter's prior beliefs, or the importance of content for a voter's decision-making. While such predetermined moderators are not randomly assigned, these tests further evaluate consistency with the political accountability logic.

5 Immediate effects of information provision

We start by examining immediate responses to receiving information at the end of the baseline survey. This enables us to assess voter-level links in the accountability logic upon receiving information and before any further interactions with other voters or political actors occur.

5.1 Voters comprehend the leaflet's information

We first verify that voters comprehended the treatment information. All respondents were asked four factual multiple-choice questions pertaining to different components of the leaflet's information.

The results in Table 2 demonstrate that most respondents comprehended the information. Columns (1) and (2) indicate that receiving any duties information increased the proportion of respondents correctly identifying the number of parliamentary committees from 5% to 71% and that deputies lack individual funds for department projects from 14% to 60%. Column (3) demonstrates that incumbent performance information increased the proportion of respondents correctly identifying the number of local projects received under the current incumbent from 8% to around 80%. Finally, column (4) shows that the benchmark leaflet increased correct answers regarding the number of debates that the previous incumbent participated in from 7% to 53%. Voters' inability to gener-

Table 2: Leaflet comprehension (baseline survey)

	Respondent correctly states...			
	...number of parliamentary committees	...deputies lack department fund	...number of incumbent's local projects	...number of previous incumbent's debates
	(1)	(2)	(3)	(4)
Duties	0.663**	0.459**		
	(0.022)	(0.022)		
Incumbent			0.729**	
			(0.024)	
Benchmark			0.717**	0.461**
			(0.025)	(0.027)
Two-sided test: Incumbent = Benchmark (p value)			0.61	
Observations	3,999	3,999	3,999	3,999
Outcome range	{0,1}	{0,1}	{0,1}	{0,1}
Control outcome mean	0.05	0.14	0.08	0.07
Control outcome std. dev.	0.22	0.35	0.26	0.26

Notes: Each regression includes randomization block and enumerator fixed effects. Observations are inversely weighted by the number of respondents surveyed in the village. Standard errors are clustered by village. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from pre-specified one-sided t tests.

ally comprehend the information provided therefore does not seem to represent a bottleneck in the political accountability process.

5.2 Voters update their beliefs in a Bayesian manner

Given that information may not be perceived as credible or relevant, voter comprehension does not necessarily imply that our leaflets' information would alter voter beliefs. Furthermore, regurgitating information may not imply sophisticated internalization.

Columns (1)-(3) in Table 3 show that, on average, voters favorably updated their posterior beliefs immediately after receiving the information—in line with hypothesis H1. Panel A indicates that treated voters who received incumbent-only, and especially benchmarked, performance information experienced around a third of a standard deviation increase in favorability toward their incumbent deputy across each assessment of their suitability for office.[13] Appendix Table I1 fur-

[13] Around 25% of treated voters favorably updated their pre-treatment beliefs; very few updated unfavorably.

Table 3: Average effects of information treatments (baseline survey)

	Incumbent evaluation outcomes					Incumbent contact request outcomes		
	Incumbent overall performance (1)	Relative performance (vs. previous) (2)	Prospective incumbent performance (3)	Incumbent vote intention (4)	Incumbent evaluation index (ICW) (5)	Request incumbent visit (6)	Request incumbent conversation (7)	Incumbent contact request index (ICW) (8)
Panel A: All information treatment conditions								
Duties	0.062	-0.043	0.066	0.003	0.003	-0.023	-0.027	-0.056
	(0.065)	(0.056)	(0.053)	(0.013)	(0.042)	(0.020)	(0.018)	(0.041)
Incumbent	0.362**	0.221**	0.239**	0.030**	0.227**	0.008	0.005	0.016
	(0.063)	(0.051)	(0.054)	(0.013)	(0.040)	(0.019)	(0.018)	(0.038)
Incumbent × Duties	-0.014	0.127*	0.044	0.002	0.052	0.061**	0.046*	0.120**
	(0.088)	(0.077)	(0.076)	(0.020)	(0.060)	(0.027)	(0.027)	(0.058)
Benchmark	0.457**	0.353**	0.376**	0.037**	0.331**	0.002	0.001	0.004
	(0.066)	(0.057)	(0.057)	(0.015)	(0.045)	(0.020)	(0.019)	(0.042)
Benchmark × Duties	-0.051	0.041	-0.098	-0.004	-0.036	0.028	0.038†	0.074
	(0.091)	(0.084)	(0.077)	(0.020)	(0.064)	(0.028)	(0.027)	(0.059)
Panel B: Pooling duties treatment conditions								
Incumbent	0.356**	0.285**	0.262**	0.031**	0.253**	0.039**	0.029**	0.076**
	(0.044)	(0.038)	(0.037)	(0.010)	(0.029)	(0.014)	(0.013)	(0.029)
Benchmark	0.432**	0.375**	0.328**	0.035**	0.313**	0.017	0.021†	0.042†
	(0.045)	(0.042)	(0.039)	(0.010)	(0.032)	(0.014)	(0.013)	(0.030)
Benchmark - Incumbent	0.076*	0.089*	0.066*	0.005	0.060*	-0.022	-0.008	-0.034
	(0.042)	(0.039)	(0.039)	(0.011)	(0.032)	(0.014)	(0.014)	(0.029)
Observations	3,942	3,932	3,928	3,999	3,891	3,999	3,998	3,998
Outcome range	{1,...,5}	{1,...,5}	{1,...,5}	{0,1}	[-2.3,1.9]	{0,1}	{0,1}	[-1.6,0.7]
Control outcome mean	2.83	3.20	3.15	0.59	0.00	0.70	0.70	-0.00
Control outcome std. dev.	1.07	0.90	1.09	0.49	1.00	0.46	0.46	1.00

Notes: Each regression includes randomization block and enumerator fixed effects and a lagged dependent variable (or pre-treatment incumbent vote intention as a proxy). Observations are inversely weighted by the number of respondents surveyed in the village. Standard errors are clustered by village. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from pre-specified one-sided t tests.

ther shows that treatment increased voters' certainty in their beliefs about the incumbent's current and future performance by nearly 0.2 standard deviations. Such updating is consistent with voters being poorly informed—Figure 5 reports a somewhat negative correlation between reported performance and pre-treatment beliefs—and receiving credible performance indicators that generally exceeded prior expectations.

Conversely, information about deputies' duties did not systematically affect voter evaluations—whether on its own or in conjunction with performance indicators. This lack of support for hypothesis H2 indicates that voters do not need additional information about incumbent duties to internalize incumbent performance information. The lack of systematic effects of information about duties is a consistent pattern throughout this study. This could reflect a plurality of respondents caring most about deputies bringing projects back to their department and already believed that incum-

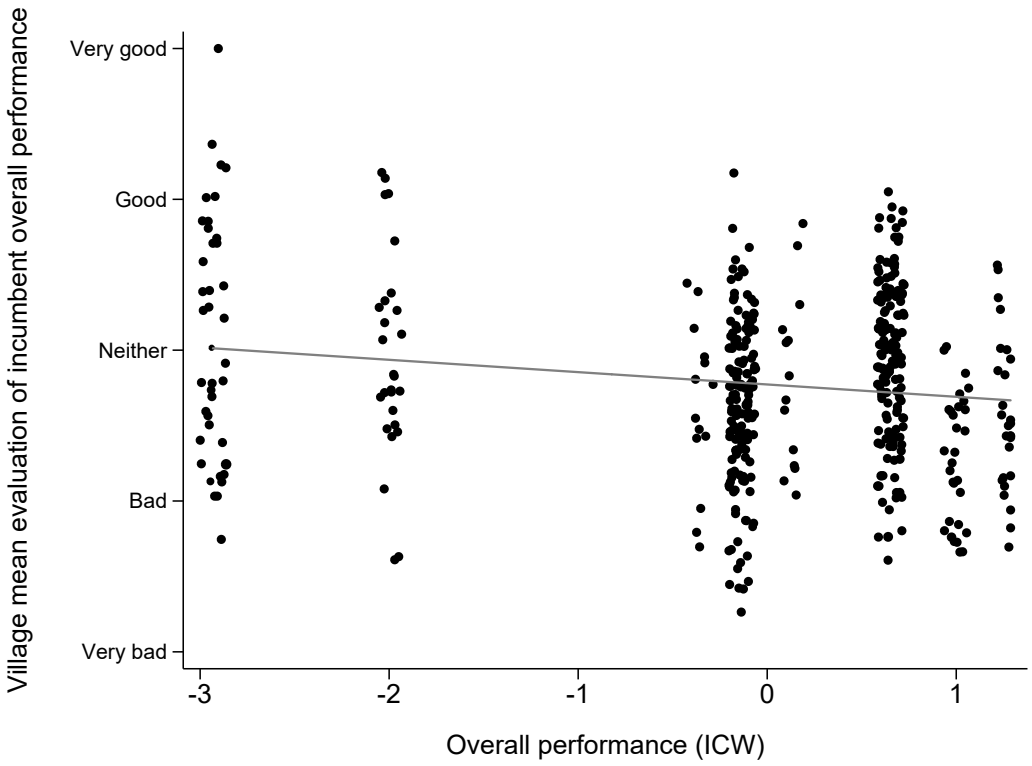

Figure 5: Correlation between pre-treatment incumbent overall performance assessments and incumbent deputy performance

Note: Villages are jittered around the overall performance (ICW) levels for each of the eight deputies.

bents were capable of influencing this process. Henceforth, we focus primarily on the comparison between benchmarked and non-benchmarked incumbent performance information by pooling the control and duties conditions.

The pooled specification in panel B of Table 3 demonstrates significant differential effects of within-department temporal benchmarks. Consistent with voters using the previous incumbents' generally worse performance to filter out department-specific effects (hypothesis H3), the tests at the foot of columns (1) and (3) show that benchmarked information increased voters' overall and prospective posterior assessments more than receiving incumbent-only information. Appendix Table I1 further shows that the benchmark also differentially increased belief precision. Suggesting that this greater precision increased the weight attached to the signal of incumbent performance in voters' posterior beliefs, column (2) also reports that the benchmarked information had a larger effect on the *relative* comparison between current and previous incumbents than incumbent-only

information.[14]

Finally, the heterogeneous effect results in Table 4 indicate that voter processing of information is consistent with Bayesian updating (hypothesis H1). Panel A first shows that voters updated significantly more favorably about their incumbent when the leaflet indicated higher performance on a standardized ICW scale combining our six reported performance indicators. Panel B further shows that voters' increased favorability almost entirely reflects the "local" projects and transfers components of the performance index, suggesting that—consistent with their own stated preference—voters are mostly concerned with the resources that deputies bring to their departments. As in Adida et al. (forthcoming), engaging in "national" legislative efforts, if anything, appears to constitute a negative signal. Moreover, Appendix Table I3 shows that the voters with the least favorable and least precise prior beliefs updated most favorably about the incumbent. Together, these findings imply sophisticated voter learning about incumbent quality, suggesting that neither cognitive capacity nor resistance to information impeded political accountability in rural Senegal.

5.3 Performance information alters vote intentions

We next examine whether these changes in beliefs translate into vote intentions. Column (4) of Table 3 indicates that receiving incumbent performance information increased intentions to vote for the incumbent deputy by 3 percentage points (hypothesis H1).[15] This effect is again larger for benchmarked than incumbent-only information, although not significantly so (hypothesis H3). We also again find no evidence to suggest that incumbent duties information directly affected vote intention or moderated the effects of performance information (hypothesis H2).

Effects on vote intentions also vary in line with the electoral accountability logic. First, changes in incumbent vote intentions are consistent with changes in voters' beliefs (hypothesis H1). Column (4) of panels A and B of Table 4 shows that treatments revealing better performance—

[14] Appendix Table I2 shows that benchmarks did not alter average prospective challenger performance evaluations. Previous incumbent performance falling below prior expectations thus does not appear to drive the differential effects of benchmarked information.

[15] Voters' already-high vote choice certainty did not significantly increase.

Table 4: Heterogeneous effects of information treatments (baseline survey)

	Incumbent evaluation outcomes					Incumbent contact request outcomes		
	Incumbent overall performance (1)	Relative performance (vs. previous) (2)	Prospective incumbent performance (3)	Incumbent vote intention (4)	Incumbent evaluation index (ICW) (5)	Request incumbent visit (6)	Request incumbent conversation (7)	Incumbent contact request index (ICW) (8)
Panel A: Heterogeneity by (standardized) reported performance level								
Incumbent	0.355**	0.285**	0.261**	0.031**	0.253**	0.039**	0.029**	0.076**
	(0.042)	(0.036)	(0.036)	(0.010)	(0.028)	(0.014)	(0.013)	(0.029)
Incumbent × Overall performance (ICW)	0.233**	0.179**	0.177**	0.025**	0.115**	0.012	0.008	0.023
	(0.050)	(0.044)	(0.044)	(0.009)	(0.029)	(0.017)	(0.017)	(0.037)
Benchmark	0.430**	0.373**	0.327**	0.035**	0.313**	0.016	0.021†	0.041†
	(0.044)	(0.041)	(0.038)	(0.010)	(0.031)	(0.014)	(0.013)	(0.030)
Benchmark × Overall performance (ICW)	0.230**	0.140**	0.156**	0.037**	0.137**	0.014	0.003	0.019
	(0.051)	(0.047)	(0.050)	(0.010)	(0.033)	(0.019)	(0.018)	(0.041)
Overall performance (ICW) range	[-2.37,1.12]	[-2.37,1.12]	[-2.37,1.12]	[-2.37,1.12]	[-2.37,1.12]	[-2.37,1.12]	[-2.37,1.12]	[-2.37,1.12]
Panel B: Heterogeneity by (standardized) local and national reported performance level								
Incumbent	0.359**	0.287**	0.264**	0.030**	0.255**	0.039**	0.028**	0.076**
	(0.039)	(0.034)	(0.035)	(0.010)	(0.026)	(0.014)	(0.013)	(0.029)
Incumbent × National performance (ICW)	0.031	0.068**	0.012	0.001	0.013	0.000	0.016	0.017
	(0.043)	(0.034)	(0.033)	(0.009)	(0.025)	(0.015)	(0.014)	(0.031)
Incumbent × Local performance (ICW)	0.317**	0.239**	0.250**	0.025**	0.201**	0.009	-0.004	0.006
	(0.045)	(0.038)	(0.038)	(0.011)	(0.030)	(0.016)	(0.015)	(0.033)
Benchmark	0.431**	0.373**	0.327**	0.035**	0.313**	0.016	0.021†	0.041†
	(0.042)	(0.038)	(0.037)	(0.010)	(0.029)	(0.014)	(0.013)	(0.030)
Benchmark × National performance (ICW)	-0.032	-0.127§§	-0.071‡	0.006	-0.063§	-0.005	0.003	-0.002
	(0.048)	(0.044)	(0.038)	(0.009)	(0.030)	(0.015)	(0.014)	(0.031)
Benchmark × Local performance (ICW)	0.236**	0.271**	0.221**	0.035**	0.220**	0.005	0.013	0.020
	(0.047)	(0.042)	(0.039)	(0.011)	(0.032)	(0.016)	(0.015)	(0.033)
National performance (ICW) range	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]
Local performance (ICW) range	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]
Panel C: Heterogeneity by importance of performance in determining vote choice								
Incumbent	0.355**	0.284**	0.261**	0.031**	0.253**	0.039**	0.029*	0.077**
	(0.044)	(0.037)	(0.037)	(0.010)	(0.029)	(0.014)	(0.013)	(0.029)
Incumbent × Performance most important	-0.007	-0.008	0.009	0.020*	0.021	0.021	0.019	0.045†
	(0.034)	(0.029)	(0.030)	(0.009)	(0.022)	(0.013)	(0.013)	(0.027)
Benchmark	0.431**	0.373**	0.328**	0.036**	0.313**	0.017	0.022	0.043
	(0.045)	(0.041)	(0.039)	(0.010)	(0.032)	(0.014)	(0.013)	(0.030)
Benchmark × Performance most important	-0.001	-0.023	-0.010	0.016†	0.007	0.024†	0.019	0.048†
	(0.031)	(0.028)	(0.030)	(0.010)	(0.022)	(0.013)	(0.013)	(0.027)
Performance most important range	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}
Panel D: Heterogeneity by preference for locally-oriented deputies								
Incumbent	0.370**	0.192**	0.198**	-0.018	0.172**	0.041†	0.035	0.085†
	(0.074)	(0.070)	(0.063)	(0.023)	(0.054)	(0.029)	(0.029)	(0.063)
Incumbent × Prefer locally-oriented deputies	-0.019	0.130*	0.089	0.068**	0.114*	-0.002	-0.009	-0.012
	(0.078)	(0.077)	(0.071)	(0.027)	(0.062)	(0.033)	(0.033)	(0.071)
Benchmark	0.502**	0.286**	0.302**	0.014	0.279**	0.028	0.028	0.062
	(0.074)	(0.074)	(0.066)	(0.022)	(0.057)	(0.028)	(0.027)	(0.060)
Benchmark × Prefer locally-oriented deputies	-0.098	0.123†	0.036	0.029	0.046	-0.016	-0.010	-0.028
	(0.078)	(0.076)	(0.070)	(0.024)	(0.060)	(0.032)	(0.032)	(0.068)
Prefer locally-oriented deputies range	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}
Observations	3,942	3,932	3,928	3,999	3,891	3,999	3,998	3,998
Outcome range	{1,...,5}	{1,...,5}	{1,...,5}	{0,1}	[-2.3,1.9]	{0,1}	{0,1}	[-1.6,0.7]
Control outcome mean	2.83	3.20	3.15	0.59	0.00	0.70	0.70	-0.00
Control outcome std. dev.	1.07	0.90	1.09	0.49	1.00	0.46	0.46	1.00

Notes: Each regression includes randomization block and enumerator fixed effects a lagged dependent variable (or pre-treatment incumbent vote intention as a proxy). Lower-order (standardized) interaction terms are included but not shown. Observations are inversely weighted by the number of respondents surveyed in the village. Standard errors are clustered by village. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from pre-specified one-sided t tests; ‡ $p < 0.1$, § $p < 0.05$, §§ $p < 0.01$ from two-sided tests when coefficients point in the opposite direction to the pre-specified hypothesis.

regarding on local projects and transfers—were substantially more likely to increase intent to vote for the incumbent. Appendix Table I3 similarly shows that voters with the least favorable and least precise prior beliefs were significantly more likely to intend to vote for the incumbent. Second, and supporting hypothesis H4, the effects of performance information were greatest among the voters for whom the information was most relevant. Panel C demonstrates that the 54% of respondents who ranked the incumbent's ability to amend laws and budgets or lobby for projects in the department as the most important determinant of their vote choice (before treatment dissemination) were significantly more likely to intend to vote for the incumbent after treatment. Similar results hold in column (4) of panel D among the 71% of voters that expressed a preference for a locally-oriented, as opposed to nationally-oriented, politician in our pre-treatment vignette.

5.4 Performance information increases incumbent contact requests

Our behavioral outcomes capturing requests to contact winning incumbents after the election demonstrate that revelations of better-than-expected performance also encouraged non-electoral accountability. Columns (6) and (7) in panel B of Table 3 show that, on average, treated respondents became significantly more willing to request a visit from, or a conversation with, incumbents if re-elected (hypothesis H1). The index outcome estimates in column (8) imply around a 0.05 standard deviation increase relative to the control group. For such behaviors, benchmarked information did not differentially increase requests, although this may have been limited by high rates of take-up among incumbent supporters. Broadly in line with vote intentions, the heterogeneous treatment effects in Tables 4 and I3 report larger treatment effects on incumbent requests where incumbent performance levels were greater and especially among the voters that cared most about performance (hypotheses H1 and H4). Although not always statistically significant, these estimates suggest that voters became more likely to seek contact with incumbent politicians after learning that the incumbent may be more responsive than expected.

