Who Debates, Wins? Experimental Evidence on Debate Participation in a Liberian Election[1]

Jeremy Bowles[2] & Horacio Larreguy[3]

September 27, 2018

Abstract

We study a nationwide debate initiative ahead of Liberia's 2017 elections for House of Representatives designed to solicit and rebroadcast policy promises from candidates. Leveraging random variation in candidates' debate participation, we shock the supply of programmatic information by candidates. We find substantively large effects on citizen learning, political engagement, and voter coordination concentrated in treated districts where a higher share of leading candidates—incumbents and their challengers—were induced to participate. In those districts, challengers decreased their on-the-ground campaigning efforts, while incumbents increased their radio exposure. The initiative electorally benefited incumbents in treated districts, particularly those who performed well and prioritized policy issues more aligned with their electorate. These incumbents, but not their challengers, selected into debate participation based on the quality of their policy platforms and competence. The results point to the importance of understanding selection into the supply of programmatic information when evaluating the effects of its provision.

[1]We are grateful to the J-PAL Governance Initiative and USAID for funding. With thanks to GeoPoll, Internews Liberia, the National Elections Commission of Liberia, NORC at the University of Chicago, the USAID Democracy, Human Rights and Governance Office and USAID Liberia for their cooperation. We thank Kate Casey, Miriam Golden, Torben Iversen, Eric Kramon, Evan Lieberman, John Marshall, Pia Raffler, and Ariel White for useful discussions, and audience at APSA, the Harvard Experiments Working Group Conference, the Harvard Political Economy of Development working group, Internews, Liberia Strategic Analysis, the NYU Center for Experimental Social Science, Twaweza, USAID Liberia, UBC and WGAPE at NYU-Abu Dhabi for their feedback. With special thanks to Morgan Holmes and Jan McArthur for their enduring support, to Anthony Chan, Michael Haines, Anna Langer, Tawedzegwa Musitini and April O'Neill for their cooperation and to Luis da Silva, Theo Harris, Alpha Senkpeni, Kate Thomas, and Jesús Ynfusino for research assistance. IRB approval granted by Harvard Committee on the Use of Human Subjects (ID: IRB17-1178) and NORC at the University of Chicago (ID: 7554.072.01). Study pre-registered with EGAP (ID: 20171024AA) and AEA (ID: AEARCTR-0002553).

[2]PhD Candidate, Department of Government, Harvard University. jbowles@g.harvard.edu.

[3]Associate Professor, Department of Government, Harvard University. hlarreguy@fas.harvard.edu.

1 Introduction

Democratic accountability relies on the effective selection of political candidates and the availability of mechanisms to monitor and incentivize them once in office. Following classic models of electoral accountability (Barro, 1973; Ferejohn, 1986; Fearon, 1999; Holmstrom, 1999), much extant empirical research focuses on the informational dimension of this problem by assessing whether providing citizens with information—generally about their incumbents—affects what they know about candidates, how they evaluate them, and ultimately their voting behavior.[4] However, this work often ignores the strategic calculus underlying the provision of programmatic information by candidates, who in weakly-institutionalized democracies can face strong incentives to reduce the amount of information they provide (Keefer, 2007; Keefer and Vlaicu, 2007). Supplying policy promises can both worsen their electoral prospects if other candidates are better-equipped for programmatic competition (Wantchekon, 2003) and could later restrict their ability to deviate once in office.

A functional media sector might limit the ability of politicians to conceal such information either by supplying it directly or by forcing them to commit to policy promises that affect sizable shares of the electorate, rather than targeting promises to groups of voters on which they can later renege relatively easily (García-Jimeno and Yildirim, 2017). But when the media lacks capacity or is captured (Djankov et al., 2003; Besley and Prat, 2006; Enikolopov et al., 2011), then low-quality, low-information equilibria persist and democratic accountability suffers (Pande, 2011). One particularly effective way to contribute to democratic accountability in these settings, therefore, could be to target both the decision of politicians to supply programmatic information and its dissemination by the media. Shocking the supply of credible policy platforms, and ensuring its dissemination to a large audience might improve candidate selection in the short run and policy delivery in the medium run. This might be particularly so if the extent to which citizens respond to candidate information depends on which candidates have selected into supplying it.

In partnership with USAID and the NGO Internews, we evaluate the impact of randomized elements of a nationwide initiative to hold debates between all 984 candidates for 73 House of Representatives seats ahead of the Liberian election of October 2017.[5] 129 standardized debates across all districts, where only 59% of all candidates running participated, were designed to solicit the policy promises of different candidates in a setting where votes are as often bought as won (Bowles et al., 2017). In the debates, participating candidates were asked a series of questions by moderating journalists on particular issues of national and local policy relevance. Rather than large townhall-style debates, the emphasis here was on soliciting concrete policy platforms and promises from the candidates to be broadcast by community radio stations.

[4] Ferraz and Finan (2008), Chang et al. (2010) and Larreguy et al. (2017) show that media revelations of past incumbent performance enhance electoral accountability. However, a series of localized information dissemination campaigns fail at consistently replicating these media effects (Banerjee et al., 2011; de Figueiredo et al., 2013; Chong et al., 2015; Bhandari et al., 2016; Dunning et al., ming). Closest to our work, recent work studies whether debates can lead citizens to more informed voting decisions (Bidwell et al., 2016; Platas and Raffler, 2017; Brierley et al., 2018).

[5] This is the first time debates have been held universally in an election in, at least, West Africa (Olukotun and Omotoso, 2017).

We randomized the encouragement to participate in the debates by varying the intensity of invitation efforts to attract candidates to attend.[6] The decision to participate in a candidate debate is clearly a strategic one, and particularly so in clientelist settings. Candidates who "win" a debate may enjoy greater publicity and net electoral gains, but *ex ante* they risk either losing a debate or restricting their ability to deviate from policy promises on the campaign trail or once in office. Especially for these contenders, the expected returns from debate participation are limited – they risk providing a platform for their opponents to attack them and gain publicity.[7] We thus intended to generate variation in the share of individuals exposed to the competence and policy promises of the contenders participating in the debates.

With this variation we evaluate a series of hypotheses drawing on a rich set of original data sources. This includes a nationwide panel survey of over 4,000 citizens before and after the intervention, a survey of over 600 candidates who ran in the election, a survey of more than 50 radio stations, full transcripts from debates, around 20 focus groups and polling station-level electoral results. We find that the intervention was successfully delivered. The invitation intervention generated substantively large treatment effects on the debate participation decision primarily of predicted contenders—incumbents and their predicted challengers—and radio stations broadcast the debates as contracted. Around 25% of our citizen survey heard their district debate. Nevertheless, citizens were significantly more exposed to the debate content in treated districts. As our qualitative accounts corroborate, this indicates that citizens were more interested in, and responsive to, debate information when these involved relevant contenders.

The exposure to information about relevant contenders via the broadcasting of debates by radio stations led to increased political engagement by citizens, who engaged in more political information acquisition, and discussion and coordination with others only in treated districts . Increased exposure to the debates also led citizens to update about the competence and policy priorities of the contenders induced to participate. Importantly, such updating was suggestively positive for the incumbents but negative for their challengers. This aligns with qualitative evidence suggesting that incumbents dominated the discussion when they participated. As did the citizens, candidates reacted to the intervention in treated districts but exhibited contrasting heterogeneity in their responses. While incumbents increased their radio campaigning aided by increased demand from radio stations, the challengers reduced their on-the ground campaigning. This suggests a deterrent effect of incumbent debate performance on challengers who fared relatively worse.

Ultimately, consistent with these results—on exposure, learning, engagement and candidate response—the intervention led to improved electoral outcomes for incumbents in the treated districts, as supported by both self-reported and polling-station voting outcomes. Remarkably, 50% of incumbents in those

[6] As discussed later, we also cross-randomized the intensity of radio coverage of the debates. However, such an intervention had no differential results, an candidates were unaware of differential future rebroadcasting efforts, we present pooled results.

[7] Prominent examples abound of incumbents avoiding electoral debates: President Jimmy Carter's in the first 1980 US presidential debate, President Yoweri Museveni in all 2016 Ugandan presidential debates, Primer Minister Theresa May in all debates with other party leaders ahead of the 2017 general election in the UK, President Uhuru Kenyatta in one of the 2017 Kenyan presidential debates, and the incumbent party's candidate Joseph Boakai in one of the 2017 Liberian presidential debates. President Kenyatta noticeably argued that "I decided that he (referring to his main challenger Raila Odinga) will debate alone because I have nothing to debate with him. I will not waste my time there."

districts won re-election compared to 43% in control districts. This vote switching reflected the aforementioned updating on incumbent competence and policy priorities, since it was concentrated in instances where incumbents performed better and supplied policy platforms that better matched the priorities of their electorate. The results, therefore, suggest that shocking the supply of policy promises by candidates led to broad increases in citizen exposure, engagement and learning, but led to uneven electoral consequences that tended to benefit incumbent candidates.