Post-election incumbent contact requests represent the one area where information about deputies' duties might complement performance indicators. Columns (6)-(8) of panel A in Table 3 show that

learning that the incumbent is generally performing better than expected primarily translated into non-electoral accountability efforts when voters were aware of what incumbents can do (hypothesis H2). However, any suggestion that voters must believe that politicians possess the capacity to respond effectively before engaging in costly requests is tentative because it did not persist at endline.

6 Longer-term effects of information provision

To understand whether the immediate increases in electoral and non-electoral efforts to hold incumbents to account translate into persisting beliefs, voting behavior, and costly attempts to contact incumbents, we turn to our endline survey and polling station electoral returns.

6.1 Voters correctly recall leaflet content type after the election

The endline results in Table 5 show that most respondents continued to recall the treatment information around a month after its dissemination. Column (1) shows that virtually all treated respondents correctly recalled receiving the LEGS-Africa leaflet, while only 7% of control respondents incorrectly recalled receiving the leaflet. Columns (2)-(4) further demonstrate that almost as many respondents correctly remembered the differentiating features of the leaflet's content. These high recall rates substantially exceed those documented in other field experiments generally using lower-intensity information dissemination mechanisms.[16] This suggests that limited recall is unlikely to represent a significant barrier to political accountability, at least within a month of information dissemination.

[16]For example, the average treated respondent across the Metaketa studies was only 7 percentage points more likely to recall the information's substance than generally-uninformed control respondents (Dunning et al. 2019).

Table 5: Leaflet recall (endline survey)

	Received leaflet	Received duties information	Received incumbent information	Received previous incumbent information
	(1)	(2)	(3)	(4)
Any treatment	0.921**			
	(0.010)			
Duties		0.881**		
		(0.009)		
Incumbent			0.920**	
			(0.009)	
Benchmark			0.937**	0.924**
			(0.009)	(0.007)
Two-sided null: Incumbent = Benchmark (p value)			0.03	
Observations	3,875	3,875	3,875	3,875
Outcome range	{0,1}	{0,1}	{0,1}	{0,1}
Control outcome mean	0.07	0.01	0.01	0.00
Control outcome std. dev.	0.25	0.09	0.09	0.06

Notes: Each regression includes randomization block and (baseline and endline) enumerator fixed effects. Observations are inversely weighted by the number of respondents surveyed in the village. Standard errors are clustered by village. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from pre-specified one-sided t tests.

6.2 Beliefs about incumbent performance persist after the election

Updated beliefs about the incumbent also largely persisted a month after treatment. Columns (1) and (2) of Table 6 show that voters who received performance information continued to register significantly higher ratings of the incumbent, and believe that the incumbent performed better than previous incumbents (hypothesis H1). Appendix Table I1 further reports that treated voters continued to express greater certainty about their beliefs. Treatment effects roughly halved relative to the immediate effect of providing incumbent performance information. Given the lack of evidence that information spilled *across villages* (see Table I8), these estimates suggest that the smaller treatment effects at endline could reflect decay in information's effects, the influence of other factors during the election campaign, and/or post-treatment interactions within villages relating to the information's provision (see below).

The persistent changes in beliefs remain generally consistent with Bayesian updating. First,

Table 6: Average effects of information treatments (endline survey)

	Incumbent evaluation outcomes					Incumbent contact request outcomes				
	Incumbent overall performance	Relative performance (vs. previous)	Incumbent vote	Incumbent vote (validated)	Incumbent evaluation index (ICW)	Request incumbent visit	Request incumbent conversation	Request hotline number	Called hotline	Incumbent contact request index (ICW)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: All information treatment conditions										
Duties	0.012	-0.018	-0.049	-0.017	-0.074	0.006	-0.004	0.014†	0.016	0.085†
	(0.053)	(0.052)	(0.032)	(0.032)	(0.062)	(0.006)	(0.038)	(0.010)	(0.022)	(0.053)
Incumbent	0.149**	0.113**	-0.042	-0.033	0.059	0.006	-0.015	0.006	0.013	0.065
	(0.050)	(0.047)	(0.031)	(0.030)	(0.055)	(0.007)	(0.039)	(0.010)	(0.020)	(0.052)
Incumbent × Duties	0.024	0.060	0.036	0.030	0.102	-0.013	-0.005	-0.001	-0.004	-0.070
	(0.070)	(0.069)	(0.043)	(0.043)	(0.081)	(0.009)	(0.051)	(0.014)	(0.029)	(0.073)
Benchmark	0.235**	0.256**	-0.017	-0.004	0.202**	0.017**	-0.017	0.004	0.056**	0.170**
	(0.049)	(0.045)	(0.032)	(0.031)	(0.056)	(0.005)	(0.047)	(0.011)	(0.022)	(0.051)
Benchmark × Duties	0.020	-0.021	0.038	0.020	0.028	-0.015‡	0.157	0.004	-0.070§	-0.164§
	(0.072)	(0.069)	(0.044)	(0.044)	(0.082)	(0.008)	(0.151)	(0.015)	(0.030)	(0.074)
Panel B: Pooling duties treatment conditions										
Incumbent	0.161**	0.144**	-0.024	-0.018	0.110**	0.000	-0.017	0.006	0.011	0.029
	(0.035)	(0.034)	(0.021)	(0.021)	(0.039)	(0.005)	(0.029)	(0.007)	(0.015)	(0.036)
Benchmark	0.246**	0.246**	0.002	0.007	0.217**	0.009**	0.063	0.006	0.021†	0.087**
	(0.036)	(0.034)	(0.022)	(0.022)	(0.041)	(0.004)	(0.054)	(0.007)	(0.015)	(0.037)
Benchmark - Incumbent	0.085**	0.102**	0.026†	0.024	0.107**	0.009*	0.080	0.000	0.010	0.058*
	(0.033)	(0.031)	(0.020)	(0.020)	(0.036)	(0.004)	(0.065)	(0.007)	(0.014)	(0.035)
Observations	3,834	3,825	3,781	3,781	3,708	3,876	3,876	3,876	3,876	3,876
Outcome range	{1,...,5}	{1,...,5}	{0,1}	{0,1}	[-2.8,1.9]	{0,1}	{0,1}	{0,1}	{0,1}	[-7.3,1.5]
Control outcome mean	3.08	3.46	0.64	0.53	0.00	0.98	0.98	0.95	0.11	-0.00
Control outcome std. dev.	0.93	0.95	0.48	0.50	1.00	0.14	0.14	0.21	0.32	1.00

Notes: Each regression includes randomization block and (baseline and endline) enumerator fixed effects and a lagged dependent variable (or pre-treatment incumbent vote intention as a proxy). Observations are inversely weighted by the number of respondents surveyed in the village. Standard errors are clustered by village. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from pre-specified one-sided t tests; ‡ $p < 0.1$, § $p < 0.05$, §§ $p < 0.01$ from two-sided tests when coefficients point in the opposite direction to the pre-specified hypothesis.

the hypothesis tests at the foot of panel B in Table 6 indicate that benchmarked information continued to induce more favorable updating than receiving incumbent-only performance indicators (hypothesis H3). Second, the heterogeneous treatment effects with respect to reported incumbent performance in Table 7 suggest that voters concentrated their attention on certain types of information over time: a comparison of panel A with panel B indicates that voters increasingly prized higher levels of local performance, and also became more likely to view national-oriented legislative activity negatively (hypotheses H1 and H4). The lack of heterogeneous effects in panels C and D again suggests that voters valuing deputy performance and priorities differently nevertheless updated similarly from the information provided. The results thus suggest that, while the effects of information persisted, voters increasingly emphasized local performance indicators and benchmarked information over time.

Table 7: Heterogeneous effects of information treatments (endline survey)

	Incumbent evaluation outcomes					Incumbent contact request outcomes				
	Incumbent overall performance (1)	Relative performance (vs. previous) (2)	Incumbent vote (3)	Incumbent vote (validated) (4)	Incumbent evaluation index (ICW) (5)	Request incumbent visit (6)	Request incumbent conversation (7)	Request hotline number (8)	Called hotline (9)	Incumbent contact request index (ICW) (10)
Panel A: Heterogeneity by (standardized) reported performance level										
Incumbent	0.161**	0.143**	-0.024	-0.018	0.109**	0.000	-0.017	0.006	0.010	0.029
	(0.035)	(0.034)	(0.021)	(0.021)	(0.039)	(0.004)	(0.029)	(0.007)	(0.015)	(0.036)
Incumbent × Overall performance (ICW)	-0.060	-0.048	0.001	-0.001	-0.056	-0.006	-0.007	0.015†	0.020	0.035
	(0.044)	(0.044)	(0.027)	(0.027)	(0.053)	(0.005)	(0.012)	(0.010)	(0.021)	(0.053)
Benchmark	0.245**	0.245**	0.002	0.006	0.215**	0.009**	0.063	0.006	0.021†	0.087**
	(0.035)	(0.034)	(0.022)	(0.022)	(0.040)	(0.004)	(0.054)	(0.007)	(0.015)	(0.037)
Benchmark × Overall performance (ICW)	-0.049	-0.040	0.021	0.013	-0.032	-0.009‡	-0.004	-0.000	0.012	-0.029
	(0.045)	(0.046)	(0.027)	(0.028)	(0.055)	(0.005)	(0.012)	(0.011)	(0.023)	(0.056)
Overall performance (ICW) range	[-2.37,1.12]	[-2.37,1.12]	[-2.37,1.12]	[-2.37,1.12]	[-2.37,1.12]	[-2.37,1.12]	[-2.37,1.12]	[-2.37,1.12]	[-2.37,1.12]	[-2.37,1.12]
Panel B: Heterogeneity by (standardized) local and national reported performance level										
Incumbent	0.162**	0.145**	-0.024	-0.018	0.111**	0.000	-0.016	0.006	0.010	0.028
	(0.034)	(0.034)	(0.021)	(0.021)	(0.039)	(0.004)	(0.028)	(0.007)	(0.015)	(0.036)
Incumbent × National performance (ICW)	-0.028	-0.021	-0.000	-0.013	-0.022	-0.008	-0.007	0.003	0.024*	0.003
	(0.034)	(0.036)	(0.024)	(0.025)	(0.046)	(0.007)	(0.029)	(0.007)	(0.013)	(0.041)
Incumbent × Local performance (ICW)	0.068**	0.097**	0.004	0.015	0.065†	0.009*	0.016	0.014*	0.012	0.097**
	(0.034)	(0.036)	(0.023)	(0.024)	(0.044)	(0.005)	(0.018)	(0.007)	(0.014)	(0.037)
Benchmark	0.245**	0.245**	0.002	0.006	0.215**	0.009**	0.065	0.006	0.020†	0.086**
	(0.035)	(0.034)	(0.022)	(0.022)	(0.040)	(0.004)	(0.054)	(0.007)	(0.015)	(0.036)
Benchmark × National performance (ICW)	-0.039	-0.080§	0.012	0.009	-0.060	-0.017§	-0.097	-0.020§	0.012	-0.113§
	(0.034)	(0.037)	(0.026)	(0.026)	(0.044)	(0.007)	(0.074)	(0.009)	(0.013)	(0.051)
Benchmark × Local performance (ICW)	0.023	0.083**	0.012	0.018	0.059†	0.011*	0.003	0.022**	0.009	0.117**
	(0.036)	(0.036)	(0.024)	(0.024)	(0.044)	(0.006)	(0.016)	(0.008)	(0.014)	(0.044)
National performance (ICW) range	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]
Local performance (ICW) range	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]
Panel C: Heterogeneity by importance of performance in determining vote choice										
Incumbent	0.161**	0.144**	-0.024	-0.018	0.110**	-0.000	-0.016	0.006	0.011	0.029
	(0.035)	(0.034)	(0.021)	(0.021)	(0.039)	(0.005)	(0.028)	(0.007)	(0.015)	(0.037)
Incumbent × Performance most important	0.021	-0.007	0.001	0.011	0.015	-0.006	0.007	-0.001	0.003	-0.026
	(0.034)	(0.033)	(0.019)	(0.019)	(0.037)	(0.004)	(0.013)	(0.007)	(0.012)	(0.035)
Benchmark	0.246**	0.246**	0.002	0.006	0.217**	0.009*	0.063	0.006	0.021	0.087*
	(0.036)	(0.034)	(0.022)	(0.021)	(0.041)	(0.004)	(0.053)	(0.007)	(0.015)	(0.037)
Benchmark × Performance most important	-0.018	0.002	0.025	0.031	0.032	-0.005	0.060	-0.003	0.028*	0.010
	(0.035)	(0.033)	(0.019)	(0.019)	(0.037)	(0.004)	(0.052)	(0.007)	(0.014)	(0.033)
Performance most important range	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}
Panel D: Heterogeneity by preference for locally-oriented deputies										
Incumbent	0.104†	0.103*	-0.095§§	-0.073§	-0.008	-0.018§	-0.004	0.006	0.018	-0.049
	(0.064)	(0.060)	(0.035)	(0.035)	(0.067)	(0.007)	(0.038)	(0.014)	(0.025)	(0.065)
Incumbent × Prefer locally-oriented deputies	0.079	0.057	0.099**	0.077*	0.164**	0.025**	-0.015	-0.001	-0.009	0.108†
	(0.079)	(0.070)	(0.042)	(0.043)	(0.080)	(0.010)	(0.045)	(0.017)	(0.028)	(0.082)
Benchmark	0.191**	0.227**	-0.048	-0.045	0.121†	-0.003	0.283	0.002	0.050*	0.062
	(0.062)	(0.062)	(0.038)	(0.038)	(0.074)	(0.006)	(0.281)	(0.013)	(0.029)	(0.060)
Benchmark × Prefer locally-oriented deputies	0.077	0.025	0.070†	0.072†	0.134†	0.016**	-0.306	0.005	-0.041	0.035
	(0.075)	(0.073)	(0.043)	(0.045)	(0.086)	(0.008)	(0.323)	(0.015)	(0.032)	(0.073)
Prefer locally-oriented deputies range	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}
Observations	3,834	3,825	3,781	3,781	3,708	3,876	3,876	3,876	3,876	3,876
Outcome range	{1,...,5}	{1,...,5}	{0,1}	{0,1}	[-2.8,1.9]	{0,1}	{0,1}	{0,1}	{0,1}	[-7.3,1.5]
Control outcome mean	3.08	3.46	0.64	0.53	0.00	0.98	0.98	0.95	0.11	-0.00
Control outcome std. dev.	0.93	0.95	0.48	0.50	1.00	0.14	0.14	0.21	0.32	1.00

Notes: Each regression includes randomization block and (baseline and endline) enumerator fixed effects and a lagged dependent variable (or pre-treatment incumbent vote intention as a proxy). Lower-order (standardized) interaction terms are included but not shown. Observations are inversely weighted by the number of respondents surveyed in the village. Standard errors are clustered by village. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from pre-specified one-sided t tests; ‡ $p < 0.1$, § $p < 0.05$, §§ $p < 0.01$ from two-sided tests when coefficients point in the opposite direction to the pre-specified hypothesis.

6.3 Performance information influences the vote choices of likely-voters

We next examine whether the beliefs that persisted through endline carried through to voting behavior. We test this crucial link in the chain of electoral accountability by examining self-reported vote choices, before analyzing polling station-level electoral returns.

The self-reported survey data provides mixed evidence that incumbent performance information ultimately enhances electoral accountability. First, panel A of Table 6 offers little systematic evidence of an increase in incumbent voting *on average*, even after column (4) applies our vote validation criteria (hypothesis H1).[17] The pooled estimates in panel B are also indistinguishable from zero. Nevertheless, consistent with voters' more favorable updating from benchmarked information, panel B indicates that the benchmark increased BBY voting by around 2.5 percentage points more than the incumbent-only performance information. Second, the heterogeneous effects in Table 7 also yield mixed evidence. While the leaflet's content did not significantly influence the average respondent's self-reported vote, panels C and particularly D suggest that the information treatments did relatively increase incumbent support among respondents that—at baseline—regarded performance information as the most important factor in determining their vote choice or preferred locally-oriented deputies (hypothesis H4).[18] These findings suggest that only a small share of the younger rural voters in our sample ultimately acted on their updated beliefs.

However, our sample of voters may respond differently to incumbent performance information than the broader, more politically-experienced, electorate. In particular, younger voters are far less likely to turn out: within our sample, a 20-year-old was more than 20 percentage points less likely to turn out than a 33-year-old. Moreover, voters that have not previously voted were significantly less likely to value performance indicators and locally-oriented politicians. Consequently, if the relatively politically-inexperienced voters in our sample often do not turn out or vote on the basis of other factors, even when they persistently update about the incumbent's performance, then our

[17] Appendix section I.10 shows that treatment did not significantly affect turnout.

[18] Appendix Table I7 shows that treatment information did not increase the importance of performance in determining vote choice.

Table 8: Effects of information treatments, among respondents those that turned out in 2012 (baseline and endline surveys)

	Incumbent overall performance (endline)		Incumbent vote intention		Incumbent vote (validated)	
	(1)	(2)	(3)	(4)	(5)	(6)
Incumbent	0.134*	0.136*	0.015	0.018	0.034	0.034
	(0.059)	(0.058)	(0.015)	(0.015)	(0.033)	(0.032)
Benchmark	0.235**	0.231**	0.022	0.025	0.034	0.041
	(0.057)	(0.056)	(0.016)	(0.016)	(0.031)	(0.030)
Incumbent × National performance (ICW)		-0.039		-0.022		-0.060†
		(0.053)		(0.014)		(0.032)
Incumbent × Local performance (ICW)		0.072		0.032†		0.058
		(0.060)		(0.016)		(0.037)
Benchmark × National performance (ICW)		-0.077		-0.013		-0.019
		(0.060)		(0.015)		(0.035)
Benchmark × Local performance (ICW)		0.027		0.036*		0.083*
		(0.060)		(0.017)		(0.035)
Observations	1,469	1,469	1,528	1,528	1,435	1,435
Outcome range	{1,...,5}	{1,...,5}	{0,1}	{0,1}	{0,1}	{0,1}
Control outcome mean	3.10	3.10	0.63	0.63	0.59	0.59
Control outcome std. dev.	0.93	0.93	0.48	0.48	0.49	0.49

Notes: Each regression includes randomization block and (baseline and, where relevant, endline) enumerator fixed effects and a lagged dependent variable (or pre-treatment incumbent vote intention as a proxy). Observations are inversely weighted by the number of respondents surveyed in the village. Standard errors are clustered by village. Given that these hypotheses were not pre-specified, † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from two-sided t tests.

theoretical framework suggests that electoral accountability may rely on more seasoned likely-voters receiving the information and updating similarly.

To better approximate electorate-level voting behavior, Table 8 first restricts our survey sample to the 38% of voters that reported turning out in the 2012 parliamentary election. Such voters were 14 percentage points more likely to report voting in 2017. The point estimates in columns (1)-(4) indicate that previous voters immediately updated their vote intentions and persistently updated their posterior beliefs similarly to our full sample of young registered voters, suggesting that any differences in behavior are unlikely to reflect differential priors or differential updating from the information received. However, validated vote choice depicts a stark contrast. Unlike the full sample of voters, columns (3) and (5) show that previous voters remained 2-3 percentage points more likely to vote for the incumbent across baseline and endline surveys (hypothesis H4),

Table 9: Effects of information treatments (polling station data)

	Incumbent vote share (proportion of turnout)		Incumbent vote share (proportion of registered voters)	
	(1)	(2)	(3)	(4)
Incumbent	0.001	-0.013	0.006	-0.006
	(0.023)	(0.025)	(0.018)	(0.020)
Benchmark	-0.003	-0.016	-0.003	-0.014
	(0.024)	(0.026)	(0.018)	(0.020)
Incumbent × National performance (ICW)		-0.000		0.012
		(0.032)		(0.021)
Incumbent × Local performance (ICW)		0.043*		0.030*
		(0.025)		(0.017)
Benchmark × National performance (ICW)		0.021		0.011
		(0.033)		(0.021)
Benchmark × Local performance (ICW)		0.020		0.023†
		(0.026)		(0.017)
Observations	284	284	284	284
Outcome range	[0.06,0.99]	[0.06,0.99]	[0.02,0.73]	[0.02,0.73]
Control outcome mean	0.71	0.71	0.41	0.41
Control outcome std. dev.	0.17	0.17	0.13	0.13

Notes: Each regression includes randomization block fixed effects and a lagged dependent variable. Observations are unweighted, and polling stations for which the village in our sample comprises less than 50% of registered voters are excluded. Robust standard errors are in parentheses. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from pre-specified one-sided t tests.

although this is imprecisely estimated in this subsample. Furthermore, in contrast with the full sample, the heterogeneous effects in column (6) shows that the persisting belief that incumbents with higher performance scores on the local performance dimension—the primary driver of differences in election-time beliefs—are better overall did translate into a significantly higher probability of treatment increasing self-reported incumbent votes. These findings suggest that incumbent performance information may have induced electoral accountability among experienced voters, who rewarded highly-performing locally-oriented incumbents.

Due to high levels of within-village information diffusion, such responses could translate into polling station-level voting outcomes. Indeed, Appendix section I.6 shows that almost 40% of our nine treated voters per village, and at least one within every village, discussed the leaflet with

others in their village. Appendix section I.8 further reports that incumbents, and to a lesser extent challengers, also disseminated treatment information through their operatives and community meetings. In the small and densely-connected rural villages that comprise our sample, treatment information thus likely reached many citizens before the election through interpersonal interactions. To examine official electoral returns that are not susceptible to self-reporting biases, we restrict our analysis to polling stations containing the 284 villages in our sample that comprise at least 50% of registered voters at their polling station.[19]

The results in Table 9 largely mirror the self-reported behavior of the survey respondents that reported voting in 2012. We do not observe a notable average treatment effect of incumbent performance information on incumbent vote share in columns (1) and (3). However, columns (2) and (4) show that treatment effects increased with performance on the local dimension that voters value most (hypothesis H1). The estimates imply that a standard deviation increase in an incumbent's local performance increased the incumbent's vote share by around three percentage points. In sum, these findings suggest that likely-voters in treated villages that did not receive leaflets directly nevertheless learned about incumbent performance and engaged in electoral accountability. In contrast, younger and first-time voters that were less likely to vote did not.

6.4 Performance information increases incumbent contact requests

While intentions to engage in electoral accountability somewhat weakened between the time of information delivery and the election itself, non-electoral requests to contact incumbents represent a different kind of effort to hold incumbents to account. The contact requests that we study are less likely than vote choice to depend on the relative importance of performance metrics to voters, and are less susceptible to campaign-based interactions between survey waves.

Consistent with these intuitions, voters receiving performance information continued to make greater effort to contact the re-elected incumbents a month after information was delivered. As

[19]Appendix Table I11 reports similar results weighting all polling stations by the share of voters at a given polling station from the experimental village.

at baseline, columns (6)-(10) of Table 6 indicate that the benchmarked information significantly increased requests to contact incumbents (hypotheses H1 and H3). Given that low-cost requests for visits, conversations, and a hotline number through which to contact incumbents were almost universally sought, the 20% increase in the costly act of actually texting or calling the hotline in column (9) provides the most compelling evidence. Aggregated as an index, column (10) of panel B reports that benchmarked performance information induced a 0.09 standard deviation increase across such behaviors.

As with electoral accountability, the heterogeneous effects in Table 7 further demonstrate that increased hotline usage was greatest in departments where the incumbent-only and benchmarked treatments reported greatest incumbent performance (hypothesis H1). In line with self-reported behaviors among likely-voters, this principally reflected local performance. Consistent with the baseline results, requests were no greater among respondents that value performance information more in making vote choices. These results indicate that treatment caused voters to durably engage in costly efforts to seek accountability from better-performing incumbents, consistent with voters expecting greater responsiveness.

7 Conclusion

Given the mixed evidence that information campaigns can support political accountability, this article examined the extent to which accountability failures reflect voter-level constraints or the types of information that voters receive. By abstracting from issues of information dissemination and take-up, we dissect the process linking the personal delivery and explanation of incumbent performance information to electoral and non-electoral efforts by voters to hold incumbent legislators to account. Our findings show that rural Senegalese voters engaged in sophisticated information processing, largely retain their updated beliefs, and regarded local projects/transfers and temporal benchmarks as particularly informative about incumbent quality. In contrast, information about national performance and incumbent duties had little systematic effect in a context where deliv-

ering "pork" is widely regarded as a legislator's primary function. Persisting beliefs and initial intentions to hold politicians to account did translate into electoral rewards for better-performing incumbents among more experienced voters, but not among unlikely-voters or voters prioritizing other issues. With respect to non-electoral accountability, voters persistently engaged in greater costly efforts to contact the best-performing winning incumbents after the election.

Taken together, these findings illustrate that voters—upon receiving information they deem credible and relevant—are able and mostly willing to hold politicians to account. This highlights the importance of understanding how factors *other than voter-level constraints* may sustain low-accountability equilibria in developing contexts. First, future research might establish both the most effective and scalable means through which information can be communicated to voters *en masse* and the factors driving demand for and supply of such information absent external campaigns. Second, our partial equilibrium focus only briefly addressed election campaign responses to information dissemination. While such responses are often documented (Banerjee et al. 2011; Cruz, Keefer and Labonne 2019), little is still known about whether or when politicians influence political accountability by complementing or refuting information dissemination campaigns. Third, our finding that voters are keen to take advantage of a rare non-electoral opportunity to engage with incumbents suggests that a lack of contact opportunities may also limit political accountability. NGOs and policymakers seeking to increase political accountability might therefore consider facilitating more of these opportunities to maximize the effectiveness of information dissemination campaigns. Finally, by focusing on information provision within election campaigns, we abstracted from the incentives and external constraints driving incumbent performance in the first place and the processes of candidate (de)selection that dictate the candidates that voters ultimately choose between. Each factor could constrain political accountability at an earlier stage and merits further examination.

Although rural areas where deputies from the president's party generally win may be exposed to fewer competing political influences, there are good reasons to believe that our anatomy of political accountability extends beyond our sample and intervention. First, sophisticated responses

among voters with low educations levels suggest that voters across the world could draw similar inferences. Second, parliamentary elections in Senegal share many features with elections in other developing democracies, including the dominance of the president's party and non-trivial levels of clientelism. Since information is likely to have weaker effects in contexts where electoral competition is more limited (Grossman and Michelitch 2018), information's capacity to enhance bottom-up accountability may be larger where party switching is common and national waves influence voting behavior less. Third, because our leaflets are similar in design to previous studies (Chong et al. 2015; Dunning et al. 2019; Gottlieb 2016; Humphreys and Weinstein 2012), our findings may help to direct researchers in other contexts toward the types of impediments to electoral accountability that we highlight here.

Nevertheless, several features of our study's context merit further research to explore the generality and broader implications of our findings. First, since current incumbents' performance generally exceeded both voters' prior beliefs and previous incumbents' performance in our sample, it is natural to wonder whether voters would respond differently to unexpectedly or comparably poor performance. Although future studies should explore this more extensively, Appendix section I.4 provides evidence that voters updated negatively and sought to sanction the current incumbent in Oussouye—a department where the previous incumbent's local performance exceeded the current incumbent's. Second, it is similarly important to establish whether voters regard *across*-party temporal benchmarks, like those that we provided, as more relevant than *within*-party benchmarks that filter out party-level factors. This may be pertinent in regimes where elections effectively select candidates from within the dominant party. Finally, further research should explore exactly what citizens demand when they perceive their representatives to be more responsive or effective, and if they ultimately receive it. Whether this constitutes standard distributive politics is important to understand for governments and donors, who often seek to use accountability mechanisms to promote programmatic policymaking.

References

Abadie, Alberto, Susan Athey, Guido W. Imbens and Jeffrey Wooldridge. 2017. "When should you adjust standard errors for clustering?" NBER Working Paper 24003, https://www.nber.org/papers/w24003 (accessed June 26, 2020).

Adida, Claire, Jessica Gottlieb, Gwyneth McClendon and Eric Kramon. forthcoming. "Breaking the Clientelistic Equilibrium: The Joint Importance of Salience and Coordination." *Comparative Political Studies* .

Aker, Jenny C., Paul Collier and Pedro C. Vicente. 2017. "Is information power? Using mobile phones and free newspapers during an election in Mozambique." *Review of Economics and Statistics* 99(2):185–200.

Anderson, Michael L. 2008. "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103(484):1481–1495.

Arias, Eric, Horacio Larreguy, John Marshall and Pablo Querubín. 2018*a*. "Does the Content and Mode of Delivery of Information Matter for Electoral Accountability? Evidence from a Field Experiment in Mexico." Working paper, https://tinyurl.com/y6tkfb7p (accessed June 26, 2020).

Arias, Eric, Horacio Larreguy, John Marshall and Pablo Querubín. 2018*b*. "Priors rule: When do malfeasance revelations help and hurt incumbent parties?" NBER Working Paper 24888, https://www.nber.org/papers/w24888 (accessed June 26, 2020).

Arias, Eric, Pablo Balán, Horacio Larreguy, John Marshall and Pablo Querubín. 2019. "Information provision, voter coordination, and electoral accountability: Evidence from Mexican social networks." *American Political Science Review* 113(2):475–498.

Aytaç, Selim Erdem. 2018. "Relative Economic Performance and the Incumbent Vote: A Reference Point Theory." *Journal of Politics* 80(1):16–29.

Banerjee, Abhijit V., Selvan Kumar, Rohini Pande and Felix Su. 2011. "Do Informed Voters Make Better Choices? Experimental Evidence from Urban India." Working paper, https://epod.cid.harvard.edu/sites/default/files/2018-02/do_informed_voters_make_better_choices_experimental_evidence_from_urban_india.pdf (accessed June 26, 2020).

Beck, Linda J. 2012. 'Senegal'. In *Countries at the Crossroads 2011: An Analysis of Democratic Governance*, ed. Jake Dizard, Christopher Walker and Vanessa Tucker. Rowman & Littlefield.

Bishop, Christopher M. 2006. *Pattern recognition and machine learning*. Springer.

Boas, Taylor C., F. Daniel Hidalgo and Marcus André Melo. 2019. "Norms versus action: Why voters fail to sanction malfeasance in Brazil." *American Journal of Political Science* 63(2):385–400.

Bussell, Jennifer. 2019. *Clients and Constituents: Political Responsiveness in Patronage Democracies*. Oxford University Press.

Campello, Daniela and Cesar Zucco, Jr. 2016. "Presidential success and the world economy." *Journal of Politics* 78(2):589–602.

Casey, Katherine. 2018. "Radical Decentralization: Does community driven development work?" *Annual Review of Economics* 10:139–165.

Chong, Alberto, Ana De La O, Dean Karlan and Leonard Wantchekon. 2015. "Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice and Party Identification." *Journal of Politics* 77(1):55–71.

Cruz, Cesi, Philip Keefer and Julien Labonne. 2019. "Buying Informed Voters: New Effects of Information on Voters and Candidates." Working paper, https://julienlabonne.files.wordpress.com/2020/04/ppcrv_web_042020.pdf (accessed June 26, 2020).

Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan Hyde, Craig McIntosh and Gareth Nellis. 2019. *Information, Accountability, and Cumulative Learning: Lessons from Metaketa I*. Cambridge University Press.

Fearon, James D. 1999. Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance. In *Democracy, Accountability, and Representation*, ed. Adam Przeworski, Susan Stokes and Bernard Manin. Cambridge University Press.

Ferraz, Claudio and Frederico Finan. 2008. "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes." *Quarterly Journal of Economics* 123(2):703–745.

Gomez, Brad T. and J. Matthew Wilson. 2006. "Cognitive heterogeneity and economic voting: A comparative analysis of four democratic electorates." *American Journal of Political Science* 50(1):127–145.

Gottlieb, Jessica. 2016. "Greater Expectations: A Field Experiment to Improve Accountability in Mali." *American Journal of Political Science* 60(1):143–157.

Grossman, Guy and Kristin Michelitch. 2018. "Information dissemination, competitive pressure, and politician performance between elections: A field experiment in Uganda." *American Political Science Review* 112(2):280–301.

Grossman, Guy, Macartan Humphreys and Gabriella Sacramone-Lutz. 2014. ""I wld like u WMP to extend electricity 2 our village": On Information Technology and Interest Articulation." *American Political Science Review* 108(3):688–705.

Humphreys, Macartan and Jeremy Weinstein. 2012. "Policing Politicians: Citizen Empowerment and Political Accountability in Uganda Preliminary Analysis." Working paper, https://tinyurl.com/y4h6uasz (accessed June 26, 2020).

Kendall, Chad, Tommaso Nannicini and Francesco Trebbi. 2014. "How do voters respond to information? Evidence from a randomized campaign." *American Economic Review* 105(1):322–353.

Koter, Dominika. 2013. "King makers: Local leaders and ethnic politics in Africa." *World Politics* 65(2):187–232.

Larreguy, Horacio and Shelley Liu. 2017. "The Effect of Education on Political Participation: Evidence from a Consolidated Developing Democracy." Working Paper, https://tinyurl.com/yy94to68 (accessed June 26, 2020).

Lieberman, Evan S., Daniel N. Posner and Lily L. Tsai. 2014. "Does information lead to more active citizenship? Evidence from an education intervention in rural Kenya." *World Development* 60:69–83.

McKenzie, David. 2012. "Beyond baseline and follow-up: The case for more T in experiments." *Journal of Development Economics* 99(2):210–221.

Meyer, Margaret A. and John Vickers. 1997. "Performance comparisons and dynamic incentives." *Journal of Political Economy* 105(3):547–581.

Powell, Jr., G. Bingham and Guy D. Whitten. 1993. "A Cross-National Analysis of Economic Voting: Taking Account of the Political Context." *American Journal of Political Science* 37(2):391–414.

Thomas, Melissa A. and Oumar Sissokho. 2005. "Liaison Legislature: The Role of the National Assembly in Senegal." *Journal of Modern African Studies* 43(1):97–117.

Zaller, John R. 1992. *The Nature and Origins of Mass Opinion*. Cambridge University Press.

SUPPORTING INFORMATION FOR:

"ABLE AND MOSTLY WILLING: AN EMPIRICAL ANATOMY OF INFORMATION'S EFFECT ON VOTER-DRIVEN ACCOUNTABILITY IN SENEGAL"

Contents

A Overview of our theoretical framework

This section more formally summarizes our simple learning framework used to dissect how voters engage in political accountability. With respect to *electoral* accountability, we consider an expressive voter i that implements the following decision rule:

$$v_i\big(\mathbb{E}_i[q_I],\mathbb{E}_i[q_C],w_i,V_i,c_i\big)=\begin{cases}\text{vote } I & \text{if } \ w_i f_i\Big(\mathbb{E}_i[q_I],\mathbb{E}_i[q_C]\Big)+(1-w_i)V_i\geq c_i\\ \text{vote } C & \text{if } \ w_i f_i\Big(\mathbb{E}_i[q_I],\mathbb{E}_i[q_C]\Big)+(1-w_i)V_i\leq -c_i\\ \text{abstain} & \text{if } \ \Big|w_i f_i\Big(\mathbb{E}_i[q_I],\mathbb{E}_i[q_C]\Big)+(1-w_i)V_i\Big|<c_i\end{cases} \tag{A1}$$

where the function $f_i(\cdot)$ increases with i's expectation of incumbent I's underlying "quality," $\mathbb{E}_i[q_I]$, and decreases with i's expectation of challenger C's underlying quality, $\mathbb{E}_i[q_C]$, and V_i is the relative utility i receives from voting for I over C from all other factors. Voter i attaches weight $w_i\in[0,1]$ to relative expectations about quality, and weight $(1-w_i)$ to other factors entering their voting calculus. If the magnitude of this weighted average of expressive benefits is positive (negative) and exceeds cost $c_i\geq 0$ of turning out, i will vote for I (C). This simple model implies that information that alters prior beliefs about incumbent quality—relative to challenger quality, and on issues that matter to voters—can alter vote choice and turnout, especially where w_i is large. The following subsection illustrates how Bayesian voters update in a Normal learning model.

With respect to *non-electoral* efforts to hold incumbents to account after elections, we instead propose that voters engage in costly efforts to contact an incumbent when the expected benefits of responsiveness exceed the cost $e_i\geq 0$ of seeking to make a request. Specifically, voter i implements the following rule:

$$r_i\big(\mathbb{E}_i[q_I],e_i\big)=\begin{cases}\text{make request from } I & \text{if } \ g_i(\mathbb{E}_i[q_I])\geq e_i\\ \text{no request} & \text{if } \ g_i(\mathbb{E}_i[q_I])<e_i\end{cases} \tag{A2}$$

where the benefits function $g_i(\cdot)$ increases with expected incumbent quality. Voters thus seek to contact incumbents when they expect a high probability of action or a more effective action by the incumbent on the voter's behalf. In contrast with voting, w_i does not influence non-electoral efforts to contact incumbents.

B The effect of providing incumbent-only and benchmarked information

We formally derive the effects of providing incumbent-only and temporally-benchmarked information, relative to each other and to receiving no information at all, on Bayesian voters' behavior. We adopt a simple Normal learning framework where a given voter learns about the current incumbent t's unobserved underlying quality, the previous incumbent $t-1$'s unobserved underlying quality, and an unobserved time-invariant district/department-specific component of performance that affects all incumbents within the district equally. Specifically, denote the voter's prior belief about the current incumbent's quality as q_t, the voter's prior belief about the previous incumbent's

quality as q_{t-1}, and the voter's prior belief about the district-specific component as q_c. We assume that our voter's prior beliefs over these quantities are given by $N(\theta_t, 1/p_t)$, $N(\theta_{t-1}, 1/p_{t-1})$, and $N(\theta_C, 1/p_c)$ respectively. We assume for simplicity that draws from these distributions are independent,[20] but impose no restrictions on the mean or precision of these Normal prior distributions. The central tendency of these priors distributions could thus be very similar or different across current and previous incumbents and may or may not reflect information that the voter may already have incorporated about current or previous incumbent performance.

Our goal is to examine the differential effect of different types of incumbent performance signal on voters' absolute and relative posterior beliefs. These are two key outcomes in our empirical analysis that directly influence vote choice and, in the case of the level, seeking to contact incumbents in our simple decision-theoretic model—equations (A1) and (A2) above.