Our results point to the crucial importance of understanding selection into the debates themselves, and more generally the supply of programmatic information in similar contexts. We show that *only* complying incumbents, generally more sophisticated than their challengers, self-selected into debate participation based on the proximity of their policy preferences to those of their constituents and somewhat their competence. In other words, incumbents who complied with the intervention were those for whom the supply of programmatic information and providing a signal of their competence made electoral sense. Other candidates, who often lacked the political sophistication to correctly assess the returns to debate participation, experienced no average gain from showing up.

While fitting into the expansive literature on information and accountability, we contribute in several ways. To begin with, our intervention and results build on, and contrast significantly from, those of previous interventions addressing the effect of localized debates (Bidwell et al., 2016; Platas and Raffler, 2017; Brierley et al., 2018).[8] First, these studies ensured that all (main) candidates were present in the debates in the selected constituencies where they were conducted. We show that candidate attendance cannot be taken for granted when scaling debate interventions at the national level—districts that were not assigned to the intensive-invitation intervention only saw 20% of the incumbents and 40% of the top challengers showing up at the debate—and our results indicate that the effect of such initiatives is conditioned by the attendance decisions of those candidates. This difference in compliance likely explains our contrasting finding that the intervention differentially helped incumbent candidates induced to participate. All previous debate interventions, in turn, find the opposite effect.

This result speaks to recent work highlighting the importance of experimentation at scale, since the effect of small-scale interventions might differ substantially when scaled (Al-Ubaydli et al., 2017; Banerjee and Walton, 2017; Muralidharan and Niehaus, 2017). In particular, Al-Ubaydli et al. (2017) point to individual non-adoption of treatment as a key factor explaining such differences in effects. This lack of compliance is more of a problem for interventions subject to political economy considerations. While development economists struggle with the low take up of agricultural and health products by individuals who would surely benefit from them, politicians represent a harder challenge since they have clear and strong incentives against participating in programs designed to move away from the low-accountability equilibrium from which they often benefit. Our results then highlight the importance of selection into programmatic initiatives to understand their effects when scaled.

[8]Through a similar logic, we inform the gap between the work showing that electoral accountability is enhanced through revelations of past incumbent performance via media (Ferraz and Finan, 2008; Chang et al., 2010; Larreguy et al., 2017), but not necessarily when those revelations are via localized dissemination campaigns (Banerjee et al., 2011; de Figueiredo et al., 2013; Chong et al., 2015; Bhandari et al., 2016; Dunning et al., ming).

Our intervention also differs in its mode of delivery. We focus on debate broadcasting via community radio stations, the dominant form of media consumption across Sub-Saharan Africa (Afrobarometer, 2018). This is relevant for two reasons. First, modes of transmission that reach large shares of constituents—through which voters become aware that many other voters have also received a given piece of information—can produce powerful effects by inducing voter coordination based on common knowledge (Morris and Shin, 2002; Enikolopov et al., 2016; Manacorda and Tesei, 2016; Adida et al., 2017; Arias et al., 2017a).[9] However, more localized modes of delivery, such as public or individual screenings, might lead to relatively greater updating and internalization of the information.[10] Our sizable effects on voter discussion and coordination, which are much weaker when assessed in other debate interventions, suggest that the mode of transmission played a key role in generating the observed effects on voting behavior. Second, the mode of transmission determines the ability and incentives of candidates to strategically respond to the release of information. Relative to localized modes of delivery, media broadcasting should undermine the capacity of candidates to respond by targeting treated communities with on-the-ground campaigning activities (Banerjee et al., 2011; Cruz et al., 2017; Arias et al., 2017b). Moreover, there might be strong complementarities to campaign strategies that follow the mode of delivery. In contrast to Bidwell et al. (2016)'s finding that debates led to increased targeting of campaign expenditure in communities subject to debates screening,[11] our negative results on on-the-ground campaigning by challengers and positive results on radio exposure by incumbents suggest important implications of the mode of delivery for candidates' responses.

Our paper also speaks to recent work on political persuasion. Kendall et al. (2015) study the effect of persuasive campaign messages in a mayoral election in a developed democracy, Italy, and show that voters updated from both valence and ideological campaign messages, and vote accordingly. Closer to the context we study, Cruz et al. (2018) show that, in the clientelistic context of the Philippines, voters who receive comparable information about campaign promises from all candidates are more likely to vote for candidates whose promises are closest to their own preferences. We see our study as complementary. While the intervention by Cruz et al. (2018) is successful at clearly conveying comparable candidate campaign promises to voters to see if these update and vote accordingly, such is at the cost of departing from real campaigns where voters have to parse and compare campaign promises delivered distinctly by each candidate, e.g., via the debates we study. Similarly, our interventions differ in the mode of delivery, which allows Cruz et al. (2018) to isolate from the mentioned important implications for candidate and voter responses that we are interested in. More importantly, our intervention was designed to capture the extent to which the results of information dissemination campaigns is conditioned by the decisions of candidates to supply that information, from which Cruz et al. (2018) abstract to isolate citizen updating

[9] For example, see the mixed evidence from the plethora of studies studying the provision of information through leaflets and scorecards (Adena et al., 2015; Arias, 2016; Arias et al., 2017b; Yanagizawa-Drott, 2014), with broadly more positive evidence on the consequences of media coverage of political issues (Ferraz and Finan, 2008; Larreguy et al., 2017).

[10] For example, while Grossman et al. (2014) find strong effects on citizen participation when directly offering a representative sample of constituents in Uganda with the opportunity to make a request to their representatives via a text message, (Grossman et al., 2017) show no effects when delivering the same opportunity via short radio ads to a comparable sample of constituents.

[11] Brierley et al. (2018) mainly focus on immediate effects of debate exposure on citizen updating and vote intention, and Platas and Raffler (2017) do not report results on candidate responses to their debate intervention.

and the drivers of their voting decisions.

Lastly, our intervention and results also contribute to the recent experimental work designed to reduce clientelistic campaigning practices in developing democracies. Fujiwara and Wantchekon (2013) show that programmatic platforms transmitted through town hall meetings reduced vote-buying and increased electoral support for involved candidates. Vicente (2014), Bobonis et al. (2017), Green and Vasudevan (2018), Hicken et al. (2018) and Larreguy et al. (2018), however, show contrasting results of interventions designed to combat clientelistic and vote-buying practices. Our results underscore the existence of candidate-level variation in the suitability for programmatic competition. This points to the drawing of citizens into candidacy and training as an overlooked determinant in shifting towards more programmatic political equilibria.

The paper proceeds as follows. In Section 2 we outline the Liberian political context, media landscape and the nationwide debates initiative in detail. In Section 3 we justify and describe the two randomized interventions implemented, before outlining our main hypotheses in Section 4. We describe our primary and auxiliary sources of data in Section 5, and report results in Section 7. We discuss our results in Section 8 and conclude in Section 9.

2 Background

2.1 Electoral context

Since its emergence from civil war in 2003, Liberia has held three presidential elections (2005, 2011 and 2017), three House of Representatives elections (2005, 2011 and 2017) and two Senatorial elections (2005 and 2014). The focus in this paper is on the House of Representatives election of October 2017, where each of 73 electoral districts elected a single representative for a six-year term in a first-past-the-post electoral system. Representatives are rewarded handsomely with an annual salary over $200,000 USD in a country with an annual per capita income of around $900 (IREDD, 2016)[12]. Combined with relatively low barriers to candidacy and a fragmented party environment, it is therefore unsurprising that lots of people run for office. In the 2017 election there were 984 candidates for House of Representatives across 26 different political parties, with between 3 and 28 candidates per district. Incumbents sought re-election in nearly 90% of cases. As one opinion piece described this profusion of candidates, "Rest assured that this is not a healthy expression of diverse opinions. Everyone wants a piece of the pie." (Glencorse and Yealue, 2017).