B.1 Incumbent-only performance information

We start with the baseline case where a representative voter receives a given realization of the incumbent-only performance signal, $\hat{s}_t$, drawn from signal distribution $N(q_t + q_c, 1/\rho_t)$, where the signal's precision ρ_t is known to the voter but its expectation $q_t + q_c$ is not. The performance signal thus reflects both the unobserved quality of the current incumbent and unobserved time-invariant characteristics of the district/department. The following proposition establishes the voter's posterior inferences about incumbent's quality q_t and the district-specific shock q_c:

Proposition 1. *(Incumbent-only performance information) Upon receiving realized signal $\hat{s}_t$, a voter's posterior expectation of current incumbent t's quality is $w_t(\hat{s}_t - \theta_c) + (1 - w_t)\theta_t$ and of the district-specific shock is $w_c(\hat{s}_t - \theta_t) + (1 - w_c)\theta_t$, where w_t and w_c are weights (defined within the proof) that both increase with ρ_t and respectively increase in p_c and p_t.*

Proof: We first define $\mathbf{q} = [q_t, q_c]'$, $\mu = [\theta_t, \theta_c]'$, $\mathbf{\Lambda}^{-1} = \begin{bmatrix} 1/p_t & 0 \\ 0 & 1/p_c \end{bmatrix}$, $\mathbf{A} = [1, 1]$, and $\mathbf{L}^{-1} = [1/\rho_t]$. Applying a standard multivariate updating result (e.g. Bishop 2006:93) implies that posterior beliefs are distributed according to:

$$p(\mathbf{q}|\hat{s}_t) \sim N\Big((\mathbf{\Lambda} + \mathbf{A}'\mathbf{L}\mathbf{A})^{-1}(\mathbf{A}'\mathbf{L}\hat{s}_t + \mathbf{\Lambda}\mu), (\mathbf{\Lambda} + \mathbf{A}'\mathbf{L}\mathbf{A})^{-1}\Big), \tag{B1}$$

where the application of matrix operations to the model in hand implies:

$$(\mathbf{\Lambda} + \mathbf{A}'\mathbf{L}\mathbf{A})^{-1} = \begin{bmatrix} p_t + \rho_t & \rho_t \\ \rho_t & p_c + \rho_t \end{bmatrix}^{-1} = \frac{1}{p_t p_c + p_t \rho_t + p_c \rho_t} \begin{bmatrix} p_c + \rho_t & -\rho_t \\ -\rho_t & p_t + \rho_t \end{bmatrix} := \mathbf{\Sigma},$$

$$(\mathbf{A}'\mathbf{L}\hat{s}_t + \mathbf{\Lambda}\mu) = \begin{bmatrix} \rho_t \hat{s}_t + p_t \theta_t \\ \rho_t \hat{s}_t + p_c \theta_c \end{bmatrix}. \tag{B2}$$

[20] Relaxing the independence assumption would alter the relative weights attached to the signals and the priors in the propositions below.

Combining these results yields probability distribution:

$$p(\mathbf{q}|\hat{s}_t) \sim N\left(\begin{bmatrix} w_t(\hat{s}_t - \theta_c) + (1-w_t)\theta_t \\ w_c(\hat{s}_t - \theta_t) + (1-w_c)\theta_c \end{bmatrix}, \mathbf{\Sigma}\right), \tag{B3}$$

where $w_t := \frac{p_c\rho_t}{p_tp_c+p_t\rho_t+p_c\rho_t}$ and $w_c := \frac{p_t\rho_t}{p_tp_c+p_t\rho_t+p_c\rho_t}$. ■

This first proposition demonstrates that current incumbent-only performance information influences the voter's beliefs about current incumbent quality to the extent that the district-specific shock-adjusted signal $(\hat{s}_t - \theta_c)$ differs from the voter's prior belief θ_t. Since the district-specific shock is also unobserved from the perspective of the voter, the voter has limited capacity to update about the value of this shock, and thus relies on their prior belief θ_c. Indeed, relative to receiving no information about incumbent performance, and thus retaining the prior belief θ_t, a voter upwardly (downwardly) updates their expectation of t's quality when $\hat{s}_t - \theta_c > (<)\theta_t$. This shows that, after netting out prior expectations of the district-specific shock, the voter updates favorably about the incumbent when the signal exceeds their prior expectation. If the signal is believed to be uncorrelated with the previous incumbent's type (as we assume for simplicity below), the same expression pertains to evaluating the posterior belief regarding the expected *difference* in t's quality relative to previous incumbent $t-1$'s quality. If the voter believes that their posterior belief about the previous incumbent is uninformative about the challenger or regards the previous incumbent as a proxy for challengers,[21] we thus expect that providing incumbent performance information will increase (decrease) the incumbent's vote share, relative to receiving no information, when $\hat{s}_t - \theta_c > (<)\theta_t$.

B.2 Benchmarked malfeasance information

Turning to our main result, we now consider the the effect of adding a benchmark, such that the voter receives a performance signal $\hat{s}_{t-1}$ pertaining to the previous incumbent *as well as* the current incumbent performance signal $\hat{s}_t$. We similarly assume that $\hat{s}_{t-1}$ is drawn from signal distribution $N(q_{t-1} + q_c, 1/\rho_{t-1})$, where the signal's precision ρ_{t-1} is again known to the voter. We assume for simplicity that the realized signals are uncorrelated.[22] This second signal enables the voter to draw more precise inferences by filtering out their more accurate updated beliefs about the district-specific component of performance, as well as potentially learn more about the performance of previous incumbents that may be informative about current challengers.

The following proposition now establishes the voter's posterior beliefs following the provision of such benchmarked performance information:

Proposition 2. *(Benchmarked performance information) Upon receiving realized signals $\hat{s}_t$ and $\hat{s}_{t-1}$, a voter's posterior expectation of current incumbent t's quality is $w_{t,s}\hat{s}_t - w_{t,c}\theta_c - w_{t,\Delta}(\hat{s}_{t-1} - \theta_{t-1}) + w_{t,t}\theta_t$, of the previous incumbent's $t-1$'s quality is $w_{t-1,s}\hat{s}_{t-1} - w_{t-1,c}\theta_c - w_{t-1,\Delta}(\hat{s}_t - \theta_t) + w_{t-1,t-1}\theta_{t-1}$, and of the district-specific shock is $w_{c,t}(\hat{s}_t - \theta_t) + w_{c,t-1}(\hat{s}_{t-1} - \theta_{t-1}) + w_{c,c}\theta_c$, where the weights are defined within the proof.*

[21] The latter is plausible in the context of our experimental sample, where all incumbent legislators were from the current president's BBY coalition and all benchmark legislators were from BBY's opposition, the previous president's Sopi coalition.

[22] Relaxing this assumption would alter the relative weights attached to the signals and prior beliefs in computing the voter's posterior belief in Proposition 2.

Proof: We first define $\hat{\mathbf{s}} = [\hat{s}_t, \hat{s}_{t-1}]'$, $\mathbf{q} = [q_t, q_{t-1}, q_c]'$, $\mu = [\theta_t, \theta_{t-1}, \theta_c]'$, $\mathbf{\Lambda}^{-1} = \begin{bmatrix} 1/p_t & 0 & 0 \\ 0 & 1/p_{t-1} & 0 \\ 0 & 0 & 1/p_c \end{bmatrix}$, $\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$, and $\mathbf{L}^{-1} = \begin{bmatrix} 1/\rho_t & 0 \\ 0 & 1/\rho_{t-1} \end{bmatrix}$. We then apply the same theorem as in the previous proof, where the application of matrix operations to the model in hand implies:

$$\begin{aligned}
(\mathbf{\Lambda} + \mathbf{A}'\mathbf{L}\mathbf{A})^{-1} &= \begin{bmatrix} p_t + \rho_t & 0 & \rho_t \\ 0 & p_{t-1} + \rho_{t-1} & \rho_{t-1} \\ \rho_t & \rho_{t-1} & p_c + \rho_t + \rho_{t-1} \end{bmatrix}^{-1} \\
&= \frac{1}{p_t\rho_I(p_{t-1} + \rho_{t-1}) + p_{t-1}\rho_{t-1}(p_t + \rho_t) + p_c(p_t + \rho_t)(p_{t-1} + \rho_{t-1})} \\
&\quad \times \begin{bmatrix} (p_{t-1} + \rho_{t-1})(p_c + \rho_t) + p_{t-1}\rho_{t-1} & \rho_t\rho_{t-1} & -\rho_t(p_{t-1} + \rho_{t-1}) \\ \rho_t\rho_{t-1} & (p_t + \rho_t)(p_c + \rho_{t-1}) + p_t\rho_t & -\rho_{t-1}(p_t + \rho_t) \\ -\rho_t(p_{t-1} + \rho_{t-1}) & -\rho_{t-1}(p_t + \rho_t) & (p_t + \rho_t)(p_{t-1} + \rho_{t-1}) \end{bmatrix} \\
&:= \mathbf{\Sigma}_B, \qquad \text{(B4)}
\end{aligned}$$

$$(\mathbf{A}'\mathbf{L}\hat{\mathbf{s}} + \mathbf{\Lambda}\mu) = \begin{bmatrix} \rho_t\hat{s}_t + p_t\theta_t \\ \rho_{t-1}\hat{s}_{t-1} + p_{t-1}\theta_{t-1} \\ \rho_t\hat{s}_t + \rho_{t-1}\hat{s}_{t-1} + p_c\theta_c \end{bmatrix}. \qquad \text{(B5)}$$

Combining these results yields the probability distribution:

$$p(\mathbf{q}|\hat{s}_t, \hat{s}_{t-1}) \sim N\left(\begin{bmatrix} w_{t,s}\hat{s}_t - w_{t,c}\theta_c - w_{t,\Delta}(\hat{s}_{t-1} - \theta_{t-1}) + w_{t,t}\theta_t \\ w_{t-1,s}\hat{s}_{t-1} - w_{t-1,c}\theta_c - w_{t-1,\Delta}(\hat{s}_t - \theta_t) + w_{t-1,t-1}\theta_{t-1} \\ w_{c,t}(\hat{s}_t - \theta_t) + w_{c,t-1}(\hat{s}_{t-1} - \theta_{t-1}) + w_{c,c}\theta_c \end{bmatrix}, \mathbf{\Sigma}_B \right), \qquad \text{(B6)}$$

where the weights are given by $w_{t,s} := \frac{\rho_t(p_{t-1}p_c + p_c\rho_{t-1} + p_{t-1}\rho_{t-1})}{D}$, $w_{t,c} := \frac{p_c\rho_t(p_{t-1}+\rho_{t-1})}{D}$, $w_{t,\Delta} := \frac{p_{t-1}\rho_t\rho_{t-1}}{D}$, $w_{t,t} := \frac{p_t(p_{t-1}p_c + p_c\rho_{t-1} + p_{t-1}\rho_{t-1} + p_{t-1}\rho_t + \rho_t\rho_{t-1})}{D}$, $w_{t-1,s} := \frac{\rho_{t-1}(p_tp_c + p_c\rho_t + p_t\rho_t)}{D}$, $w_{t-1,c} := \frac{p_c\rho_{t-1}(p_t+\rho_t)}{D}$, $w_{t-1,\Delta} := \frac{p_t\rho_t\rho_{t-1}}{D}$, $w_{t-1,t-1} := \frac{p_{t-1}(p_tp_c + p_c\rho_t + p_t\rho_t + p_t\rho_{t-1} + \rho_t\rho_{t-1})}{D}$, $w_{c,t} := \frac{p_t\rho_t(p_{t-1}+\rho_{t-1})}{D}$, $w_{c,t-1} := \frac{p_{t-1}\rho_{t-1}(p_t+\rho_t)}{D}$, and $w_{c,c} := \frac{p_c(p_t+\rho_t)(p_{t-1}+\rho_{t-1})}{D}$, where $D := [p_t\rho_t(p_{t-1} + \rho_{t-1}) + p_{t-1}\rho_{t-1}(p_t + \rho_t) + p_c(p_t + \rho_t)(p_{t-1} + \rho_{t-1})]$ and all weights are positive. ■

This proposition demonstrates that the voter's posterior beliefs about the level of current incumbent quality increase with the extent to which indicators of performance exceed expectations that now explicitly adjust for updated beliefs about the district-specific shock. In contrast with incumbent-only information, a Bayesian voter now also uses the benchmarked signal to better account for the possibility that high incumbent performance could reflect a high realization of the district-specific common shock, i.e. $\hat{s}_{t-1} - \theta_{t-1}$. Consequently, relative to receiving no information, benchmarked performance information will induce upward (downward) updating when: $w_{t,s}\hat{s}_t - w_{t,c}\theta_c - w_{t,\Delta}(\hat{s}_{t-1} - \theta_{t-1}) + w_{t,t}\theta_t (<) > \theta_t$. The voter will thus update favorably about the current incumbent when performance indicators, adjusted for updated expectations of the district-specific shock, exceed prior expectations of quality. The same logic applies to evaluations of the previous incumbent. The voter's posterior belief about the common shock itself, $w_{c,t}(\hat{s}_t - \theta_t) + w_{c,t-1}(\hat{s}_{t-1} - \theta_{t-1}) + w_{c,c}\theta_c$, is intuitively increasing in the extent to which perfor-

mance signals exceed prior expectations.

B.3 Comparing posterior beliefs under incumbent-only and benchmarked signals

Combining our two propositions, benchmarked information induces a more favorable (unfavorable) posterior expectation of incumbent quality than an incumbent-only signal when:

$$w_{t,s}\hat{s}_t - w_{t,c}\theta_c - w_{t,\Delta}(\hat{s}_{t-1} - \theta_{t-1}) + w_{t,t}\theta_t > (<) w_t(\hat{s}_t - \theta_c) + (1 - w_t)\theta_t. \tag{B7}$$

There are thus two primary forces pushing a voter to update favorably about the current incumbent upon receiving a benchmarked performance signal in addition to an incumbent performance signal: (1) the increased weight attached to the current incumbent's performance signal (it is easy to show that $w_{t,s} > w_t$), when $\hat{s}_t > \theta_t$; and (2) when the weights attached to the signal and prior beliefs do not drastically differ (i.e. $w_t \approx w_{t,s} \approx w_{t,c} \approx 1 - w_{t,t}$, and thus $\hat{s}_t$, θ_c, and θ_t cancel out), voters will generally update favorably when $\hat{s}_{t-1} < \theta_{t-1}$, i.e. when the previous incumbent performed worse than expected. Force (1) reflects the sharper inferences that can be drawn from a given signal when it is benchmarked, while force (2) reflects the second signal inducing the voter to infer that there was a larger-than-expected district-specific shock.

Turning to relative evaluations and vote choice, voter behavior is instead likely to reflect a *relative* comparison between posterior beliefs about the current incumbent and current *challengers*. If the voter does not update about current challengers from the information received, then benchmarked information will increase the incumbent's vote share to the extent that voters' posterior beliefs are more favorable than under incumbent-only information, as in equation (B7). However, voters may associate challengers with previous incumbents, especially in contexts like ours where the previous incumbents were all from what is now the main opposition party facing the BBY's current incumbents. If the signals received are informative about current challengers, benchmarked information induces a larger difference in expected quality between the current incumbent and the previous incumbent—which may proxy for challengers, when it comes to vote choice—relative to incumbent-only information when:

$$w^*_{t,s}\hat{s}_t + w_{t,t}\theta_t - w^*_{t-1,s}\hat{s}_{t-1} - w_{t-1,t-1}\theta_{t-1} > w_t(\hat{s}_t - \theta_c - \theta_{t-1}) + (1 - w_t)(\theta_t - \theta_{t-1}), \tag{B8}$$

where the district-specific shock is identically accounted for when comparing posterior beliefs about the current and previous incumbents (but adjusts the weighting coefficients to account for extracting the district-specific shock). As with absolute beliefs, a voter will become relatively more favorable toward the incumbent approximately when: (1) $w^*_{t,s} > w_t$, where $\hat{s}_t > \theta_t$; and (2) $\hat{s}_{t-1} < \theta_{t-1} + \theta_c$, where the weights on comparable terms are similar in magnitude.

B.4 Empirical implications

With respect to absolute posterior beliefs, we expect to observe the following predictions and comparative statics:

- The effect of incumbent-only information vs. no information on overall beliefs about current incumbent quality is positive when $\hat{s}_t > \theta_t + \theta_c$, and is increasing in $(\hat{s}_t - \theta_t)$.

- The effect of benchmarked information vs. no information on overall beliefs about current incumbent quality is positive when, approximately, $\hat{s}_t > \theta_t + \mathbb{E}[q_c|\hat{s}_t, \hat{s}_{t-1}]$, and is increasing in $(\hat{s}_t - \theta_t)$ and decreasing in $(\hat{s}_{t-1} - \theta_{t-1})$.

- The effect of benchmarked information vs. incumbent-only information on beliefs about current incumbent quality is positive when, approximately, $\hat{s}_{t-1} < \theta_{t-1}$, and is increasing in $(\hat{s}_t - \theta_t)$ and decreasing in $(\hat{s}_{t-1} - \theta_{t-1})$.

"Approximate" relationships are cases where weights are assumed not to meaningfully differ.

With respect to relative comparisons between current and previous incumbents—a plausible proxy for challengers, and thus vote choices—incumbent-only and benchmarked information provision implies the following comparative statics:

- The effect of incumbent-only information vs. no information on incumbent vote share is positive when $\hat{s}_t > \theta_t + \theta_c$, and is increasing in $(\hat{s}_t - \theta_t)$.

- The effect of benchmark information vs. no information on incumbent vote share is positive when, approximately, $\hat{s}_t > \hat{s}_{t-1}$, and is decreasing in $(\hat{s}_t - \hat{s}_{t-1})$.

- The effect of benchmark vs. incumbent-only information on incumbent vote share is positive when, approximately, $\hat{s}_{t-1} < \theta_{t-1} + \theta_c$, and is increasing in $(\hat{s}_t - \theta_t)$ and decreasing in $(\hat{s}_{t-1} - \theta_{t-1})$.

C Senegalese parliamentary electoral and non-electoral political engagement in comparative context

Figure C1 reports national turnout rates across sub-Saharan Africa in the most recent parliamentary elections, while Figure C2 shows the percentage of Afrobarometer (round 6) respondents that have contacted a member of parliament within the last year. Consistent with the limited role of the Assemblée Nationale, turnout rates are relatively low in Senegalese parliamentary elections.

D Additional details about sample selection

D.1 Selection of departments

The five departments—Fatick, Foundiougne, Kanel, Oussouye, and Ranérou Ferlo—were selected because they satisfied four criteria that prior theoretical arguments suggest would increase the likelihood of performance information helping voters hold incumbents to account: (1) only a single incumbent was seeking re-election through the majoritarian vote (with the exception of Kanel where two were standing); (2) there were no incumbents from the proportional list attached to the department (with the exception of Kanel); (3) the incumbent's performance could be compared with the previous incumbent(s), because no incumbent was seeking re-election for a second time and the department was not a newly-created administrative unit; and (4) given the preceding criteria, the selected departments have the lowest number of deputies representing the department. Oussouye and Ranérou Ferlo had only one incumbent deputy, although Oussouye had two in the

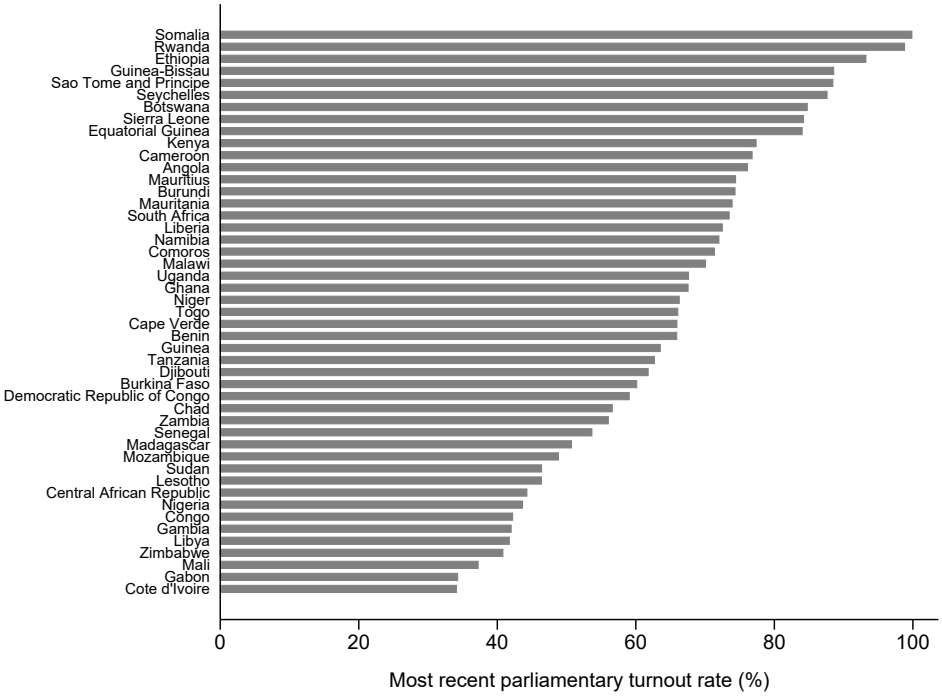

Figure C1: Parliamentary turnout rates in the most recent election across sub-Saharan Africa

Note: All data was downloaded from International Institute for Democracy and Electoral Assistance.

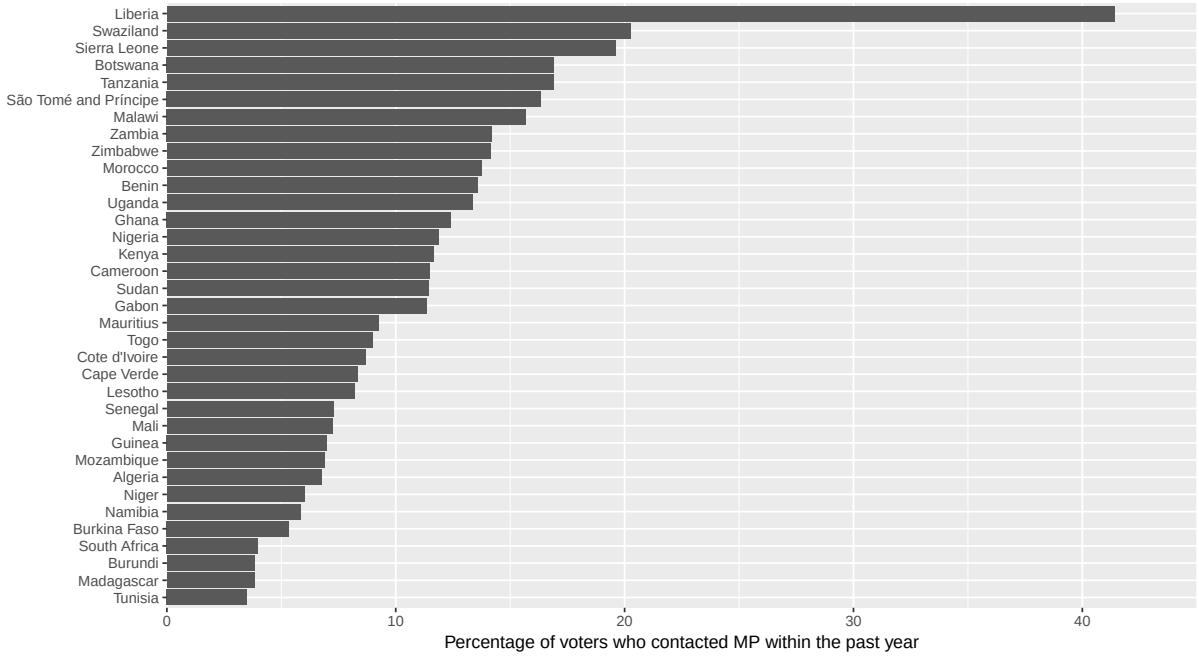

Figure C2: Parliamentary deputy contact rates across sub-Saharan Africa

Note: All data is from Afrobarometer round 6.

previous legislature. Fatick, Foundiougne, and Kanel had two majoritarian deputies. Deputies from the proportional list are not assigned to particular departments.

D.2 Selection of villages

Across our five departments, we selected 450 rural villages for our sample. Starting from the 859 possible villages in these departments, we excluded all villages with fewer than 200 people and all villages with more than 4,000 people. Logistical concerns and access to newly-constructed schools further restricted the set of potential villages. Of logistical concerns, 19 villages were dropped because they were too expensive to reach, e.g. because they are located on islands. In the hope of leveraging cross-cohort variation in access to schooling following a 2002 school construction program (to instrument for educational attainment and identify heterogeneous effects by educational attainment, see more below), we excluded villages where the first post-2002 school was built between 2006 and 2010. We ignore this cross-cutting variation because access to new schools did not robustly increase educational attainment among our survey respondents. By virtue of our randomization, access to schooling is orthogonal to our informational treatments.

D.3 Selection of young voters as respondents

Our sampling strategy stratified the sample into three age groupings of roughly equal size within each village (i.e. 3 respondents per village from each category): 20-26, 27-31, and 32-38. The only logistical restriction was that respondents must have a cellphone number, which virtually all young Senegalese satisfy. Eligible citizens were identified and located with the assistance of the village chief, and response rates were high (as demonstrated by very low rates of endline attrition). The sampling strategy was not a function of treatment. No enumerator re-interviewed a respondent that they interviewed at baseline, and enumerators were not informed of endline respondents' treatment status.

This sampling strategy reflected our pre-registered intention to examine the effects of educational attainment. In particular, we aimed to leverage a difference-in-differences or regression discontinuity design exploiting cross-cohort variation in access to schools constructed as part of a 2002 secondary schooling expansion program as an instrument for educational attainment (Larreguy and Liu 2017). The 20-26 age grouping contained cohorts that were counted as fully treated if a school been constructed within 6km of their village, while the 27-31 age grouping comprised partially treated respondents (who were already in secondary school at the time of the reform) if a school has been constructed within 6km of their village, and the 32-38 category is a control group. Unfortunately, we were unable to obtain a first stage showing that the instrument robustly increased educational attainment using either the difference-in-differences or regression discontinuity approaches.

D.4 Sample characteristics relative to national averages

Table D1 compares 2013 Census data in our sample of 450 villages with the Senegalese national averages.

Table D1: Sample summary statistics

	Data from our 2017 sample			Data from Senegal 2013 census			
	Weighted by...			Entire	From our	20-38 age group	20-38 age group,
Variable	Unweighted	Population	Baseline resps	census	villages	all villages	from our villages
Average age	28.51	28.59	28.51	22.08	21.40	27.47	27.65
% female	36.31	36.82	36.27	50.83	50.89	52.35	54.24
% with some primary education	52.58	57.25	52.40	33.79	25.92	34.57	20.98
% with some secondary education	35.32	37.59	35.21	14.10	5.39	24.36	10.50
% read/write French				37.28	29.01	35.40	21.36
% read/write Wolof				1.47	0.53	2.20	1.05
% read/write Pular				1.19	1.26	1.92	2.22
% read/write Serer				0.31	1.59	0.47	2.72
% read/write Mandingue				0.15	0.09	0.22	0.20
% read/write Diola				0.08	0.05	0.13	0.10
% read/write Soninke				0.03	0.04	0.05	0.05
% Muslim	89.14	89.32	89.15	95.66	89.80	95.42	90.93
% Christian	9.18	9.50	9.18	3.91	7.88	4.19	7.35
% with piped water	58.26	66.71	58.15	53.18	27.25	57.37	26.31
% with electricity	29.16	39.69	29.08	38.65	3.43	44.43	4.09
Average number of bedrooms	7.33	7.43	7.33	4.96	4.88	4.92	4.93
% from rural villages	100	100	100	59.27	100	53.43	100

E Additional information about information treatments

As noted above, some of the departments in our sample were represented by more than one legislative deputy in a given parliamentary session. To homogenize the form of our treatment, we provided information about only one current deputy seeking re-election and one deputy from the previous legislature. As explained in the main text, the deputies about which voters in a given randomization block were informed was completely randomized across randomization blocks within departments. Accordingly, all eligible deputies were shown to some voters, and the probability of a given deputy being reported on was equal among current incumbents and equal among previous incumbents within each department. Table E1 reports the distribution of blocks across each configuration of current and previous deputies.

To save space, Figure 4 in the main paper only provides an example of the "duties + benchmark" treatment. Figures E1-E4 illustrate the other four treatment conditions, again with examples from the department of Oussouye.

F Compliance and experimental validation checks

We were unable to conduct surveys in 7 of our 450 villages. In three cases we were refused entry, while the remaining cases reflected a lack of identity cards among villagers, inability to locate the village, heavy rain, and a village falling under judicial control. Given that we conduced surveys in all villages and did not allocate different treatment assignments to villages according to their characteristics, our inability to conduct surveys in these villages should be unaffected by treatment assignment.

We estimate equation (1) to demonstrate that treatment assignments for the 443 villages where surveys were conducted are indeed orthogonal to predetermined covariates, and thus that the ran-

Table E1: Distribution of current and previous deputies about which information was disseminated

		Current deputy:					
		Papa Biram Toure	Adama Sylla	Daouda Dia	Maïrame Kane	Aimé Assine	Aliou Demba Sow
Previous deputy:	Abdoulaye Ndour	Fatick (8 blocks)					
	Abdoulaye Sene	Fatick (8 blocks)					
	Fatou Diouf	Fatick (9 blocks)					
	Pape Dib Sarr		Foundiougne (12 blocks)				
	El Hadji Famara Senghor		Foundiougne (12 blocks)				
	Demba Diop			Kanel (4 blocks)	Kanel (3 blocks)		
	Bassirou Doro Ly			Kanel (3 blocks)	Kanel (4 blocks)		
	Sékou Sambou					Oussouye (5 blocks)	
	Ousame Ba						Ranérou Ferlo (7 blocks)

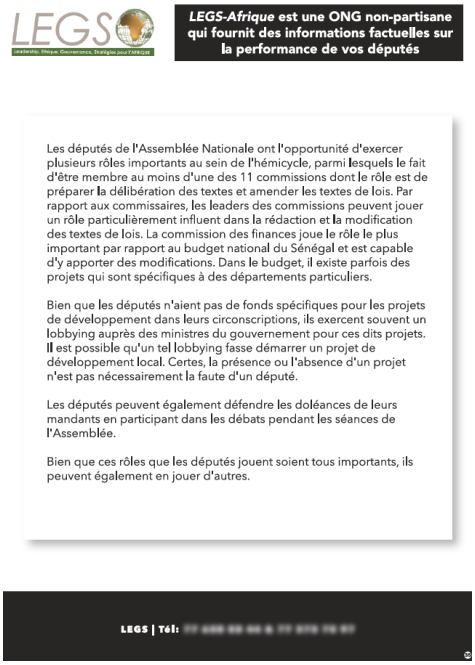

Figure E1: “duties” treatment

Figure E2: Example of "incumbent" treatment in Oussouye

Figure E3: Example of "duties + incumbent" treatment in Oussouye

Figure E4: Example of "benchmark" treatment in Oussouye

domization's integrity was maintained after dropping the seven villages that we could not access. Table F1 shows that 90 predetermined individual- and village-level covariates are well-balanced across treatment conditions at baseline: the two-sided joint F test of the restriction that each treatment group is indistinguishable from the others was rejected at the 10% level in 13 cases.

Analysis of endline data is generally more complex, since estimates using endline data could be confounded by selective attrition in response to treatment. As noted in the main text, we successfully re-interviewed 96% of the baseline sample. This remarkably high recontact rate for a telephone followup survey may have reflected the low frequency with which rural Senegalese voters have opportunities to express their views to survey teams, especially those that offered to pass on requests to politicians at baseline. Unsurprisingly, given this low rate of attrition, there is no significant difference in attrition rates across treatment groups (an F test of equality of endline responses across treatment conditions yielded a $p = 0.21$). Moreover, the balance we observed at baseline continues to hold within the endline response sample: for only 15 of 102 predetermined (baseline and endline) variables do we observe significant differences at the 10% level across treatment conditions.

G ICW index construction

Following Anderson (2008), a given inverse-covariance weighted index for individual i is defined by $(\mathbf{1}'\Sigma^{-1}\mathbf{1})^{-1}(\mathbf{1}'\Sigma^{-1}\tilde{\mathbf{x}}_i)$, where Σ is the $K \times K$ covariance matrix between items $x_{i1},...,x_{iK}$, $\tilde{\mathbf{x}}_i$ is the K-vector of standardized items, and $\mathbf{1}$ is a K-vector of 1s. The covariance matrix is computed