Once in office there is varied, but generally poor, legislator performance. While incumbents attend an average of slightly less than 80% of legislative sessions, some have close to perfect attendance and others attend as few as 45% of sessions (IREDD, 2016). There is, however, widespread dissatisfaction

[12]Representatives often additionally gatekeep access to the rich natural resource wealth of the country by controlling access to concessionary agreements. A major corruption scandal in 2016, for example, revealed that both the Speaker of the House of Representatives and the Chair of the incumbent Unity Party had conspired to illegally broker an iron ore concession in exchange for bribes (Global Witness, 2016).

with incumbent performance, with 67% of citizens reporting mistrusting their representative, 38% believing that representatives never listen to what citizens say, and 68% disapproving of their incumbent's performance (Afrobarometer, 2015). This is not the result of citizens being unaware of who their legislator is—a remarkable 92% of our citizen survey correctly name their legislator—but the paucity of mechanisms to hold politicians to account for poor performance. Largely as a result of an underdeveloped media sector, there is a paucity of credible information about political activity. Moreover, access to politicians is rendered difficult by both poor infrastructural conditions and a limited telecommunications network. For example, 51% of citizens report only seeing their representatives at election time every six years (USAID, 2018).

2.2 Who runs for office?

Combining high personal returns to public office with low oversight can affect selection of entrants into the political market (Brollo et al., 2013). We draw on evidence from an original survey of candidates to provide some descriptive evidence on their characteristics.[13] Throughout this paper we distinguish between three *predicted contenders* per district and *other candidates* for theoretical, practical and measurement reasons. We further split *predicted contenders* into *incumbents* and *predicted challengers*. The objective was not to predict actual election outcomes but to facilitate analysis by identifying three candidates per district who had genuine chances of success – qualitatively, there exists a long tail of candidates who run primarily to enhance their profiles and secure post-electoral favors (Spatz and Thaler, 2018). Moreover, the definition of *actual contenders*, those whose vote share ranked in the top three of their district, might be endogenous to our intervention. This assignment of 219 *predicted contenders*—64 *incumbents* and 155 *predicted challengers*—and 765 *other candidates* is described fully in Appendix A.1.

Table 1: Candidate characteristics

Candidate type	Age	University educated	Ran before	Govt. job before	NGO job before	Advocacy experience	Campaign expenditure	Radio station
Incumbent	55.8	0.68	1.00	0.48	0.35	0.87	$61,458	0.16
Challenger	48.9	0.64	0.43	0.30	0.38	0.88	$41,282	0.06
Other	47.7	0.53	0.22	0.32	0.42	0.85	$30,083	0.03

Table displays mean values of column variables across incumbent, challenger and other candidate surveys. 'Age:' candidate age in years; 'University educated:' candidate has completed university; 'Ran before:' candidate ran for office before; 'Govt. job before:' candidate has held non-elected government job before; 'NGO job before:' candidate has worked for an NGO before; 'Advocacy experience:' candidate reports having worked on an advocacy campaign before; 'Campaign expenditure:' self-reported campaign spending in USD; 'Radio station:' candidate either owns or manages a radio station.

In Table 1 we provide descriptive statistics by candidate category. Candidates generally come from Liberia's elite, being relatively more educated, and are overwhelmingly male (84%). Incumbents are much older and possess higher levels of education than challenger candidates. They are much more likely to possess prior experience in a non-elected government job and less likely to have experience working for an NGO. Almost a third of all candidates have previously run for office, and nearly all report

[13] We successfully surveyed 612 candidates, and the response rate is balanced across treatment groups.

experience in advocacy campaigns in their districts. Candidates report spending substantial amounts – on average above $30,000 – on their campaigns. Incumbents, however, report spending 50% more than predicted challengers and 100% more than other candidates. Additionally, 16% of incumbents reported either directly managing or owning a radio station in their district. Such differences are consistent with a substantial literature on the resource advantages enjoyed by incumbents in developing democracies.

2.3 Campaigning and policy promises

Electoral campaigns are marked by local rallies where candidates travel between communities distributing gifts in cash or kind to generate support, while making local non-credible policy promises to build local infrastructure and anything else that they expect to generate votes. Nearly 80% of surveyed candidates reported visiting most or all communities in their district, while nearly half of surveyed candidates reported distributing gifts in most or all communities. During campaigning season incumbents, especially, orchestrate the mass turnout-buying and trucking of voters from the capital to their districts (Bowles et al., 2017). A USAID survey in 2015 indicated that 49% of citizens believe that “many" or “almost everyone" accepted gifts from parties in exchange for their vote and that 35% of respondents were personally given money in exchange for their vote (USAID, 2015).

In this clientelist context, candidates face few incentives to make and widely disseminate policy promises. Candidates themselves point to systematic differences in the credibility of policy promises delivered at local rallies versus over the radio. In Table 2, we show that candidates point to differences in the *types* of promises made on radio versus at rallies, and are overall more likely to believe that promises made on the radio are more credible compared to those made at rallies. Interestingly, incumbents appear to be more sophisticated in this regard, and candidates overall point to the low likelihood of campaign promises being kept.

Table 2: Candidate attitudes towards policy promises

Candidate type	Different promises	Rally credibility	Radio credibility
Incumbent	0.73	0.19	0.26
Challenger	0.70	0.12	0.14
Other	0.67	0.16	0.15

Table displays mean values of binary column variables across incumbent, challenger and other candidate surveys. ‘Different promises:’ candidate believes that different promises are made on radio versus in-person campaigning; ‘Rally credibility:’ candidate believes that promises made at rallies are likely to be fulfilled; ‘Radio credibility:’ candidate believes that promises made on radio are likely to be fulfilled.

Consistent with these figures indicating the lack of incentives to publicize policy promises, the wide dissemination of policy platforms across remains extremely rare. For an illustration, one of the country’s most prominent newspapers—the Daily Observer—built a ‘promises tracker’ ahead of the election where candidates could specify their policy platforms to appeal to voters and commit to implementing specific

projects. No incumbents did this. As such, candidate campaigns broadly lack policy platforms but rather focus on promises of local development delivered through on-the-ground campaigning, on which they usually renege, and vote buying.

This broad absence of programmatic information is facilitated by a fractured media landscape. Radio stations are an potentially important source of access to politically-relevant information: radio ownership is high at 83%, and 62% of Liberian respondents report listening to news on the radio every day (Afrobarometer, 2015). However, with a lack of regulation in an unconsolidated market with over 100 community radio stations, sporadic access to electricity and scarce sources of commercial revenue, radio stations frequently become the mouthpieces of particular political figures and local firms (Kamara, 2017). Indeed, as Table 1 shows, many incumbents actually own their own radio stations and many more candidates are informally connected to other stations.

3 Candidate debates

3.1 Debate structure

Internews Liberia led an unprecedented nationwide debates initiative in the run-up to the October 2017 elections for House of Representatives,[14] to push back against Liberia's clientelist equilibrium and towards the beginnings of a programmatic one. First, Internews partnered with several Liberian journalist associations to organize debates across different parts of the country.[15] In each district, a local journalist was responsible for conducting research about the issues relevant to constituents, publicizing the debate, and moderating it. In districts with a high number of candidates multiple debates were held, generally on the same day, with candidates randomly assigned to a specific debate. The first debates took place in mid-August, and ran through until mid-September 2017. In total, 129 debates were held across all 73 districts.

Debate venues were mostly administrative buildings, town halls and schools. Debates ran for 90-120 minutes with a simple, uniform structure of up to five questions being asked to all candidates in attendance. Every candidate was given an opportunity to respond to each question, with time limits on responses of three minutes. Candidates were asked to start by outlining their campaign promises. The first question in each debate was related to the management of the County Social Development Fund (CSDF), which is poorly managed with little oversight or input from citizens. Second, candidates were asked about how they would spend their Legislative Support Project (LSP) funds. After these standardized questions, candidates were asked 2-3 questions based on research conducted by the moderator in the district about locally-relevant issues. Moderators specifically intervened to prevent candidates from making personal attacks on other candidates.

Second, Internews also organized the dissemination of the debate content to a broader electorate

[14]Some debates were also held for the presidential race by other organizations which are not the focus here.

[15]These partners were the Press Union of Liberia (PUL), Liberia Media for Democratic Initiatives (LMDI) and the Center for Media Studies and Peacebuilding (CEMESP).

Figure 1: District debate (Montserrado D3)

beyond the few voters attending.[16] Each debate had at least one community radio partner present to broadcast and later rebroadcast its content. Internews chose 43 radio stations to rebroadcast the debates, selecting based on the signal strength of the station to maximize audience sizes and discounting any stations which were owned by candidates running for office in that district.[17] Using geographical data on signal coverage and the geocoded 2008 census, we estimate that nearly 90% of the population was covered by a signal from the station broadcasting that district's debate.

To evaluate the importance of the supply of programmatic information by contenders, we randomized debate invitation effort at the between-district level.[18] We also cross-randomized the extent of debate rebroadcasting, which ultimately had no effect. The relative absence of effects was due to citizens frequently hearing their district debate even in districts without intensive rebroadcasting, likely reflecting the level of citizen interest in the debates. As a result, and since candidates were unaware of differential future rebroadcasting efforts,[19] we present results where we pool over rebroadcasting intensity for clarity of exposition.

We designed the interventions to speak to our theoretical motivation without depriving candidates or voters from opportunities they would have otherwise received. In particular, rather than experimentally varying the extensive margin of whether candidates were invited at all, we significantly increased the intensity of activities already planned to invite candidates to evaluate the relevance of debate attendance

[16]Total audiences were around 100 people. Election-related violence is a concern in Liberia and so, to minimize the risk of conflict, in-person audience for the debates was kept small. Each candidate was given 5 tickets to invite their supporters.

[17]The debates were broadcast by fewer than 73 radio stations since some had the ability to broadcast debates in more than one district. Very few stations were discounted due to political affiliations, primarily because they could not be guaranteed to replay the debates in full with no editing.

[18]Randomizing invitation effort at the candidate-level, while cleaner experimentally, would have raised serious concerns in terms of ethics and fairness to candidates.

[19]In our candidate survey, candidates believed that debates would be rebroadcast roughly two times with no statistical difference by treatment assignment.

by leading contenders.

3.2 Invitation intervention

Whether to participate in a debate represents a strategic decision by candidates to select themselves into the provision of programmatic policy information. Targeting this supply-side decision, the intervention was intended to induce random variation in the debate participation of candidates across districts. To do this, together with Internews, we randomized the intensity of debate invitations to candidates. Candidates in control districts were contacted by the relevant Liberian partner who invited them to attend and provided logistical information about the debate. These invitations frequently did not reach candidates until very close to the debate date, and in some instances not at all.[20]

In treatment districts, we added three components in addition to the partner invitation. First, we sent official invitations and emails from Internews with USAID branding as far ahead of the debates as possible with logistical details and clear contact information for candidates to contact Internews if they had any doubts about the debates.[21] Second, we made phone calls to all candidates around two days before each debate to persuade them to attend. These were mostly conducted by a high-profile Liberian radio journalist who is widely known and respected by local politicians. In these calls, candidates were reminded why they should attend the debates, and their concerns about any elements of debate organization were addressed. Third, we sent SMS reminders to all candidates on either the evening before, or the morning of, the debate with information on where to go.

As such, our invitation intervention carried both behavioral elements (e.g. reminding candidates about the debates in the middle of a busy campaigning season) and more persuasive ones (e.g., reducing fears about the bias of debate moderators and arguing why it is their democratic duty to participate in the debates). Further, the invitations reduced the range of uncertainty candidates faced regarding their participation decision. This was through providing much more information on what to expect and clarifying that the structure of the debates was entirely focused on policy platforms rather than exchanging attacks.

While the intervention was at the debate-level, due to fairness and ethical concerns associated with individual-level randomization, we did not expect homogeneity in the response of candidates within a race to the treatment. In particular, as highlighted by our pre-analysis plan, we expected that the treatment was more likely to affect the marginal decision of serious contenders. For these candidates there were both higher campaigning opportunity costs to debate participation, as well as potential downside from performing poorly in the debate. For less-relevant candidates with more limited resources, debate participation offered a much clearer positive expected return and so we expected that the intervention would have less effect on them.

[20] In our candidate survey over 50% of non-participating candidates cited late notice as their reason for not being present at their debate, while 30% claim that they did not receive an invitation.

[21] USAID was the donor funding the overall initiative. We expected their branding to be persuasive since our candidate survey indicates that more than 43% of the candidates report having worked for an international NGO.

3.3 Debate participation

While the debates were overall well-attended, there was substantial variation across districts. A total 59% of candidates participated in the debates, with attendance across districts varying between 11% and 100%. 48% of incumbents participated compared to 60% of challengers and other candidates. As shown in Table 3, the reasons cited for participation by our candidate survey respondents also varied across candidate types. While non-incumbent candidates were most likely to cite their democratic duty as explaining their attendance decision, incumbents were far more likely to cite the opportunity to showcase their policy platforms to voters. Candidates equally viewed the debates as providing publicity for their campaigns, while challengers put more emphasis on the debates allowing them to demonstrate their competence compared to other candidates. Only a small share of candidates pointed to radio broadcasting as explaining their participation, or admitted to being induced to attend by the possibility of attacking other candidates.

Table 3: Reasons cited for debate participation

Candidate type	Duty	Policies	Competence	Publicity	Radio	Attack
Incumbent	0.40	0.80	0.27	0.40	0.07	0.07
Challenger	0.61	0.48	0.37	0.35	0.02	0.07
Other	0.54	0.52	0.25	0.43	0.01	0.09

The table displays mean values of binary column variables across incumbent, challenger and other candidate surveys. The candidates were able to cite more than one reason for debate participation. 'Duty:' cited democratic duty; 'Policies:' cited opportunity to present policy platform; 'Competence:' cited opportunity to show off competence; 'Publicity:' cited opportunity for free campaign publicity; 'Radio:' cited the benefits of radio broadcasting reaching a large audience; 'Attack:' cited opportunity to attack other candidates.

Our candidate survey is also informative about the reasons that explain why candidates were not present at their debate. Over 50% of non-participating candidates cited late notice, while 30% claimed that they did not receive an invitation. The weak organizational capacity of some of the partner organizations implementing the debates effectively resulted in frequent short-notice changes to debate logistics. Nearly 20% mentioned road conditions to justify their absence, which is unsurprising since Liberia's rainy season renders many rural areas near-impassable. For the many candidates who live in the capital, this generated difficulties in accessing the debates.

The performance of those attending the debates varied substantially. Table 4, using analysis of the debate transcripts, shows that the unbiased rules of debate moderation were generally kept and candidates were given equal time to outline their policy priorities. However, incumbents spoke substantially more about both policy issue-focused questions—the County Social Development Fund (CSDF) and the Legislative Support Project (LSP) funds—likely reflecting their increased experience. Further, they were much more likely to both be attacked by other candidates and attack others, reflecting their attendance likely acting as a focal point for other debate participants.

Qualitative evidence from focus groups suggests that citizens were affected by candidate participation in the debates, as well as the novelty of focusing on concrete policy platforms compared to typical

Table 4: Transcript descriptive statistics

Candidate type	Intro words	CSDF words	LSP words	Attacked	Attacker
Incumbent	340.3	398.2	224.0	0.19	0.15
Challenger	352.0	284.7	218.0	0.04	0.04
Other	345.9	269.8	203.7	0.03	0.03

Table displays mean values of column variables across incumbent, challenger and other candidate surveys. 'Intro words': number of words spoken in introduction; 'CSDF words': number of words spoken about ways to improve management of County Social Development Funds; 'LSP words': number of words spoken about priorities for spending Legislative Support Projects funds; 'Attacked': candidate was verbally attacked by another candidate; 'Attacker': candidate verbally attacked another candidate.

campaigning. As one participant said, "before the debate, the word 'platform' was a strange word to me."[22] Many commented on the debates increasing the accessibility of candidates, indicating that "in the past, there was no opportunity created for voters to engage candidates in understanding their platforms, and most candidates were not accessible to the electorate."[23] As a result, it is not surprising that citizens noted participation decisions, highlighting that "we wanted to see all the six candidates at this debate but only two appeared, which is not good because we are not hearing from other four candidates."[24] Similarly, participants stressed that "all the candidates did not appear for the debate, which was not a good sign,"[25] and that "there should be a law binding all candidates to attend the debate... You can't be somebody who wants to represent me if you don't turn up."[26]

Our qualitative evidence also suggests that the debates caused changes in the assessment of different candidates by voters. One participant highlighted that "the debate changed my attitude toward candidates and helped me discover the hidden secret of some candidates."[27] Similarly, another participant mentioned that "for me, when I reached there, the first person I wanted to vote for ... well, my mind did not go on him. When I entered inside the debate and heard them speak my mind started going on another candidate."[28] Several focus groups pointed to the the lack of specificity and the mixed quality of policy platforms. As an example, one participant indicated that "some of the candidates were not detailed in their explanation on how they going to tackle these sectors."[29] In particular, some respondents argued that the policy platforms of challengers were often weaker than those of the incumbents, as exemplified by one indicating that "I did not hear anything new from candidates contesting against the incumbent because the incumbent was already doing most of these things."[30] Finally, citizens in our focus groups pointed to the importance of the mode of delivery for future accountability by highlighting that "the radio broadcast is a way of record-keeping. If you break your promise once elected, you will be on record for that."[31]

[22] Vai Town, 26 September 2017.
[23] Foya, 20 September 2017.
[24] Massabolahun, 21 September 2017.
[25] Voinjama, 21 September, 2017.
[26] Vai Town, 26 September, 2017.
[27] Kolahun, 18 September.
[28] Klay, 27 September 2017.
[29] Voinjama, 12 September.
[30] Kolahun, 18 September 2017.
[31] Klay, 27 September, 2017.