Table F1: Baseline sample balance tests

Outcome	Observations	Control mean	Control std. dev.	Duties		Incumbent		Incumbent and Duties		Benchmark		Benchmark and duties		F test (two-sided p value)
Individual-level predetermined baseline survey covariates														
Female (pre)	3,999	0.37	0.48	0.006	(0.031)	-0.027	(0.033)	0.006	(0.032)	-0.060*	(0.036)	-0.020	(0.033)	0.36
Married (pre)	3,999	0.66	0.48	-0.009	(0.030)	-0.056*	(0.029)	-0.034	(0.028)	-0.045	(0.029)	-0.012	(0.029)	0.31
Age (pre)	3,999	28.34	5.78	0.444*	(0.249)	0.173	(0.235)	0.133	(0.248)	0.204	(0.246)	0.300	(0.243)	0.59
Years of education (pre)	3,998	4.82	5.46	0.049	(0.350)	0.528	(0.354)	0.229	(0.346)	0.112	(0.356)	-0.191	(0.353)	0.36
Diola ethnicity (post)	3,999	0.07	0.25	0.006	(0.006)	-0.007	(0.012)	0.000	(0.005)	0.000	(0.008)	0.004	(0.006)	0.73
Pulaar ethnicity (post)	3,999	0.18	0.39	-0.021	(0.020)	0.002	(0.020)	-0.018	(0.021)	-0.022	(0.020)	-0.005	(0.021)	0.70
Peul ethnicity (post)	3,999	0.16	0.37	0.002	(0.020)	0.018	(0.024)	0.008	(0.024)	-0.020	(0.026)	0.003	(0.022)	0.79
Serer ethnicity (post)	3,999	0.41	0.49	-0.004	(0.025)	0.017	(0.028)	0.009	(0.023)	-0.016	(0.029)	-0.014	(0.025)	0.84
Toucouleur ethnicity (post)	3,999	0.01	0.08	-0.005	(0.006)	0.013	(0.010)	0.007	(0.006)	0.015*	(0.009)	-0.008	(0.007)	0.05**
Wolof ethnicity (post)	3,999	0.16	0.37	0.023	(0.023)	-0.027	(0.024)	0.000	(0.021)	-0.046**	(0.023)	0.029	(0.025)	0.07*
Christian (post)	3,997	0.10	0.30	-0.002	(0.022)	-0.014	(0.025)	-0.012	(0.023)	-0.011	(0.022)	-0.011	(0.023)	0.99
Muslim (post)	3,997	0.88	0.32	-0.001	(0.023)	0.027	(0.026)	0.023	(0.023)	0.000	(0.022)	0.010	(0.023)	0.75
Household has electricity (post)	3,999	0.32	0.47	0.007	(0.048)	-0.057	(0.049)	-0.041	(0.045)	-0.012	(0.046)	-0.053	(0.051)	0.62
Household has water (post)	3,999	0.61	0.49	0.014	(0.043)	-0.001	(0.042)	-0.066	(0.042)	-0.016	(0.043)	-0.048	(0.044)	0.38
Number of bedrooms (post)	3,889	7.08	5.22	0.638	(0.397)	0.406	(0.312)	0.300	(0.321)	0.206	(0.310)	0.420	(0.280)	0.59
Income scale (post)	3,454	1.72	1.86	-0.067	(0.114)	-0.265**	(0.106)	-0.208*	(0.113)	-0.165	(0.114)	-0.207*	(0.110)	0.12
Frequency discuss politics (pre)	3,996	2.06	0.80	-0.035	(0.046)	-0.014	(0.047)	-0.065	(0.049)	0.022	(0.045)	0.025	(0.046)	0.37
Interest in public affairs (pre)	3,999	1.97	1.01	0.049	(0.053)	0.082	(0.051)	-0.013	(0.054)	0.040	(0.051)	0.083*	(0.049)	0.31
Radio news frequency (pre)	3,999	4.01	2.23	-0.006	(0.146)	0.251*	(0.132)	-0.065	(0.138)	0.269**	(0.123)	0.059	(0.143)	0.01**
Television news frequency (pre)	3,999	2.42	2.49	0.139	(0.207)	0.165	(0.195)	-0.042	(0.197)	0.045	(0.197)	-0.087	(0.203)	0.69
Newspaper news frequency (pre)	3,999	0.67	1.62	-0.003	(0.093)	-0.069	(0.100)	-0.024	(0.102)	-0.038	(0.093)	-0.047	(0.092)	0.97
Satisfied with National Assembly (pre)	3,999	2.01	0.99	0.028	(0.060)	0.036	(0.058)	0.027	(0.059)	-0.047	(0.055)	0.080	(0.056)	0.33
Believe deputies listen to voters (pre)	3,999	0.58	0.73	0.033	(0.044)	0.030	(0.043)	0.032	(0.043)	-0.017	(0.045)	0.050	(0.047)	0.70
Believe deputies respond to requests (pre)	3,999	1.99	0.89	0.033	(0.053)	0.020	(0.055)	-0.067	(0.053)	-0.040	(0.052)	-0.036	(0.055)	0.28
Frequency of contacting deputy (pre)	3,999	0.13	0.48	-0.009	(0.023)	-0.013	(0.023)	-0.029	(0.025)	0.011	(0.030)	-0.016	(0.027)	0.81
Turnout in 2012 (pre)	3,999	0.42	0.49	-0.030	(0.029)	-0.043	(0.027)	-0.044	(0.029)	-0.049*	(0.028)	-0.031	(0.030)	0.57
Incumbent vote in 2012 (pre)	3,999	0.32	0.47	-0.049*	(0.026)	-0.066^{+}	(0.025)	-0.062**	(0.026)	-0.089^{+}	(0.024)	-0.064**	(0.027)	0.01**
Believe deputy is from own commune (pre)	3,997	0.28	0.45	-0.024	(0.033)	-0.058	(0.038)	-0.038	(0.037)	-0.070*	(0.039)	-0.034	(0.037)	0.52
Believe deputy is from own village (pre)	3,997	0.04	0.20	-0.006	(0.013)	0.005	(0.014)	0.006	(0.013)	0.036*	(0.022)	0.011	(0.018)	0.43
Believe deputy is of same ethnicity (pre)	3,990	0.57	0.50	-0.017	(0.028)	-0.039	(0.028)	-0.007	(0.028)	-0.067**	(0.031)	-0.054*	(0.031)	0.16
Know incumbent party (pre)	3,999	0.64	0.48	-0.056**	(0.026)	0.020	(0.026)	-0.005	(0.028)	-0.025	(0.028)	-0.029	(0.029)	0.10*
Know incumbent name (pre)	3,999	0.35	0.48	-0.021	(0.027)	-0.004	(0.027)	0.014	(0.028)	0.014	(0.029)	0.008	(0.028)	0.81
Know incumbent commune (pre)	3,999	0.66	0.47	-0.014	(0.035)	0.049	(0.035)	0.044	(0.037)	0.098^{+}	(0.037)	-0.021	(0.039)	0.00^{+}
Know incumbent village (pre)	3,999	0.91	0.29	-0.003	(0.021)	0.004	(0.021)	0.010	(0.020)	0.014	(0.021)	-0.021	(0.027)	0.84
Know incumbent ethnicity (pre)	3,999	0.54	0.50	-0.040	(0.029)	-0.036	(0.031)	-0.041	(0.032)	0.000	(0.033)	-0.016	(0.032)	0.58
Know deputies make laws (pre)	3,999	0.46	0.50	-0.027	(0.030)	0.023	(0.030)	0.025	(0.030)	0.040	(0.031)	0.004	(0.031)	0.35
Know deputies approve budget (pre)	3,999	0.54	0.50	0.020	(0.030)	0.018	(0.034)	0.000	(0.031)	0.011	(0.032)	-0.014	(0.032)	0.90
Know deputies do not select local projects (pre)	3,999	0.15	0.35	0.021	(0.022)	0.057**	(0.024)	0.045*	(0.025)	0.024	(0.023)	0.030	(0.024)	0.27
Believe proposing laws is a main role (pre)	3,999	0.13	0.33	0.005	(0.020)	0.012	(0.021)	0.009	(0.022)	-0.009	(0.021)	0.028	(0.021)	0.70
Believe passing laws is a main role (pre)	3,999	0.22	0.41	-0.008	(0.027)	-0.012	(0.027)	0.009	(0.029)	-0.015	(0.027)	0.009	(0.028)	0.91
Believe committees are a main role (pre)	3,999	0.05	0.22	-0.004	(0.013)	0.014	(0.015)	0.014	(0.015)	-0.023*	(0.013)	0.019	(0.014)	0.01**
Believe budgeting is a main role (pre)	3,999	0.07	0.26	-0.006	(0.014)	0.002	(0.016)	0.003	(0.014)	-0.024*	(0.014)	0.009	(0.016)	0.17
Believe constituency petitions are a main role (pre)	3,999	0.20	0.40	0.002	(0.019)	-0.019	(0.019)	0.002	(0.019)	-0.021	(0.020)	-0.015	(0.021)	0.59
Believe local transfer lobbying is a main role (pre)	3,999	0.16	0.37	0.016	(0.021)	0.035	(0.022)	-0.004	(0.021)	0.007	(0.023)	0.002	(0.020)	0.47
Believe local project lobbying is a main role (pre)	3,999	0.38	0.49	0.018	(0.029)	0.051	(0.031)	0.041	(0.029)	0.029	(0.031)	0.034	(0.030)	0.63
Believe local project implementation is a main role (pre)	3,999	0.23	0.42	-0.023	(0.021)	0.025	(0.020)	-0.006	(0.023)	-0.014	(0.019)	-0.004	(0.021)	0.27
Passing laws is a main role (pre)	3,999	0.08	0.27	-0.020	(0.016)	0.021	(0.016)	0.003	(0.016)	-0.036**	(0.016)	-0.008	(0.015)	0.03**
Passing laws is a main role (pre)	3,999	0.21	0.40	-0.004	(0.020)	0.029	(0.020)	0.018	(0.021)	-0.014	(0.023)	0.022	(0.022)	0.18
Prefer nationally-oriented deputies (pre)	3,999	0.24	0.42	-0.005	(0.023)	-0.016	(0.024)	0.022	(0.024)	0.025	(0.024)	-0.026	(0.024)	0.17
Prefer locally-oriented deputies (pre)	3,999	0.73	0.45	-0.016	(0.027)	0.022	(0.026)	-0.043*	(0.026)	-0.035	(0.026)	0.010	(0.026)	0.08*
Deputy's village or community is among three most important voting factors	3,999	0.33	0.47	0.017	(0.026)	-0.016	(0.025)	0.047*	(0.026)	0.019	(0.025)	0.022	(0.025)	0.24
Deputy's ethnicity or religion is among three most important voting factors	3,999	0.14	0.35	0.032	(0.020)	0.005	(0.020)	0.023	(0.021)	-0.012	(0.020)	0.016	(0.020)	0.27
Deputy's education or profession is among three most important voting factors	3,999	0.28	0.45	0.001	(0.025)	-0.025	(0.024)	-0.003	(0.025)	-0.023	(0.026)	0.003	(0.024)	0.77
Deputy's party is among three most important voting factors	3,999	0.21	0.41	-0.047**	(0.023)	-0.013	(0.022)	-0.040*	(0.023)	-0.065^{+}	(0.022)	-0.027	(0.024)	0.03**
Deputy's political experience is among three most important voting factors	3,999	0.36	0.48	0.046*	(0.025)	0.030	(0.025)	-0.005	(0.027)	0.046	(0.028)	0.024	(0.026)	0.23
Deputy's amending/approving of laws is among three most important voting factors	3,999	0.33	0.47	-0.025	(0.022)	-0.007	(0.026)	-0.030	(0.025)	0.011	(0.024)	0.005	(0.022)	0.46
Deputy's parliamentary lobbying is among three most important voting factors	3,999	0.74	0.44	0.032	(0.025)	0.036	(0.024)	0.024	(0.024)	0.037	(0.026)	0.009	(0.024)	0.54
Deputy's campaign promises is among three most important voting factors	3,999	0.20	0.40	-0.012	(0.021)	0.022	(0.023)	0.017	(0.021)	-0.007	(0.022)	-0.002	(0.022)	0.51
Deputy's election gifts is among three most important voting factors	3,999	0.08	0.27	-0.008	(0.013)	0.003	(0.014)	0.009	(0.014)	0.009	(0.013)	-0.006	(0.015)	0.68
No listed factor is among most important voting factor	3,999	0.21	0.41	-0.010	(0.022)	-0.020	(0.022)	-0.032	(0.020)	-0.004	(0.024)	-0.014	(0.023)	0.62
Village-level predetermined covariates														
Turnout (2012)	3,999	0.59	0.10	0.008	(0.010)	0.003	(0.010)	0.015	(0.011)	0.010	(0.011)	0.023**	(0.010)	0.19
Incumbent vote share (2012)	3,999	0.68	0.17	0.000	(0.019)	0.035*	(0.018)	-0.026	(0.019)	0.020	(0.017)	0.023	(0.019)	0.01**
Village latitude	3,999	440,370.08	147,848.55	222.972	(2,654.099)	-748.300	(2,709.726)	1,990.090	(2,857.238)	735.476	(2,622.135)	-1,755.535	(2,654.622)	0.76
Village longitude	3,999	1,583,885.99	81,743.16	-1,551.907	(2,091.086)	-4,416.385**	(2,147.208)	-2,476.073	(1,920.76)	-3,647.528*	(1,982.189)	-4,983.795**	(2,083.830)	0.15
Village population	3,999	863.05	686.09	-138.344*	(79.288)	-36.580	(84.031)	-130.982	(80.607)	0.282	(86.294)	-63.088	(83.418)	0.29
Village has a nearby middle school	3,999	0.45	0.50	0.052	(0.061)	0.050	(0.064)	0.052	(0.057)	0.117**	(0.057)	0.054	(0.056)	0.50
Distance to nearest school	2,792	4.52	2.68	0.152	(0.446)	0.072	(0.425)	-0.364	(0.432)	-0.505	(0.442)	0.482	(0.430)	0.17
Share of village completing middle school	3,999	0.04	0.06	0.003	(0.005)	0.005	(0.005)	-0.002	(0.005)	0.002	(0.006)	-0.005	(0.006)	0.29
Share of village households with a good toilet	3,999	0.06	0.11	0.016	(0.014)	-0.010	(0.012)	-0.008	(0.013)	-0.008	(0.011)	-0.003	(0.011)	0.46
Share of village households with piped toilet	3,999	0.08	0.11	-0.008	(0.014)	0.004	(0.015)	0.011	(0.015)	-0.008	(0.014)	0.004	(0.014)	0.75
Share of village households with electricity	3,999	0.01	0.04	-0.006	(0.006)	-0.002	(0.008)	-0.007	(0.006)	0.010	(0.010)	0.010	(0.009)	0.25
Share of village households with good walls	3,999	0.75	0.32	0.008	(0.028)	-0.012	(0.028)	-0.025	(0.029)	0.020	(0.028)	0.017	(0.032)	0.51
Share of village households with a good roof	3,999	0.03	0.07	0.001	(0.008)	0.002	(0.008)	0.015*	(0.009)	-0.001	(0.009)	0.007	(0.008)	0.46
Share of village households with good floors	3,999	0.22	0.20	0.030	(0.025)	0.013	(0.024)	0.027	(0.024)	0.025	(0.025)	0.068^{+}	(0.024)	0.11
Share of village households with a radio	3,999	0.73	0.18	-0.035*	(0.019)	-0.053^{+}	(0.019)	-0.041**	(0.018)	-0.045**	(0.020)	-0.041**	(0.018)	0.05*
Share of village households with a good television	3,999	0.03	0.04	0.001	(0.006)	0.003	(0.006)	0.007	(0.005)	0.007	(0.006)	0.003	(0.006)	0.82
Share of village households with a car	3,999	0.02	0.06	0.001	(0.007)	-0.004	(0.007)	-0.004	(0.007)	-0.005	(0.006)	0.002	(0.007)	0.85
Bambara share of village	3,999	0.03	0.09	-0.005	(0.013)	-0.016	(0.011)	-0.008	(0.012)	-0.007	(0.015)	-0.018	(0.011)	0.51
Diola share of village	3,999	0.07	0.25	0.001	(0.006)	-0.012	(0.008)	-0.006	(0.004)	-0.011	(0.007)	-0.002	(0.005)	0.20
Lebou share of village	3,999	0.00	0.00	0.000	(0.000)	0.000	(0.000)	0.001	(0.000)	0.000	(0.000)	0.001*	(0.000)	0.39
Manding share of village	3,999	0.03	0.11	-0.021*	(0.011)	-0.022**	(0.011)	-0.006	(0.013)	-0.006	(0.010)	-0.016	(0.011)	0.17
Manjag share of village	3,999	0.00	0.00	0.000	(0.000)	0.000	(0.000)	0.001	(0.001)	0.000	(0.000)	0.000	(0.000)	0.89
Maure share of village	3,999	0.00	0.01	-0.001	(0.001)	0.000	(0.002)	-0.001	(0.001)	-0.001	(0.001)	-0.001	(0.001)	0.86
Peul share of village	3,999	0.21	0.38	0.018	(0.018)	0.030*	(0.018)	0.019	(0.018)	0.010	(0.018)	0.019	(0.021)	0.66
Pulaar share of village	3,999	0.06	0.17	-0.018	(0.017)	0.009	(0.018)	0.003	(0.018)	0.014	(0.020)	0.014	(0.018)	0.40
Serer share of village	3,999	0.39	0.43	0.002	(0.020)	0.028	(0.019)	0.003	(0.018)	0.014	(0.024)	-0.010	(0.018)	0.62
Soce share of village	3,999	0.01	0.04	0.001	(0.004)	-0.002	(0.004)	0.003	(0.006)	0.004	(0.004)	0.001	(0.004)	0.83
Soninke share of village	3,999	0.01	0.07	0.001	(0.010)	-0.007	(0.007)	-0.003	(0.008)	0.012	(0.010)	-0.007	(0.007)	0.34
Toucouleur share of village	3,999	0.04	0.13	0.002	(0.017)	-0.010	(0.015)	-0.002	(0.017)	0.006	(0.018)	-0.018	(0.016)	0.67
Wolof share of village	3,999	0.15	0.30	0.021	(0.019)	-0.001	(0.020)	-0.001	(0.018)	-0.038**	(0.019)	0.037**	(0.017)	0.01**

Notes: Each row is a single regression including block and enumerator fixed effects. Observations are inversely weighted by the number of respondents surveyed in each village. Standard errors are clustered by village. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from two-sided tests.

using data from all observations. Unreported results averaging across standardized items produce similar results.

H Deviations from pre-analysis plan

All reported analyses follow our pre-analysis plan, with the following minor exceptions:

1. Although we pre-specified that standard errors would be clustered by randomization block, we instead cluster standard errors by village. This change was implemented to reflect best practice (e.g. Abadie et al. 2017), and—in practice—hardly affects the size of standard errors. The substantive conclusions are not affected by this choice.

2. Our pre-analysis plan proposed both first-differencing and adjusting for a lagged dependent variable. We ultimately chose only the latter due to its greater statistical efficiency (McKenzie 2012). Again, the results are not substantively affected by this choice.

3. Although we pre-specified that requesting an incumbent poster (in the baseline survey) would be included in the ICW index of behavioral outcomes, we ultimately decided to exclude it because of its weak conceptual fit alongside the other behavioral measures of non-electoral accountability. In particular, we in hindsight regard it as a measure of *support* for the incumbent, rather than *seeking to contact* the incumbent. Nevertheless, the results are not substantively altered by including this indicator in either the baseline incumbent evaluation index or the incumbent contact request index.

4. Although we did not pre-specify that we would restrict our polling station-level analysis to polling stations that contain experimental villages that comprise at least 50% of the polling station's registered voters, we believe that this is a natural restriction to minimize the estimation imprecision that arises from including polling stations that contain a small number of villagers that could have received the treatment information through via within-village information diffusion. As a robustness check, Table I11 shows similar results when using all polling stations, but weighting observations by the share of registered voters at a given polling station that are from the village in our experimental sample.

I Additional results

I.1 Effects on the precision of voters beliefs

Table I1 shows that the information treatments increased the precision of respondent beliefs about the incumbent at both baseline and endline.

I.2 Effects on evaluations of challenger parties

Table I2 examines how the information treatments affected voter beliefs about prospective challenger performance in office (if they were to be elected). Since the direction of the effect did not have a clear theoretical expectation, we use two-sided tests. Columns (1) and (2) indicate that voters receiving the benchmark also updated positively about challengers, albeit far less positively than about incumbents. In the context of the model in Appendix section B, this suggests that—to the extent that challengers and previous incumbents are believed to be correlated—previous incumbent performance information exceeded expectations. However, the heterogeneous effects in column (3) cast doubt on this interpretation, given that treatment effects on prospective challenger evaluations are not generally increasing in previous incumbent performance. These results indicate that benchmarked information did not substantially or systematically affect perceptions

Table I1: Effects of information treatments on posterior belief precision

	Baseline survey				Endline survey	
	Incumbent overall performance precision	Relative performance (vs. previous) precision	Prospective incumbent performance precision	Incumbent vote precision	Incumbent overall performance precision	Relative performance (vs. previous) precision
	(1)	(2)	(3)	(4)	(5)	(6)
Incumbent	0.351**	0.291**	0.305**	-0.032	0.302**	0.156*
	(0.102)	(0.095)	(0.096)	(0.048)	(0.077)	(0.092)
Benchmark	0.517**	0.650**	0.486**	0.024	0.448**	0.536**
	(0.104)	(0.099)	(0.099)	(0.050)	(0.076)	(0.090)
One-sided null: Incumbent $\geq$ Benchmark (p value)	0.06	0.00	0.04	0.13	0.02	0.00
Observations	3,963	3,942	3,945	3,615	3,852	3,844
Outcome range	{1,...,10}	{1,...,10}	{1,...,10}	{1,...,10}	{1,...,10}	{1,...,10}
Control outcome mean	6.75	6.74	6.87	8.75	5.86	6.11
Control outcome std. dev.	2.79	2.60	2.63	1.87	2.64	2.83

Notes: Each regression includes randomization block and (baseline and, where relevant, endline) enumerator fixed effects and a lagged dependent variable (baseline) or adjust for the corresponding pre-treatment outcome (endline). Observations are inversely weighted by the number of respondents surveyed in the village (at baseline or endline). Standard errors are clustered by village. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from pre-specified one-sided t tests.

of challengers. This suggests that the differential effects of benchmarked information on relative evaluations, including vote choice, likely reflect the increased weight attached to incumbent performance signals when an accompanying benchmark increases the precision of the signal, rather than the difference between posterior and prior expectations of challenger performance.

I.3 Additional heterogeneous treatment effects

Tables I3 and I4 show how treatment effects vary with four additional variables relating to the process of belief updating. Panel A reports heterogeneous effects by an incumbent performance index weighted by whether voters stated that national or local factors were most important to them before receiving treatment.[23] Panel B builds on our national and local performance comparison by further interacting the benchmark treatment with the previous incumbent's level of performance. Panels C and D respectively interact treatments with the position and precision of voter prior beliefs. The results support the Bayesian updating interpretation: voters are somewhat more sensitive to more relevant performance indicators, and—especially at baseline—update their beliefs more favorably when their had low or imprecise prior expectations. Moreover, the results—especially for the outcomes at baseline—show that voters that received the benchmark treatment also updated more favorably when the current incumbent outperformed the previous incumbent.

[23]The relevance-weighted performance index assigns a respondent the national, local, or both performance indicators corresponding to which they listed among the three most important factors determining their vote choice. An indicator for respondents not listing national or local performance as important is also interacted with treatments.

Table I2: Effects of information treatments on prospective challenger performance evaluations (baseline survey)

	Prospective challenger performance		
	(1)	(2)	(3)
Incumbent	0.044	0.051	0.049
	(0.046)	(0.032)	(0.032)
Duties	-0.006		
	(0.044)		
Incumbent × Duties	0.015		
	(0.064)		
Benchmark	0.074‡	0.112§§	0.114§§
	(0.039)	(0.030)	(0.029)
Benchmark × Duties	0.073		
	(0.060)		
Incumbent × National previous performance (ICW)			0.099**
			(0.045)
Benchmark × National previous performance (ICW)			0.026
			(0.043)
Incumbent × Local previous performance (ICW)			-0.011
			(0.032)
Benchmark × Local previous performance (ICW)			-0.038
			(0.031)
Observations	3,888	3,888	3,888
Outcome range	{1,...,5}	{1,...,5}	{1,...,5}
Control outcome mean	3.42	3.42	3.42
Control outcome std. dev.	0.88	0.88	0.88
National previous performance (ICW) mean			-0.00
National previous performance (ICW) std. dev.			1.00
Local previous performance (ICW) mean			0.00
Local previous performance (ICW) std. dev.			1.00

Notes: Each regression includes randomization block and enumerator fixed effects. Lower-order interaction terms are included but not shown. Observations are inversely weighted by the number of respondents surveyed in the village. Standard errors are clustered by village. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from pre-specified one-sided t tests; ‡ $p < 0.1$, § $p < 0.05$, §§ $p < 0.01$ from two-sided tests when coefficients point in the opposite direction to the pre-specified hypothesis.

I.4 Negative updating in Oussouye

While our results for the full sample imply that the distribution of prior beliefs induced voters to generally updated favorably about their current incumbents, Figure 3 also demonstrates that there is variation within and across departments in performance along reported dimensions. Most notably, the previous incumbent in Oussouye outperformed the current incumbent in terms of the projects and transfers about which many voters care most. This is thus the department where unfavorable updating is most likely to be observed. We examine this more formally by restricting our sample to respondents from Oussouye. The average treatment effects reported in Tables I5 and I6 show that—particularly at baseline—voters updated unfavorably about the current incumbent in Oussouye. This is most pronounced among the voters that received the benchmarked information showing that

Table I3: Heterogeneous effects of information treatments by relevance-weight leaflet content and priors beliefs (baseline survey)

	Incumbent evaluation outcomes					Incumbent contact request outcomes		
	Incumbent overall performance (1)	Relative performance (vs. previous) (2)	Prospective incumbent performance (3)	Incumbent vote intention (4)	Incumbent evaluation index (ICW) (5)	Request incumbent visit (6)	Request incumbent conversation (7)	Incumbent contact request index (ICW) (8)
Panel A: Heterogeneity by (standardized) relevance-weighted reported performance level								
Incumbent	0.361**	0.289**	0.265**	0.030**	0.256**	0.039**	0.028**	0.075**
	(0.041)	(0.035)	(0.035)	(0.010)	(0.027)	(0.014)	(0.013)	(0.029)
Incumbent × Relevant performance (ICW)	0.240**	0.205**	0.188**	0.030**	0.152**	-0.002	-0.004	-0.007
	(0.039)	(0.036)	(0.035)	(0.011)	(0.027)	(0.015)	(0.015)	(0.032)
Benchmark	0.438**	0.379**	0.332**	0.035**	0.317**	0.017	0.021†	0.042†
	(0.043)	(0.040)	(0.038)	(0.010)	(0.031)	(0.014)	(0.013)	(0.030)
Benchmark × Relevant performance (ICW)	0.200**	0.165**	0.136**	0.027**	0.138**	-0.007	0.009	0.001
	(0.038)	(0.037)	(0.037)	(0.011)	(0.029)	(0.017)	(0.016)	(0.036)
Observations	3,942	3,932	3,928	3,999	3,891	3,999	3,998	3,998
Relevant performance (ICW) range	[-2.94,2.73]	[-2.94,2.73]	[-2.94,2.73]	[-2.94,2.73]	[-2.94,2.73]	[-2.94,2.73]	[-2.94,2.73]	[-2.94,2.73]
Panel B: Heterogeneity by (standardized) local and national reported performance level, including difference performance relative to benchmark								
Incumbent	0.359**	0.287**	0.264**	0.030**	0.255**	0.039**	0.028**	0.076**
	(0.039)	(0.034)	(0.034)	(0.010)	(0.026)	(0.014)	(0.013)	(0.029)
Incumbent × National performance (ICW)	0.026	0.061*	0.007	0.000	0.007	0.000	0.017	0.018
	(0.043)	(0.032)	(0.033)	(0.009)	(0.025)	(0.015)	(0.014)	(0.032)
Incumbent × Local performance (ICW)	0.318**	0.242**	0.251**	0.025**	0.203**	0.009	-0.004	0.006
	(0.045)	(0.037)	(0.038)	(0.011)	(0.030)	(0.016)	(0.015)	(0.033)
Benchmark	0.428**	0.367**	0.326**	0.034**	0.310**	0.016	0.021†	0.041†
	(0.042)	(0.037)	(0.037)	(0.010)	(0.029)	(0.014)	(0.014)	(0.030)
Benchmark × National performance (ICW)	-0.134	-0.365§§	-0.173‡	-0.005	-0.207§§	-0.006	0.027	0.022
	(0.115)	(0.103)	(0.100)	(0.023)	(0.071)	(0.038)	(0.035)	(0.078)
Benchmark × Previous national performance (ICW)	0.172	0.383§	0.170	0.020	0.236§	0.003	-0.040	-0.039
	(0.172)	(0.158)	(0.160)	(0.036)	(0.105)	(0.061)	(0.057)	(0.127)
Benchmark × Local performance (ICW)	0.310**	0.430**	0.295**	0.043**	0.319**	0.007	-0.004	0.004
	(0.086)	(0.076)	(0.080)	(0.017)	(0.053)	(0.032)	(0.030)	(0.068)
Benchmark × Previous local performance (ICW)	0.014	-0.094**	-0.012	0.008	-0.029	0.012	-0.001	0.013
	(0.048)	(0.047)	(0.041)	(0.010)	(0.030)	(0.013)	(0.012)	(0.026)
Observations	3,942	3,932	3,928	3,999	3,891	3,999	3,998	3,998
National performance (ICW) range	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]
Previous national performance (ICW) range	[-2.44,2.62]	[-2.44,2.62]	[-2.44,2.62]	[-2.44,2.62]	[-2.44,2.62]	[-2.44,2.62]	[-2.44,2.62]	[-2.44,2.62]
Local performance (ICW) range	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]
Previous local performance (ICW) range	[-1.39,1.29]	[-1.39,1.29]	[-1.39,1.29]	[-1.39,1.29]	[-1.39,1.29]	[-1.39,1.29]	[-1.39,1.29]	[-1.39,1.29]
Panel C: Heterogeneity by (standardized) prior belief level								
Incumbent	0.361**	0.295**	0.268**	0.034**	0.259**	0.043**	0.032**	0.084**
	(0.043)	(0.037)	(0.037)	(0.010)	(0.029)	(0.014)	(0.014)	(0.030)
Incumbent × Prior index (ICW)	-0.132**	-0.085**	-0.096**	-0.043**	-0.116**	-0.031**	-0.023*	-0.061**
	(0.037)	(0.033)	(0.036)	(0.011)	(0.029)	(0.014)	(0.014)	(0.030)
Benchmark	0.443**	0.382**	0.334**	0.039**	0.319**	0.019†	0.024*	0.048†
	(0.044)	(0.042)	(0.039)	(0.010)	(0.033)	(0.014)	(0.014)	(0.031)
Benchmark × Prior index (ICW)	-0.163**	-0.081**	-0.054†	-0.030**	-0.089**	-0.007	-0.008	-0.017
	(0.040)	(0.037)	(0.038)	(0.011)	(0.029)	(0.015)	(0.015)	(0.033)
Observations	3,908	3,906	3,905	3,922	3,891	3,922	3,921	3,921
Prior index (ICW) range	[-2.23,2.10]	[-2.23,2.10]	[-2.23,2.10]	[-2.23,2.10]	[-2.23,2.10]	[-2.23,2.10]	[-2.23,2.10]	[-2.23,2.10]
Panel D: Heterogeneity by (standardized) prior belief precision								
Incumbent	0.363**	0.296**	0.287**	0.045**	0.278**	0.036**	0.035**	0.078**
	(0.044)	(0.038)	(0.039)	(0.011)	(0.030)	(0.014)	(0.014)	(0.031)
Incumbent × Prior precision index (ICW)	-0.005	0.002	-0.042	-0.031**	-0.051*	-0.011	-0.009	-0.022
	(0.038)	(0.031)	(0.037)	(0.012)	(0.027)	(0.015)	(0.016)	(0.033)
Benchmark	0.445**	0.397**	0.360**	0.048**	0.343**	0.019†	0.031**	0.056*
	(0.045)	(0.042)	(0.040)	(0.011)	(0.034)	(0.015)	(0.014)	(0.031)
Benchmark × Prior precision index (ICW)	-0.029	-0.056*	-0.044	-0.025**	-0.069**	-0.008	-0.017	-0.027
	(0.042)	(0.033)	(0.040)	(0.013)	(0.028)	(0.016)	(0.015)	(0.032)
Observations	3,636	3,636	3,626	3,667	3,609	3,667	3,666	3,666
Prior precision (ICW) range	[-3.19,1.38]	[-3.19,1.38]	[-3.19,1.38]	[-3.19,1.38]	[-3.19,1.38]	[-3.19,1.38]	[-3.19,1.38]	[-3.19,1.38]
Outcome range	{1,...,5}	{1,...,5}	{1,...,5}	{0,1}	[-2.3,1.9]	{0,1}	{0,1}	[-1.6,0.7]
Control outcome mean	2.83	3.20	3.15	0.59	0.00	0.70	0.70	-0.00
Control outcome std. dev.	1.07	0.90	1.09	0.49	1.00	0.46	0.46	1.00

Notes: Each regression includes randomization block and enumerator fixed effects and a lagged dependent variable (in columns (5)-(8), pre-treatment incumbent vote is used as a proxy). Lower-order (standardized) interaction terms are included but not shown. Observations are inversely weighted by the number of respondents surveyed in the village. Control outcome means and standard deviations are for the sample in panels A and B. Standard errors are clustered by village. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from pre-specified one-sided t tests; ‡ $p < 0.1$, § $p < 0.05$, §§ $p < 0.01$ from two-sided tests when coefficients point in the opposite direction to the pre-specified hypothesis.

Table I4: Heterogeneous effects of information treatments by relevance-weighted leaflet content and importance of performance information for vote choice (endline survey)

	Incumbent evaluation outcomes					Incumbent contact request outcomes				
	Incumbent overall performance	Relative performance (vs. previous)	Incumbent vote	Incumbent vote (validated)	Incumbent evaluation index (ICW)	Request incumbent visit	Request incumbent conversation	Request hotline number	Called hotline	Incumbent contact request index (ICW)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Heterogeneity by (standardized) relevance-weighted reported performance level										
Incumbent	0.162**	0.145**	-0.024	-0.018	0.111**	-0.000	-0.017	0.006	0.011	0.029
	(0.035)	(0.034)	(0.021)	(0.021)	(0.039)	(0.005)	(0.029)	(0.007)	(0.014)	(0.036)
Incumbent × Relevant performance (ICW)	0.038	0.027	-0.003	-0.003	0.001	-0.001	0.010	0.019**	0.024*	0.076**
	(0.037)	(0.037)	(0.022)	(0.022)	(0.045)	(0.004)	(0.016)	(0.008)	(0.014)	(0.039)
Benchmark	0.248**	0.248**	0.003	0.007	0.219**	0.009**	0.063	0.006	0.021†	0.089**
	(0.036)	(0.034)	(0.022)	(0.021)	(0.041)	(0.004)	(0.053)	(0.007)	(0.015)	(0.036)
Benchmark × Relevant performance (ICW)	0.007	0.020	0.001	-0.006	-0.015	-0.004	-0.000	-0.001	0.027*	0.018
	(0.037)	(0.038)	(0.022)	(0.022)	(0.045)	(0.005)	(0.016)	(0.007)	(0.016)	(0.042)
Observations	3,834	3,825	3,781	3,781	3,708	3,876	3,876	3,876	3,876	3,876
Overall performance (ICW) range	[-2.94,2.73]	[-2.94,2.73]	[-2.94,2.73]	[-2.94,2.73]	[-2.94,2.73]	[-2.94,2.73]	[-2.94,2.73]	[-2.94,2.73]	[-2.94,2.73]	[-2.94,2.73]
Panel B: Heterogeneity by (standardized) local and national reported performance level, including difference in performance relative to benchmark										
Incumbent	0.162**	0.145**	-0.025	-0.018	0.110**	0.000	-0.016	0.006	0.010	0.028
	(0.034)	(0.034)	(0.021)	(0.021)	(0.039)	(0.004)	(0.028)	(0.007)	(0.015)	(0.036)
Incumbent × National performance (ICW)	-0.027	-0.020	-0.001	-0.013	-0.021	-0.008	-0.011	0.003	0.024*	0.001
	(0.034)	(0.036)	(0.024)	(0.025)	(0.046)	(0.007)	(0.029)	(0.007)	(0.013)	(0.041)
Incumbent × Local performance (ICW)	0.067**	0.096**	0.004	0.015	0.064†	0.010*	0.017	0.014*	0.012	0.097**
	(0.034)	(0.036)	(0.023)	(0.024)	(0.045)	(0.005)	(0.019)	(0.007)	(0.014)	(0.037)
Benchmark	0.247**	0.248**	0.002	0.007	0.218**	0.008**	0.059	0.006	0.020†	0.085**
	(0.035)	(0.034)	(0.022)	(0.022)	(0.040)	(0.004)	(0.050)	(0.007)	(0.015)	(0.036)
Benchmark × National performance (ICW)	-0.036	-0.025	0.005	0.023	-0.033	-0.026§	-0.194	-0.018	0.015	-0.155
	(0.084)	(0.086)	(0.056)	(0.058)	(0.093)	(0.012)	(0.157)	(0.020)	(0.040)	(0.106)
Benchmark × Previous national performance (ICW)	-0.009	-0.089	0.015	-0.021	-0.043	0.016	0.166	-0.002	-0.005	0.068
	(0.137)	(0.135)	(0.080)	(0.087)	(0.143)	(0.013)	(0.143)	(0.033)	(0.070)	(0.147)
Benchmark × Local performance (ICW)	0.018	0.046	0.019	0.009	0.042	0.017*	0.075	0.021	0.007	0.145*
	(0.067)	(0.070)	(0.043)	(0.046)	(0.078)	(0.009)	(0.064)	(0.018)	(0.032)	(0.084)
Benchmark × Previous local performance (ICW)	-0.036	0.002	0.020	0.010	0.005	-0.002	0.039	0.001	0.002	-0.006
	(0.035)	(0.034)	(0.021)	(0.022)	(0.035)	(0.004)	(0.036)	(0.007)	(0.018)	(0.041)
Observations	3,834	3,825	3,781	3,781	3,708	3,876	3,876	3,876	3,876	3,876
National performance (ICW) range	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]	[-1.42,2.21]
Previous national performance (ICW) range	[-2.44,2.62]	[-2.44,2.62]	[-2.44,2.62]	[-2.44,2.62]	[-2.44,2.62]	[-2.44,2.62]	[-2.44,2.62]	[-2.44,2.62]	[-2.44,2.62]	[-2.44,2.62]
Local performance (ICW) range	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]	[-1.39,1.35]
Previous local performance (ICW) range	[-1.39,1.29]	[-1.39,1.29]	[-1.39,1.29]	[-1.39,1.29]	[-1.39,1.29]	[-1.39,1.29]	[-1.39,1.29]	[-1.39,1.29]	[-1.39,1.29]	[-1.39,1.29]
Panel C: Heterogeneity by (standardized) prior belief level										
Incumbent	0.166**	0.152**	-0.028†	-0.021	0.110**	0.001	-0.018	0.006	0.014	0.038
	(0.035)	(0.034)	(0.021)	(0.021)	(0.040)	(0.005)	(0.029)	(0.007)	(0.015)	(0.037)
Incumbent × Prior index (ICW)	-0.023	-0.011	0.011	0.027	-0.004	0.009	-0.022	0.009	-0.007	0.056
	(0.039)	(0.037)	(0.021)	(0.020)	(0.042)	(0.006)	(0.026)	(0.008)	(0.014)	(0.043)
Benchmark	0.243**	0.251**	-0.003	-0.000	0.219**	0.009**	0.060	0.005	0.022†	0.088**
	(0.035)	(0.034)	(0.022)	(0.022)	(0.041)	(0.004)	(0.050)	(0.007)	(0.015)	(0.038)
Benchmark × Prior index (ICW)	-0.041	-0.025	-0.005	0.023	-0.032	0.004	-0.007	0.010	-0.000	0.044
	(0.039)	(0.037)	(0.020)	(0.022)	(0.042)	(0.005)	(0.018)	(0.007)	(0.015)	(0.040)
Observations	3,834	3,825	3,781	3,781	3,708	3,876	3,876	3,876	3,876	3,876
Prior index (ICW) range	[-2.29,2.10]	[-2.29,2.10]	[-2.29,2.10]	[-2.29,2.10]	[-2.29,2.10]	[-2.29,2.10]	[-2.29,2.10]	[-2.29,2.10]	[-2.29,2.10]	[-2.29,2.10]
Panel D: Heterogeneity by (standardized) prior belief precision										
Incumbent	0.153**	0.152**	-0.027	-0.018	0.109**	0.002	-0.002	0.008	0.009	0.043
	(0.036)	(0.035)	(0.022)	(0.022)	(0.041)	(0.005)	(0.005)	(0.007)	(0.015)	(0.038)
Incumbent × Prior precision index (ICW)	0.059	0.030	-0.002	-0.001	0.031	-0.002	-0.006	-0.002	-0.002	-0.017
	(0.039)	(0.039)	(0.022)	(0.022)	(0.045)	(0.006)	(0.006)	(0.008)	(0.014)	(0.043)
Benchmark	0.245**	0.264**	0.001	0.006	0.226**	0.008*	0.004	0.006	0.024†	0.089**
	(0.038)	(0.035)	(0.023)	(0.023)	(0.043)	(0.004)	(0.005)	(0.008)	(0.016)	(0.039)
Benchmark × Prior precision index (ICW)	-0.005	0.035	0.005	0.011	0.017	0.005	-0.002	0.011	-0.028*	0.011
	(0.040)	(0.038)	(0.022)	(0.022)	(0.045)	(0.005)	(0.005)	(0.008)	(0.016)	(0.044)
Observations	3,834	3,825	3,781	3,781	3,708	3,876	3,876	3,876	3,876	3,876
Prior precision (ICW) range	[-3.19,1.38]	[-3.19,1.38]	[-3.19,1.38]	[-3.19,1.38]	[-3.19,1.38]	[-3.19,1.38]	[-3.19,1.38]	[-3.19,1.38]	[-3.19,1.38]	[-3.19,1.38]
Outcome range	{1,...,5}	{1,...,5}	{0,1}	{0,1}	[-2.8,1.9]	{0,1}	{0,1}	{0,1}	{0,1}	[-7.5,1.3]
Control outcome mean	3.08	3.46	0.64	0.53	0.00	0.98	0.98	0.95	0.11	-0.00
Control outcome std. dev.	0.93	0.95	0.48	0.50	1.00	0.14	0.14	0.21	0.32	1.00

Notes: Each regression includes randomization block and (baseline and endline) enumerator fixed effects and a lagged dependent variable (in columns (5)-(10), pre-treatment incumbent vote is used as a proxy). Lower-order interaction terms are included but not shown. Observations are inversely weighted by the number of respondents surveyed in the village. Control outcome means and standard deviations are for the sample in panels A and B. Standard errors are clustered by village. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from pre-specified one-sided t tests; ‡ $p < 0.1$, § $p < 0.05$, §§ $p < 0.01$ from two-sided tests when coefficients point in the opposite direction to the pre-specified hypothesis.

Table I5: Average effects of information treatments on beliefs about incumbent performance, intention to vote for the incumbent, and requests from the incumbent in Oussouye (baseline survey)

	Incumbent evaluation outcomes					Incumbent contact request outcomes		
	Incumbent overall performance (1)	Relative performance (vs. previous) (2)	Prospective incumbent performance (3)	Incumbent vote intention (4)	Incumbent evaluation index (ICW) (5)	Request incumbent visit (6)	Request incumbent conversation (7)	Incumbent contact request index (ICW) (8)
Incumbent	-0.142	-0.038	-0.113	0.001	-0.065	0.067	0.090*	0.173†
	(0.124)	(0.083)	(0.086)	(0.027)	(0.066)	(0.045)	(0.041)	(0.092)
Benchmark	-0.124	-0.579**	-0.262*	0.014	-0.269**	0.021	0.020	0.046
	(0.154)	(0.101)	(0.103)	(0.027)	(0.074)	(0.051)	(0.046)	(0.106)
Observations	267	269	269	270	267	270	270	270
Outcome range	{1,...,5}	{1,...,5}	{1,...,5}	{0,1}	[-2.3,1.9]	{0,1}	{0,1}	[-1.6,0.7]
Control outcome mean	2.58	2.89	2.64	0.16	-0.82	0.27	0.24	-0.99
Control outcome std. dev.	1.23	0.98	1.13	0.37	0.78	0.45	0.43	0.97

Notes: Each regression includes randomization block and enumerator fixed effects and a lagged dependent variable (in columns (5)-(8), pre-treatment incumbent vote is used as a proxy). Observations are inversely weighted by the number of respondents surveyed in the village. Standard errors are clustered by village. Given that this subsample analysis was not pre-specified, † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from two-sided t tests.

the current incumbent performed worse than the previous incumbent on salient dimensions. These findings suggest that voters react similarly to information that positively and negatively updates their beliefs, although our estimates are tentative due to the limited statistical power associated with our small sample of respondents in this single department.

I.5 The importance voters attach to incumbent legislative performance does not change

While voters' evaluations of incumbents were persistently affected, the provision of incumbent performance information could also influence the relative weight attached to incumbent legislative performance in making voting decisions. Any changes in voting behavior might then reflect changes in salience, rather than changes in beliefs. To examine such salience effects, we asked voters what the three most important factors in determining their vote choice in the 2017 election were. Table I7 shows that the treatments did not affect the likelihood of reporting that national or local legislative performance is one of the three most important, or the most important, factor in determining vote choice. This suggests that the effects on vote choice are unlikely to reflect voters placing greater weight on the considerations that the treatment information related to.