4 Hypotheses

We assess the consequences of this intended shock to the supply of policy promises by contenders, to be subsequently disseminated by the media, in a context characterized by limited programmatic competition. We evaluate the impact of this random variation along a series of dimensions which relate to candidate participation in the debates, citizen exposure to the debates, citizen beliefs about candidates' competence and policy promises, citizen political engagement, candidate campaigning, and resulting voting behavior.

Generally, following pre-registered approach, we speak about overall treatment effects and, for simplicity, refer to high-intensity invitation districts as treated districts.[32] However, we expected *predicted contender* candidates—incumbents and challengers, as described in Section 2.2—to be differentially affected by the invitation treatment. We represent the overall flow of our pre-registered hypotheses in Figure 2, where specific causal links are associated with our central hypotheses stemming from the invitation treatment. We document the limited divergences from our pre-analysis plan in Appendix A.3.

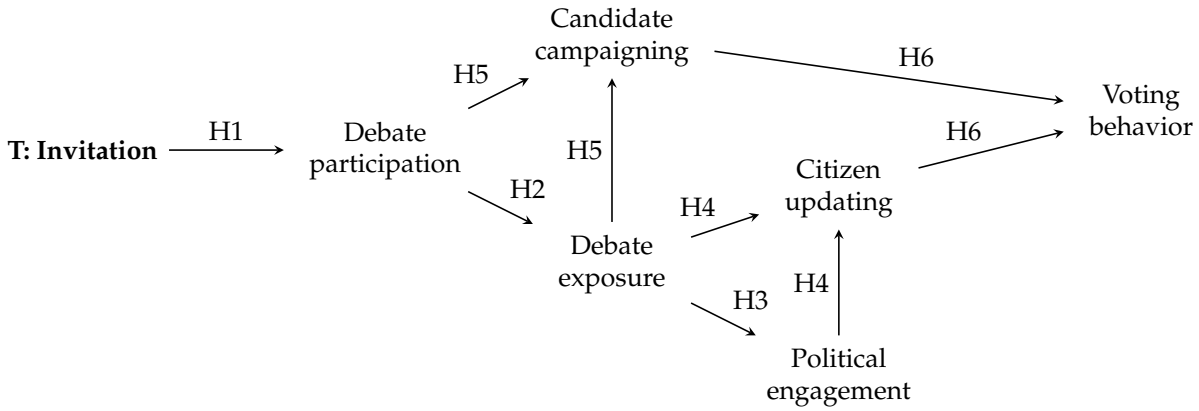

Figure 2: Hypothesized effects of intervention

H1. A higher share of candidates—especially predicted contenders in the election—attended the debate in treated districts.

As discussed in Section 3.2, the invitation intervention tried to induce candidate participation in the debates by ensuring they were fully informed, persuading them to participate, and allaying any concerns. Our first hypothesis, therefore, is that this added recruitment effort led to increased participation in the debates. We hypothesized that a greater share of the candidates in treated districts would participate, and that this effect would be concentrated among the contenders in each race. Incumbents and challengers faced higher opportunity costs and risks of participation, and we anticipated that the invitation intervention would have a larger effect on their participation decisions.

[32] In our pre-registered analysis, we made hypotheses based on the intensity of the intervention, as a function of both the invitation and rebroadcasting treatments, rather than referring to the effects of each intervention individually. As our result, our hypothesis are largely unchanged despite pooling the rebroadcasting treatments.

H2. Citizens were more exposed to the debates in treated districts.

We anticipated that citizens in treated districts were more exposed to the debates since a higher share of contenders participating would increase citizen interest in debate content. This hypothesis, while important, is not obvious because the campaigning season is fraught, with the airwaves full of content from news programs and campaigns, and our intervention represented a relatively small signal in a great deal of electoral noise. We nonetheless expected that this novel form of political communication would attract a great deal of citizen interest when contenders attended the debates. Further, with each debate rebroadcasted an average of 6 times leading up to the election, we expected that citizens would have sufficient opportunity to listen to the content.

H3. Citizens increased their political engagement in treated districts.

Due to the novelty of the debates initiative we anticipated broad effects on political engagement. Our primary expectation was that the mode of delivery would, here, make a difference. Based on the large body of research on the political role of media, we hypothesized that the broadcasting of information about contenders would lead to increased discussion of the debates with others and more coordinated vote choices. Affecting common beliefs about political information (Morris and Shin, 2002; Arias et al., 2017a), as well as increasing discussion of how candidates performed in the debate, should amplify the direct exposure effects of individuals to the debate content. Further, we expected that exposure to the new form of political communication would increase citizen demand for political information, both through their social networks and from the media.

H4. Citizens' evaluations of candidate policy promises and competence were affected by debate exposure in treated districts.

We anticipated that exposure to the debates in treated districts would lead citizens to learn more about candidates. With so many candidates and political parties vying for victory, we expected that listening to a set of candidates answer standardized questions would offer an unusual opportunity to compare across them and update about their policy positions and competence. We hypothesized, therefore, that citizens in treated districts would learn more about the policy proposals and competence of the candidates affected by the intervention.

H5. Citizens in treated districts did not experience more intensive on-the-ground campaigning by candidates in the run-up to the election.

In contrast to localized debate interventions (Bidwell et al., 2016), we hypothesized that the ability of candidates to increase their on-the-ground campaigning efforts in response to the debates would be limited. Since the debates were broadcast across the entire district, we considered it unlikely that candidates would be able to spatially target ground campaigning effort towards areas were more people

were exposed to the debate. Further, since candidates were unaware of rebroadcasting plans, candidates might have had little time to respond to the rebroadcasts themselves.

If at all, debates could crowd out on-the-ground campaigning by candidates who attended to the debates. By attending, candidates also commit themselves—at some level—to campaigning on more programmatic basis rather than focusing on the distribution of cash and promises at the village-level. Similarly, effects on candidate campaigning response could also occur through a deterrence effect since, if the debates benefit only certain candidates, their opponents might reduce campaigning effort.

> **H6a.** Citizens updated their vote choice towards well-performing candidates in treated districts.
>
> **H6b.** Citizens updated their vote choice towards candidates who matched their preferences more so in treated districts.

We anticipated that citizens in treated districts would be more likely to change their vote choices based on the debates. This was due to a direct exposure channel—we expected exposure to the debates to increase access to information about the participating candidates. As such, we anticipated that participating candidates would often benefit electorally due to the debates increasing citizen awareness of their competence and policy priorities. However, as we discuss above, debate participation might not have obvious returns for candidates since they risk losing. Therefore, we are most likely to observe citizens updating their vote choice towards candidates when they not only participated, but also performed well in their debate. Similarly, we expected that citizens would change their vote towards candidates who better matched their policy preferences.

5 Data

5.1 Data sources

Our primary data source is a panel survey of registered 4,060 voters conducted across all 73 electoral districts in the country. In these interviews, enumerators used tablet computers while making phone calls to respondents sampled from the universe of active cell phone numbers for the country's largest mobile network. The distribution of observations per electoral district naturally reflects cell phone penetration and rurality, with an average number of endline observations per electoral district of 73.3. As the descriptive statistics shown in Table 6 indicate, the sample is older, more male and better educated than the average Liberian.

In Figure 3 we show a timeline of the debates and data collection. Our survey began in early August, right before the first debates. Most data collection was completed by early September but concluding the baseline survey in the final electoral districts took several more weeks.[33] The overlap of the baseline

[33]Since the cellphone-number sample was stratified at the county-level, sampling within particular districts proved difficult especially when one county contained both urban and rural districts since, in these cases, most calls went to those in the urban districts, and so achieving sufficient sample in the more rural districts took longer than anticipated.

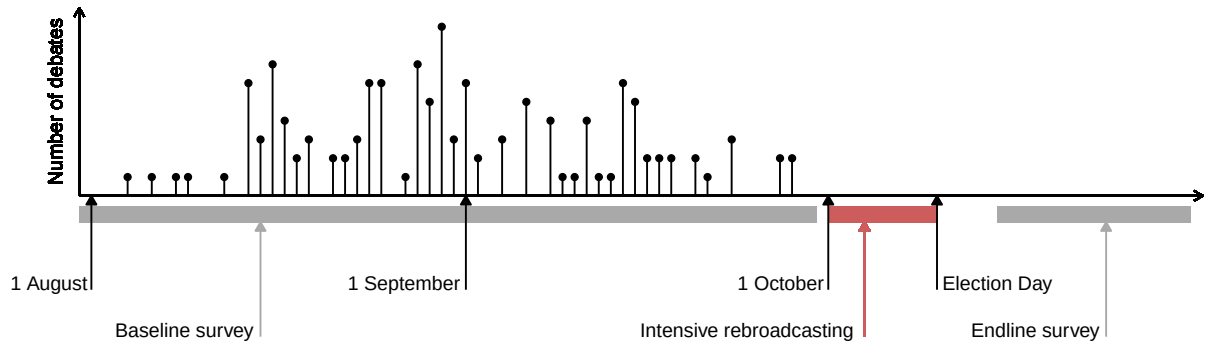

Figure 3: Timeline of debates and data collection

survey and the live debates is not a major concern. First, when we conduct a panel analysis – which we do whenever the outcome variable was collected in both the baseline and endline surveys – we control for baseline debate exposure using the date on which respondents were interviewed. Second, for the rest of the variables which were only collected during the endline survey, the timing of the baseline survey is irrelevant. Lastly, the intensive rebroadcasting of debates took place beginning October 1, by which time 88% of baseline data had been collected. The bulk of respondents actually heard the debates during this rebroadcasting period.

We use several other auxiliary data sources. First, we use polling station-level election results to assess effects on electoral outcomes. Second, we conducted a survey of more than 60% of the candidates who ran in the election. We employ this data to provide descriptive evidence on candidacy, validate important aspects of the intervention, and provide further evidence on the mechanism behind our differential effects by incumbency. Third, we use debate transcripts from each debate. Internews partnered with the Daily Observer newspaper and hired trained journalists to transcribe each debate and extract promises from each candidate. Fourth, with Internews we conducted a survey of over 50 radio stations to gather more descriptive evidence about these stations and to validate their frequency of rebroadcasting the debates.

5.2 Outcome variables

To assess whether the debates and their rebroadcasting, as well as the invitation intervention, were properly implemented, we exploit two main pieces of data. For candidate debate participation, we use data from administrative debate reports as well as full transcripts of each debate. For radio rebroadcasting we use data from the rebroadcast schedules contracted with each of the radio stations and Naymote, a youth organization, which was hired to tune into each scheduled transmission to ensure debates were being played on schedule. We complement this data with responses to our survey of radio stations to assess whether contracted and non-contracted stations also rebroadcasted the debates or related content at other times.

To measure our key outcome variables, we rely on our voter survey and polling station-level data. For all outcome variables, we provide general descriptions in the relevant regression tables and details on their construction in Appendix A.4. Whenever relevant, we aggregate related outcome variables using standardized z-scores. For debate exposure, we asked respondents a battery of questions about both self-reported hearing the debates, as well as factual questions to validate their exposure. We assess effects on political engagement using questions about their demand for political information and coordination within their social networks, and on turnout using both survey and administrative data.

At the respondent-candidate level, we only asked respondents about three *predicted contenders* in each district as per Section 2.2, whose coding we detail in Appendix A.1. This is both because asking about up to 28 candidates would have made the survey prohibitively time-consuming, and because we had theoretical reasons to expect that the invitation intervention should differentially affect the attendance decision of the most relevant candidates. We ask about exposure to different campaigning efforts of each of these predicted contenders and respondent beliefs about the competence and policy platforms of these candidates. For all respondent-candidate dyads, we split the analysis into a pooling of all three predicted contenders, the incumbent, and the predicted challengers. Ultimately we study respondent vote choices, which we validate using polling station-level data.

5.3 Interaction variables

As per H6a and H6b, we expected that voting outcomes would be affected by two key variables: candidate performance in the debate, and the extent of preference matching between respondents and candidates. In our survey analysis, we measure debate performance based on an estimator of how many *other* respondents in the same district responded that a given candidate won the debate. For our polling station-level analysis, we also construct a district-level measure of debate performance, simply defined as the share of respondents in that district who named a given candidate as their debate winner. In contrast with other debate studies (Bidwell et al., 2016; Platas and Raffler, 2017), using an expert panel to measure performance was unfeasible here due to the number of debates actually held.

We measure the extent of preference alignment between respondents and candidates using data from our baseline survey where we ask respondents to name the top three policy issues in their district, as well as to name the perceived top three policy priorities for each of the three predicted contenders. We aggregate this latter measure to the district-level across respondents to create a measure of that candidate's policy priority issues. For the individual-level analysis, we then calculate the share of the respondent's top issues which are shared with a given candidate to create a measure of preference alignment—a maximal value of 1 means that respondents share all their top issues with the top priorities of a given candidate. As with debate performance, we create a district-level version where we calculate the average of this variable at the district-level. In contrast with the district performance measure, due to the differences between our respondent sample and the average Liberian that we show in Table 6, we consider this to be a very noisy measure of the match between candidates' policy priorities and *all* voters

in the district, but provide those results for completeness.

6 Estimation

6.1 Treatment assignment and balance

We randomly assign all 73 districts into two treatment conditions according to whether they received low or high invitation effort, which we respectively refer to as to control and treated districts. To assign treatment conditions, first, we pre-stratified based on which of the debate partners was running that district's debate. This is because the capacity of the debate organizers varied substantially in terms of their ability to attract candidates and organize the logistics of the debates, and in their quality of moderation. Second, we blocked on a set of pre-treatment covariates at the district-level to maximize power.[34] This strategy left us with 19 blocks with between 3 and 4 districts per block. 38 districts were assigned to treatment and 35 to control. Pre-treatment covariates at the district, individual, polling station levels are well-balanced across treatment groups. We also have no meaningful differential attrition in the citizen survey by treatment group, or differential response rates in our candidate survey by treatment group, as well as candidate type. Full details are provided in Appendix A.2.

6.2 Estimating equations

Taking the case where the respondent-candidate is the unit of observation, we estimate:

$$y_{icd} = \beta T + \mathbf{X}_i + \mathbf{Z}_d + \eta_b + \theta_e + \epsilon_{icd}, \tag{1}$$

where y_{icd} is the outcome for respondent i regarding candidate c in district d.[35] T is an indicator for districts where there was high invitation effort, i.e. treated districts. η_b are randomization block fixed effects and θ_e are survey enumerator fixed effects. Throughout, we include both individual-level covariates $\mathbf{X}_i$ and district-level covariates $\mathbf{Z}_d$. While we have good balance on these covariates we include them both to improve precision and to assist with any slight imbalance that we might have. We cluster standard errors at the district level,[36] our level of treatment assignment, and report three pre-registered approaches to weighting. At the individual level, we report specifications which are unweighted, with observations weighted by the inverse of the number of respondents in that district-wave ('1/Obs'), or by the number of registered voters in that district divided by the number of respondents in that district-wave ('Reg/Obs'). At the polling station-level, we report specifications which are unweighted,

[34] We blocked on variables measuring the initially-planned week of the debate, the number of candidates, whether the incumbent was seeking re-election, the log of registered voters in that district, the number of debates to be held in district, the vote share for top 3 candidates in 2011, the vote share Herfindahl index in 2011, the turnout in 2011, the share of candidates who ran in 2011, the log population density, cell phone signal coverage, the share of citizens who owned a radio, and the share of citizens who frequently get news from the radio. Descriptive statistics for all these variables are shown in Panel A of Table 6.

[35] This estimation approach extends to cases where the respondent is the unit of observation, y_{id}, and where the candidate is the unit of observation, y_{cd}.

[36] The only exception is when we report results where the unit of observation is the district in which we use heteroskedasticity-robust standard errors.

with observations weighted by the inverse of the number of polling stations in that district ('1/PS'), or by the number of registered voters at that polling station ('Reg').

Whenever we have a panel for a given question where the outcome is continuous, we consider the continuous change in that variable between baseline and endline as an outcome Δy_{icd}. When the outcome is binary, we construct an indicator for whether the coded response changed between waves. The estimating equation remains the same aside from controlling controlling for whether respondents were interviewed at baseline before or after the first broadcast of their district debate and its interaction with treatment assignment. Lastly, we also make use of specifications where we interact treatment assignment with covariates X_{cd} that might vary within-district, which applies to the interactions discussed in Section 5.3.

7 Results

Our results suggest rich and consistent consequences of the debates initiative on political outcomes. First, we show that the debate intervention was successfully delivered since candidates, and in particular predicted contenders, were more likely to attend their debates in treatment districts. Second, we also find strong evidence of citizens being more exposed to the debates in the intensive-invitation districts, suggesting that the attendance of candidates more relevant to voters was central for voter exposure.