I.6 Within-village information diffusion

The endline survey shows substantial voter engagement with the leaflets within their village. While almost exactly 0% of control group respondents reported discussing the leaflet with others, this share rises to 37% and 39% in the incumbent and benchmark treatment groups. Unreported regression estimates indicate that the difference from the control group is statistically significant, and suggests that substantial information diffusion occurred. This may account for the fact that

Table I6: Average effects of information treatments on beliefs about incumbent performance, reported vote for the incumbent, and requests from the incumbent in Oussouye (endline survey)

	Incumbent evaluation outcomes					**Incumbent contact request outcomes**				
	Incumbent overall performance	Relative performance (vs. previous)	Incumbent vote	Incumbent vote (validated)	Incumbent evaluation index (ICW)	Request incumbent visit	Request incumbent conversation	Request hotline number	Called hotline	Incumbent contact request index (ICW)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Incumbent	-0.074	-0.076	-0.016	-0.037	-0.048	-0.035	-0.035	-0.026	0.024	-0.205
	(0.097)	(0.100)	(0.079)	(0.082)	(0.153)	(0.027)	(0.027)	(0.022)	(0.029)	(0.138)
Benchmark	-0.045	-0.112	0.047	0.029	-0.076	-0.042	-0.042	-0.064*	0.029	-0.323†
	(0.092)	(0.101)	(0.091)	(0.088)	(0.145)	(0.029)	(0.029)	(0.028)	(0.029)	(0.189)
Observations	261	261	233	233	232	262	262	262	262	262
Outcome range	{1,...,5}	{1,...,5}	{0,1}	{0,1}	[-2.8,1.9]	{0,1}	{0,1}	{0,1}	{0,1}	[-7.3,1.5]
Control outcome mean	2.77	3.23	0.55	0.55	-0.36	1.00	1.00	0.93	0.02	-0.09
Control outcome std. dev.	0.68	0.77	0.50	0.50	0.93	0.00	0.00	0.26	0.15	0.66

Notes: Each regression includes randomization block and (baseline and endline) enumerator fixed effects and a lagged dependent variable (in columns (5)-(10), pre-treatment incumbent vote is used as a proxy). Observations are inversely weighted by the number of respondents surveyed in the village. Standard errors are clustered by village. Given that this subsample analysis was not pre-specified, † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from two-sided t tests.

Table I7: Effects of information treatments on self-reported importance of performance in making vote choice (endline survey)

	Performance is one of the three most important factors in vote choice			Performance is the most important factor in vote choice		
	(1)	(2)	(3)	(4)	(5)	(6)
Duties	0.005			0.001		
	(0.014)			(0.021)		
Incumbent	-0.018	-0.017		-0.009	-0.024	
	(0.015)	(0.011)		(0.022)	(0.016)	
Incumbent × Duties	0.001			-0.029		
	(0.022)			(0.031)		
Benchmark	0.010	0.002		0.001	-0.006	
	(0.014)	(0.010)		(0.022)	(0.016)	
Benchmark × Duties	-0.015			-0.013		
	(0.020)			(0.032)		
Performance			-0.008			-0.015
			(0.009)			(0.014)
One-sided null: Incumbent≥Benchmark (p value)		0.04			0.15	
Observations	3,876	3,876	3,876	3,876	3,876	3,876
Outcome range	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}
Control outcome mean	0.89	0.89	0.89	0.50	0.50	0.50
Control outcome std. dev.	0.31	0.31	0.31	0.50	0.50	0.50

Notes: Each regression includes randomization block and (baseline and endline) enumerator fixed effects and a lagged dependent variable. Observations are inversely weighted by the endline number of respondents surveyed in the village. Standard errors are clustered by village. Given that these hypotheses were not pre-specified, † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from two-sided t tests.

directly providing leaflets to less than 2% of registered voters still resulted in some discernible polling station-level effects.

I.7 Cross-village informational spillovers

Another possibility is that information spilled from treated to control villages. This would underestimate the effects of the information treatments if control villages similarly became more positive about the incumbent. We estimate spillovers among the 75 pure control villages by defining spillover potential as the number of villages within *x*km of a treated village receiving performance (incumbent or benchmark) information. Panels A, B, and C of Table I8 indicate that for treated villages respectively within 1km, 2.5km, and 5km of a control village, there is no systematic evidence suggesting that proximity to treatment significantly affected endline voter beliefs or voting behavior, conditional on the number of villages within our sample within the same distance. This applies both on average, as well as by the level of reported incumbent performance. Unreported results show that leaflet recall is unaffected by treatment assignment.

I.8 Party responses to information dissemination

Politicians rarely stand by when potentially influential information is released (e.g. Arias et al. 2018*b*; Cruz, Keefer and Labonne 2019). Consequently, a possible explanation for the lack of a persistent average treatment effect on incumbent electoral support, but positive effects when interacted with the information content, is that challenger parties were particularly effective at counteracting information that generally increased favorability toward the incumbent. Incumbents may also respond by highlighting positive information, although—to the extent that it is effective—this should reinforce the favorable immediate updating of voters. Another channel through which strategic responses could explain our findings is if incumbents (challengers) reallocate resources from treatment (control) to control (treatment) villages upon learning that favorable information had already been disseminated.

We investigated such equilibrium campaign responses to information dissemination by using our endline survey to gauge two types of party or candidate action. First, we asked respondents if, and how, the incumbent or challenger parties (or their agents) responded specifically to the leaflet's provision. Second, we used a list experiment to measure the extent of vote buying, in order to assess whether party electoral strategies change, even without explicitly mentioning the leaflets.[24]

As shown in columns (1) and (5) of Table I9, challengers and especially incumbents responded directly to the intervention. As the almost-zero control group mean indicates, responses were concentrated in treated villages. Decomposing candidate responses by type, the vast majority of incumbent responses involved a community meeting or talking with the village chief, while challenger parties held community meetings or had party operatives visit voters. If incumbent responses are at least as effective as challenger responses, it is hard to account for the zero average effects observed at the individual and polling station levels. To better understand what parties did, we followed up with respondents in December 2017 to ask about what actions parties took

[24]Half the sample was subject to a list experiment including incumbent vote buying as the omitted option from the list; vote buying by a challenger party was omitted for the other half of the sample.

Table I8: Effects of information spillovers on voter beliefs, self-reported vote choices, and precinct-level vote choices (endline survey and polling station data)

	Incumbent overall performance		Relative performance (vs. previous)		Incumbent vote		Incumbent vote (validated)		Incumbent vote share (turnout)		Incumbent vote share (registered)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Spillovers within 1km												
Performance spillover	0.017	-0.026	-0.065	-0.070	-0.169†	-0.137†	-0.070	-0.039	-0.152	-0.029	-0.099	0.026
	(0.176)	(0.152)	(0.136)	(0.141)	(0.096)	(0.070)	(0.076)	(0.062)	(0.097)	(0.106)	(0.121)	(0.079)
Performance spillover × Performance index (ICW)		0.288		-0.092		-0.343*		-0.268		0.248		0.240
		(0.365)		(0.294)		(0.133)		(0.179)		(0.274)		(0.226)
Spillover mean	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.14	0.14	0.14	0.14
Panel B: Spillovers within 2.5km												
Performance spillover	-0.006	0.075	-0.099	-0.150	-0.103	-0.079	-0.091	-0.081	-0.070	-0.060	-0.018	0.002
	(0.118)	(0.168)	(0.124)	(0.172)	(0.067)	(0.073)	(0.066)	(0.067)	(0.069)	(0.102)	(0.044)	(0.057)
Performance spillover × Performance index (ICW)		-0.229		0.167		-0.163		-0.116		-0.069		-0.108
		(0.372)		(0.363)		(0.162)		(0.147)		(0.270)		(0.147)
Spillover mean	0.54	0.54	0.54	0.54	0.54	0.54	0.54	0.54	0.56	0.56	0.56	0.56
Panel C: Spillovers within 5km												
Performance spillover	-0.148	-0.134	-0.163	-0.194	0.133	0.048	0.157†	0.086	0.198†	0.220**	0.038	0.019
	(0.146)	(0.226)	(0.127)	(0.195)	(0.099)	(0.106)	(0.082)	(0.087)	(0.101)	(0.068)	(0.086)	(0.077)
Performance spillover × Performance index (ICW)		-0.031		-0.083		-0.100		-0.072		0.117*		0.029
		(0.122)		(0.127)		(0.070)		(0.062)		(0.058)		(0.052)
Spillover mean	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
Observations	635	635	635	635	632	632	632	632	72	72	72	72
Control outcome mean	3.08	3.08	3.46	3.46	0.64	0.64	0.53	0.53	0.70	0.70	0.41	0.41
Control outcome std. dev.	0.93	0.93	0.95	0.95	0.48	0.48	0.50	0.50	0.18	0.18	0.13	0.13

Notes: Each regression includes as covariates the number of experimental villages with a given distance (and its interaction with performance, where relevant) and department and (baseline and endline) enumerator fixed effects. Lower-order interaction terms are included but not shown. Observations in columns (1)-(8) are inversely weighted by the baseline number of respondents surveyed in the village. Standard errors are clustered by village. Given that these hypotheses were not pre-specified, † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from two-sided t tests.

Table I9: Effects of information treatments on incumbent and challenger responses (endline survey)

	Incumbent response			Challenger response		
	(1)	(2)	(3)	(4)	(5)	(6)
Incumbent	0.066**	0.066**	0.067**	0.042**	0.042**	0.042**
	(0.011)	(0.011)	(0.011)	(0.008)	(0.008)	(0.008)
Benchmark	0.078**	0.078**	0.078**	0.047**	0.047**	0.047**
	(0.011)	(0.011)	(0.011)	(0.009)	(0.009)	(0.009)
Incumbent × Overall performance (ICW)		-0.010			-0.000	
		(0.015)			(0.008)	
Benchmark × Overall performance (ICW)		-0.029†			-0.020†	
		(0.016)			(0.011)	
Incumbent × National performance (ICW)			-0.029**			-0.023**
			(0.008)			(0.007)
Benchmark × National performance (ICW)			-0.031**			-0.031**
			(0.009)			(0.007)
Incumbent × Local performance (ICW)			0.027**			0.018*
			(0.010)			(0.007)
Benchmark × Local performance (ICW)			0.018†			0.010
			(0.010)			(0.008)
Observations	3,875	3,875	3,875	3,875	3,875	3,875
Outcome range	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}
Control outcome mean	0.01	0.01	0.01	0.01	0.01	0.01
Control outcome std. dev.	0.08	0.08	0.08	0.08	0.08	0.08

Notes: Each regression includes randomization block and (baseline and endline) enumerator fixed effects. Lower-order interaction terms are included but not shown. Observations are inversely weighted by the baseline number of respondents surveyed in the village. Standard errors are clustered by village. Given that these hypotheses were not pre-specified, † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from two-sided t tests.

and whether they were effective. Voters that reported incumbent-held community meetings or discussions with the chief were convinced to vote for the incumbent 70-80% of the time, while the less-frequent challenger community meetings and party visits rarely convinced or even encouraged voters to support them.

The interactions with national and local incumbent performance, in column (4), suggest that incumbents capitalized on positive local performance information. Given such responses were compelling to voters, and likely reached a broader electorate that was more likely to turn out and which Table 8 found to be more receptive to local performance than our survey respondents, this could explain the positive effects of treatment on incumbent vote share at the polling station level where local performance was strongest. The lack of an effect on average at the polling station could then reflect effective but relatively sparse incumbent responses. Column (8) indicates that challengers sought to counteract such efforts, but—as noted above—these were rarely seen as effective. In contrast, both incumbents and challengers respond more to national performance information when they performed poorly, although this is not a major factor determining vote choices.

Although vote buying is prevalent, we were not able to detect a systematic indirect response to information dissemination through vote buying. Table I10 uses a list experiment to examine the effects of the information treatments on vote buying: half the sample received the control list containing three items, while one quarter of the sample received a 4-item list either containing

incumbent vote buying or challenger vote buying. The results of the list experiment in columns (1) and (5) indicate that 22% of voters reported receiving a gift from the incumbent, while another 22% reported receiving a gift from a challenger. Although such vote buying was a little lower in treated villages, especially among challengers where local incumbent performance was strong, the estimates are too small and imprecise to conclude that the substitution of vote buying across villages can account for our findings.

I.9 Weighted polling station level estimates

Table I11 reports the polling station level results for the sample of experimental villages, weighting observations by the fraction of the registered voter pool residing in an experimental village.[25] The results are qualitatively similar to those in Table 9, but are smaller in magnitude. The smaller coefficient values likely reflect assigning greater than zero weight to observations from polling stations where very few voters could have been exposed to treatment within their village (and thus treatment effects would unsurprisingly be small).

I.10 Effects on electoral turnout

Table I12 shows that polling station level turnout was not significantly affected by disseminating incumbent performance information. This is consistent with the effects of the information treatments on self-reported turnout in Table I13, which reports little evidence to suggest that turnout decisions were influenced by the information.

[25] We were unable to obtain complete electoral returns in four villages.

Table I10: Effects of information treatments on vote buying (endline survey)

	Incumbent list experiment				Challengers list experiment			
	Items listed (1)	Items listed (2)	Items listed (3)	Items listed (4)	Items listed (5)	Items listed (6)	Items listed (7)	Items listed (8)
Treatment list	0.215**	0.223**	0.225**	0.228**	0.219**	0.247**	0.243**	0.243**
	(0.029)	(0.049)	(0.049)	(0.049)	(0.031)	(0.052)	(0.052)	(0.051)
Treatment list × Incumbent		-0.040	-0.041	-0.044		-0.079	-0.076	-0.076
		(0.070)	(0.070)	(0.070)		(0.076)	(0.076)	(0.076)
Treatment list × Benchmark		0.016	0.014	0.009		-0.005	-0.000	0.002
		(0.074)	(0.074)	(0.074)		(0.074)	(0.073)	(0.074)
Treatment list × Incumbent × Overall performance (ICW)			0.033				0.023	
			(0.059)				(0.078)	
Treatment list × Benchmark × Overall performance (ICW)			0.081				0.058	
			(0.076)				(0.068)	
Treatment list × Incumbent × National performance (ICW)				-0.030				-0.061
				(0.080)				(0.087)
Treatment list × Benchmark × National performance (ICW)				-0.003				0.026
				(0.081)				(0.076)
Treatment list × Incumbent × Local performance (ICW)				-0.018				-0.081
				(0.079)				(0.079)
Treatment list × Benchmark × Local performance (ICW)				-0.020				-0.032
				(0.081)				(0.080)
Observations	2,893	2,893	2,893	2,893	2,933	2,933	2,933	2,933
Control outcome mean	1.59	1.59	1.59	1.59	1.62	1.62	1.62	1.62
Control outcome std. dev.	0.72	0.72	0.72	0.72	0.73	0.73	0.73	0.73

Notes: Each regression includes randomization block and (baseline and endline) enumerator fixed effects. Lower-order interaction terms are included but not shown. Observations are inversely weighted by the baseline number of respondents surveyed in the village. Standard errors are clustered by village. Given that these hypotheses were not pre-specified, † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from two-sided t tests.

Table I11: Effects of information treatments on polling station-level incumbent vote share, by leaflet content and weight by the share of registered voters within a polling station's experimental village (polling station data)

	Incumbent vote share (proportion turnout)		Incumbent vote share (proportion of registered voters)	
	(1)	(2)	(3)	(4)
Incumbent	-0.001	-0.006	0.003	-0.004
	(0.019)	(0.019)	(0.015)	(0.016)
Benchmark	0.004	0.001	0.003	-0.001
	(0.019)	(0.019)	(0.014)	(0.015)
Incumbent × National performance (ICW)		-0.010		0.004
		(0.023)		(0.015)
Benchmark × National performance (ICW)		0.006		-0.001
		(0.023)		(0.015)
Incumbent × Local performance (ICW)		0.038*		0.028**
		(0.020)		(0.014)
Benchmark × Local performance (ICW)		0.013		0.018†
		(0.020)		(0.014)
Observations	440	440	440	440
Control outcome mean	0.70	0.70	0.41	0.41
Control outcome std. dev.	0.17	0.17	0.13	0.13

Notes: Each regression includes randomization block fixed effects and a lagged dependent variable. Lower-order interaction terms are included but not shown. Observations are weighted by the share of registered voters at the polling station that are registered in the associated experimental village. Robust standard errors are in parentheses. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from pre-specified one-sided t tests.

Table I12: Effects of information treatments on polling station-level turnout, by leaflet content (polling station data)

	Turnout	
	(1)	(2)
Incumbent	0.001	-0.006
	(0.015)	(0.015)
Benchmark	-0.009	-0.011
	(0.011)	(0.012)
Incumbent × National performance (ICW)		0.013
		(0.012)
Benchmark × National performance (ICW)		0.002
		(0.012)
Incumbent × Local performance (ICW)		0.012
		(0.013)
Benchmark × Local performance (ICW)		0.002
		(0.011)
Observations	284	284
Control outcome mean	0.58	0.58
Control outcome std. dev.	0.11	0.11

Notes: Each regression includes randomization block fixed effects and a lagged dependent variable. Lower-order interaction terms are included but not shown. Observations are not weighted, and polling stations where the village in our sample comprises less than 50% of registered voters at the polling station are excluded. Robust standard errors are in parentheses. Given that these hypotheses were not pre-specified, † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from two-sided t tests.

Table I13: Effects of information treatments on turnout (endline survey)

	Turnout (self-reported)						Turnout (validated)					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Incumbent	-0.014	-0.014	-0.014	-0.015	-0.014	-0.001	-0.016	-0.016	-0.016	-0.018	-0.010	-0.032
	(0.018)	(0.017)	(0.017)	(0.018)	(0.019)	(0.025)	(0.020)	(0.020)	(0.020)	(0.021)	(0.021)	(0.028)
Benchmark	0.005	0.005	0.005	0.003	0.008	-0.009	0.009	0.008	0.008	0.003	0.016	-0.025
	(0.018)	(0.017)	(0.018)	(0.018)	(0.019)	(0.025)	(0.021)	(0.021)	(0.021)	(0.021)	(0.022)	(0.030)
Incumbent × Overall performance (ICW)		0.019						-0.000				
		(0.018)						(0.027)				
Benchmark × Overall performance (ICW)		0.021						0.012				
		(0.016)						(0.028)				
Incumbent × National performance (ICW)			-0.000						-0.004			
			(0.016)						(0.023)			
Benchmark × National performance (ICW)			0.005						0.005			
			(0.017)						(0.022)			
Incumbent × Local performance (ICW)			0.012						0.026			
			(0.017)						(0.022)			
Benchmark × Local performance (ICW)			0.007						0.016			
			(0.017)						(0.022)			
Incumbent × Prior index (ICW)				-0.020						0.021		
				(0.018)						(0.019)		
Benchmark × Prior index (ICW)				-0.009						0.022		
				(0.018)						(0.021)		
Incumbent × Prior precision index (ICW)					-0.021						-0.011	
					(0.020)						(0.022)	
Benchmark × Prior precision index (ICW)					0.002						0.003	
					(0.021)						(0.022)	
Incumbent × Performance most important						-0.023						0.030
						(0.034)						(0.038)
Benchmark × Performance most important						0.026						0.062
						(0.033)						(0.038)
Observations	3,874	3,874	3,874	3,801	3,551	3,874	3,876	3,876	3,876	3,803	3,553	3,876
Outcome range	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}	{0,1}
Control outcome mean	0.74	0.74	0.74	0.74	0.74	0.74	0.53	0.53	0.53	0.53	0.53	0.53
Control outcome std. dev.	0.44	0.44	0.44	0.44	0.44	0.44	0.50	0.50	0.50	0.50	0.50	0.50

Notes: Each regression includes randomization block and (baseline and endline) enumerator fixed effects and adjust for the corresponding pre-treatment outcome. Lower-order interaction terms are included but not shown. Observations are inversely weighted by the baseline number of respondents surveyed in the village. Standard errors are clustered by village. Given that these hypotheses were not pre-specified, † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$ from two-sided t tests.