Third, we present effects on political engagement measured by political information acquisition, discussion and coordination with other voters, and turnout. Mirroring the results on debate exposure, we find that effects of the intervention were concentrated in treated districts. The strength of these results aligns well with our focus group evidence and suggests that the intervention politically activated citizens in those districts. These results reinforce that citizens were more interested in the debate content in districts where the most relevant candidates participated.

Fourth, at the citizen-candidate level, we show that citizens learned more about incumbent candidates' policy priorities and competence in treated districts, and to a lesser extent about challengers' policy priorities. There is also suggestive evidence of positive updating about incumbent competence and negative updating about challenger competence in treated districts, which again aligns with the qualitative evidence from our focus groups that suggests that incumbents dominated their debates. Fifth, challengers decreased their on-the-ground campaigning efforts, while incumbents increased their radio campaigning in treated districts. Finally, as a result, in treated districts citizens switched their vote towards their incumbent, but not towards challengers, when they performed well or matched baseline voter policy preferences.

7.1 Intervention implementation

We start by testing H1 and assessing whether the intervention had its intended effect. Table 7 shows treatment effects on the attendance of candidates at the debates. Column 1 in Panel A suggests that

the invitation intervention led to a 7.7 percentage point (pp) (14% relative to the control mean) increase in the share of total candidates running in treated districts. In Panel B, where we subset down to the attendance decision of the predicted contenders in each race, there is a substantially larger treatment effect of 20.1 pp (45%). Decomposing the attendance decision of predicted contenders, Panels C and D respectively show that incumbents were 21.2 pp (76%) and 21.2 pp (43%) more likely to attend in treated districts. Lastly, Panel E shows no treatment effect on other candidates, then reinforcing that the invitation intervention mainly affected contender attendance.[37]

7.2 Exposure to debates

We next report the effects of treatment assignment on outcome variables intended to measure citizen exposure to the debates (H2), which should reflect their differential interest in the debate information about the contenders more likely to attend the debates in treated districts. In Panel A of Table 8, we use a standardized index of our measures of direct exposure to the debates. The results indicate that citizens in treated districts had standardized exposure around 0.3 standard deviations (sd) higher than those in control districts.[38]

In Panel B, we consider as an outcome the change in hearing the debate between baseline and endline. The results suggest a significant and sizable treatment effect on the probability that citizens heard the debate between surveys. This effect, compared to a control mean of 8.4 pp, is around 8 pp. In Panel C we consider as an outcome an indicator for whether the respondent heard the debate and also find the same result. Citizens were around 4 pp (20%) more likely to have heard the debate at endline in treated districts. In Panel D we assess how many times citizens report hearing the debates. There is a significant treatment effect of around 0.09 pp, relative to a control mean of 0.42 pp. Lastly, in Panel E we consider as an outcome a standardized index of the change in how many factual questions about CSDF management, which was discussed during the debates, respondents answered correctly between baseline and endline. There is a significant treatment effect of 0.042 pp (17%), which suggests that citizens internalized debate content which conveyed information about policy issues.[39] These results provide strong support that the invitation intervention meaningfully affected citizen exposure to the debate content through increasing the attendance of leading contenders.[40] As discussed, this was not obvious since the debates were rebroadcasted during a very busy period of campaigning.

[37] In Appendix Table A9, we further show that the share of actual contenders attending, whether the eventual election winner attended, and the share of actual challengers attending was substantially affected by treatment assignment as expected.

[38] Importantly for our ability to pool districts assigned to different rebroadcasting intensity, there were no differential treatment effects on citizen exposure to the debates both when pooling districts assigned to different rebroadcasting intensity, and when restricting to districts assigned to high invitation intensity.

[39] In Appendix Table A10 we validate these results with a series of auxiliary outcomes. We find that, in treated districts, citizens' stated debate winners were much more likely to have actually attended the debate, citizens more accurately name the number of participants, and broadly zero effects on whether citizens report attending their district debates in person.

[40] This is also consistent with more deliberation between relevant candidates increasing voters' attention (Wantchekon et al., ming).

7.3 Effects on political engagement

The qualitative evidence we presented in subsection 3.3 suggests that the debates had a catalyzing effect on political engagement due to the novelty of the initiative. In this section we quantitatively test whether, by increasing the attendance of contenders, the invitation intervention indeed affected the political engagement of citizens and ultimately their turnout (H3). Panel A of Table 9 shows a 0.169 sd increase in treated districts on a standardized information demand index capturing how much political information citizens demanded just before the election through listening to the radio, discussing with friends, and seeking other forms of political information.

In line with these results, in Panel B there are strong treatment effects on a standardized coordination index capturing whether citizens discussed the debates, and were more likely to believe that these discussions led them to coordinate their vote choices in districts assigned to the invitation intervention. Specifically, citizens exhibit a 0.145 sd increase in coordination in treated districts. Both panels then provide strong evidence that, when exposed to debates where candidates more relevant to voters participated, citizens demanded more political information and coordinated their vote choices.

Next, we provide evidence that the citizen activation induced by the intervention ultimately led to higher turnout. In Columns 1-2 of Table 10, we use administrative polling station-level data on turnout. Turnout in the House of Representatives election was 2.2 pp higher in treated districts, compared to a control mean of 70.8 percent. These patterns align remarkably with our survey evidence in Columns 3-8. In Columns 3-5, the evidence indicates that citizens were significantly more likely to state that the debate changed their mind about whether they were going to vote in the election by 1.9 pp (28%) in treated districts.

Similarly, in Columns 6-8, we show that citizens in treated districts were 3.4 pp (27%) more likely to respond that the debates changed *who* they were going to vote for. We show below that this vote switching in treated districts differentially happened towards incumbent candidates when they performed well in the debates, or their policy priorities matched those of their constituents. Overall, the results provide strong evidence for the assignment to debates with increased candidate participation causing changes in political engagement of citizens which mapped onto actual turnout decisions.

7.4 Effects on beliefs about candidates

We next analyze the effect of the intervention on beliefs about the competence and priority policies of contenders (H4). We assess whether citizens learned more about the contenders induced to attend the debates by the increased recruitment effort. For this, we first assess treatment effects on the standardized change in citizens' reported certainty about the competence (columns 1-3) and priority issues (columns 4-6) of predicted contenders in Table 11. Panel B shows that citizens in treated districts became significantly more certain about incumbent competence and priority issues, with standardized treatment effect sizes of 0.251 and 0.275 sd respectively. In Panel C there is no evidence that citizens in treated districts became

more certain about the competence of challengers, and variable evidence that those citizens became more certain about challengers' priority issues.

We then assess treatment effects on the standardized change in citizens' beliefs about the competence (columns 1-3) and priority issues (columns 4-6) of predicted contenders in Table12. These treatment effects directionally suggest positive updating regarding incumbents and negative updating for their challengers, but are imprecisely estimated. Panel B shows suggestive sizable, but statistically insignificant, treatment effects of citizens positively updating about incumbent competence or correctly learning about the policy priorities of their incumbents. Specifically, for both outcomes, the treatment effect is around 0.1 sd. In contrast, Panel C indicates that citizens updated negatively about the competence of their challengers, while they did not learn about the policy priorities of those candidates in treated districts. The treatment effect on challenger competence is 0.1 sd and marginally significant.[41]

The results provide evidence that aligns with the focus group evidence in Section 3.3. Inducing the participation of contenders increased certainty regarding their competence and priority issues among their electorate. This increase was concentrated with respect to the incumbent, who spoke substantially more in response to the policy issue-focused questions where they possessed more experience. Similarly, the suggestive positive treatment effects on citizens perceptions about incumbent competence, but negative about challenger competence, suggest that incumbents outperformed challengers when participating in the debates. Overall the evidence supports that, when induced to attend the debates, in contrast to the challengers, incumbents benefited both in terms of increased voter information about their policy priorities and their relative competence.

7.5 Candidate response and campaigning

Next, we consider treatment effects on the campaigning responses of candidates (H5). In Table 13, we report results where we split up the corresponding survey items into standardized indices of "on-the-ground" campaigning by candidates in respondents' towns (Columns 1-3) and "radio" campaigning (4-6). The on-the-ground campaigning index incorporates candidates' visits, distribution of leaflets and vote buying in respondents' localities. The radio campaigning index captures candidate presence on the radio.

In Panel B, there is a significant increase in incumbent exposure on the radio in treated districts, but no significant treatment effect on on-the-ground campaigning by incumbents. Specifically, there is a positive treatment effect of .082 sd in incumbent radio campaigning. In contrast, Panel C shows evidence of negative treatment effects on on-the-ground campaigning by challengers, but no treatment effect on challenger radio exposure. Challengers reduced their on-the-ground campaigning by .04 sd in treated districts.

This pattern of results suggests that the presence of incumbents in the debates deterred challengers.

[41] In Appendix Table A11 we test for treatment effects on whether citizens name particular candidates as their debate winner. Possibly due to the substantial heterogeneity in voter responses, we find no significant evidence that treated citizens are significantly more likely to name incumbents or challengers as their debate winner.

Moreover, incumbent debate performance led them to increase the use of this strategy to complement the rebroadcasting of the debates. Suggestively, in our candidate survey, 77% of incumbents believe that the debates led radio stations to ask them to appear on more shows compared to 63% of challengers, and 68% of incumbents report that the debates led them to change their campaigning strategy compared to 52% of challengers. Restricting this to candidates who said the debates affected their mode of campaigning, 60% of incumbents cited shifting towards more radio appearances compared to 40% for on-the-ground campaigning. By contrast, only 27% of challengers cited shifting towards more radio appearances compared to 73% for on-the-ground campaigning.

7.6 Voting Behavior

Our results then indicate that the debates, and in particular in treated districts, benefited participating incumbents the most. While both incumbent and challenger candidates were more likely to attend in those districts, politically motivated citizens updated relatively more and more positively about incumbent competence and policy priorities than that of their challengers. In turn, incumbents increased their campaigning efforts over the radio, while challengers reduced theirs on the ground. Here, we provide evidence that this ultimately affected voting behavior.[42]

Table 14 presents a set of specifications testing for whether there were effectively treatment effects on vote choice, defined as switching towards a given candidate between baseline and endline surveys. Columns 1-3 present the main effects of treatment assignment. Columns 4-6 present specifications where we interact treatment assignment with the standardized measure of debate performance described in Section 5.3 (H6a). In columns 7-9, we interact treatment assignment with the standardized measure of preference alignment between the citizen and the candidate described in the same section (H6b).

In Panel B, focusing on the incumbent, we find insignificant, but sizable, main treatment effects on vote switching. Incumbents experienced a 4.4 pp (22%) insignificant increase in vote share in treated districts. We do, however, find evidence of interaction effects with assignment to treated districts. Results indicate that the interaction with *either* debate performance or policy priority match is significantly positive. In contrast, focusing on predicted challengers in Panel C, there are insignificant but broadly negative main effects and little evidence of the interaction terms increasing vote switching towards them. Challengers experienced a 2.9 pp (23%) insignificant drop in vote switching in treated districts.

In Table 15, we use polling station-level data to replicate the survey results, but use the debate performance and priority match measures averaged over our sample of district respondents. In Panel B, we find that incumbent vote share in treated continues to be insignificantly higher than in control districts, though smaller in magnitude, and that this significantly depended upon their debate performance. However, we find no significant interaction with our district measure of priority matching, which we attribute to the fact that the average our sample of district respondents might be a noisy measure of

[42] Appendix Table A12 shows that citizens in treated districts show no differential change in the reasons they cited for their vote choice, which cannot then account for our results on voting.

the true district priority matching. In Panel C, we continue to find that challenger vote share in treated districts is insignificantly smaller than in control districts, and no evidence of interactions with debate performance or proximity in policy priorities between challengers and their constituents. The polling station results broadly support the survey results, particularly suggesting positive treatment effects on electoral support for incumbents that performed well in the debates.

8 Discussion

Our results suggest that electoral gains accrued strikingly to incumbent candidates in treated districts, and particularly so when they performed well in their debates and show that their policy priorities match those of their constituents. Importantly, the results from reported voting and polling station outcomes also show in the overall election outcomes, where 50% of incumbents in treated districts won re-election compared to 43% in control districts. This striking difference is not surprising given that the winning margin in more than 35% of races was less than 5 percentage points. However, especially in a context where approval and performance of incumbents is generally low, as discussed in Section 2.1, these results might seem counterintuitive. Understanding why this happened demands an analysis of compliance with the invitation treatment and hence candidate selection into the debates themselves. Few incumbents attended in control districts—just 35%—and with additional effort this increased to around 50% in treated districts, which still leaves a sizable share of incumbents not attending.

Next, we show here that these incumbents positively self-selected into debate participation to a greater extent than other candidates in terms of both preference alignment with citizens. To begin with, in both our citizen and candidate surveys, we asked respondents to name what they considered to be the three top issues in their districts.[43] We leverage this information to assess the degree of preference congruence between candidates and citizens. For candidate c in district d we define their degree of priority alignment, $align_{cd}$, with the citizens in their district as the average overlap between the candidate's top three priorities and their citizens' top three priorities. More formally, we calculate:

$$align_{cd} = \frac{1}{3N_d} \sum_{i \in d} |p_j \cap p_i|, \tag{2}$$

where N_d is the number of respondents in district d, p_j is the set of the top three priorities of candidate c and p_i represents the set of the top three priorities of citizen i in candidate c's district.[44] In Table 5 we present the mean value of this $align_{cd}$ variable by candidate category where we subset by whether candidates participated in their debate (P) or did not participate (P'). In Panel A we measure candidate

[43] Citizens (in decreasing order of popularity) cited roads, health, education, water and electricity most frequently. Candidates cited education, health, roads, youth representation and agriculture.

[44] This is closely related to the district-level version of the interaction variable for preference matching we describe in Section 5.3, but our use of the candidate survey enables us to generate the measure for candidates beyond the three predicted contenders per district who were asked about in the citizen survey.

priorities using their responses to our survey, while in Panel B we use the measure drawn from the citizen survey (i.e. exactly the measure used in our district-level interactive specifications). In the third column we present p-values testing the difference in alignment measures *within* a candidate category. We present the regressions underlying these statistics in Appendix A.5.

Table 5: Selection into debate participation

	P	**P′**	p-value
A. Policy alignment (Candidate survey)			
Incumbent	0.349	0.191	(0.008)
Challenger	0.272	0.269	(0.922)
Other candidate	0.254	0.252	(0.900)
B. Policy alignment (Citizen survey)			
Incumbent	0.482	0.401	(0.0001)
Challenger	0.450	0.449	(0.948)
C. Candidate competence			
Incumbent	3.64	3.49	(0.057)
Challenger	3.42	3.33	(0.002)

This table compares statistics among candidates who participated in their debate (P) and those who did not (P′). Panel A and B display mean $align_{cd}$ measures as per Equation 2. In Panel A we measure candidate priorities using the candidate survey, in Panel B we use citizen perceptions of a given candidate's priorities. In Panel C we measure candidate competence using citizen perceptions of candidate competence at baseline. The third column presents the p-value of the difference between these figures for a given candidate type based on the regressions described in Appendix A.5.

The results indicate that participating incumbents self-selected into debate participation positively in terms of preference alignment. Participating incumbents, based on their own stated policy priorities, were 15.8 pp (83%) more aligned with citizen preferences in their districts than non-participating incumbents, which is significant at the 1% level. There is no evidence of within-candidate type selection among challengers or other candidates. Panel B replicates this pattern by showing that citizens, at baseline, were 8.1 pp (20%) more aligned with (what they perceived to be) the priorities of incumbents who ended up participating compared to those who did not. This difference is again significant at the 1% level and no similar pattern exists for challenger candidates. Interestingly, citizens uniformly overestimate how closely candidate priorities map onto their own priorities compared to the responses of the candidates themselves.

Second, we test for differences in selection *across* candidate categories. We find that non-participating incumbents were significantly *less* aligned with citizen priorities than participating challengers ($p = 0.023$ using candidate survey, $p = 0.002$ using citizen survey) or participating other candidates ($p = 0.05$). Furthermore, participating incumbents were more aligned than participating challengers ($p = 0.11$ using candidate survey, $p = 0.007$ using citizen survey) or participating other candidates ($p = 0.043$).

In Panel C we find similar evidence of such differential selection into debate participation based on perceived candidate competence. We proxy for such competence using the mean of citizens' perceptions of a given candidate's competence at baseline, on a five-point scale. There is statistically significant evidence of selection into participation within both types of candidates with respect to competence

($p = 0.057$ for incumbents, $p = 0.002$ for challengers). Incumbents were generally perceived to be more competent than their challengers but, while suggestive, the difference between non-participating incumbents and participating challengers is statistically insignificant at conventional levels ($p = 0.266$).

The evidence then suggests that incumbents strategically selected into debate participation, based on their comparative advantage at programmatic competition, while their challengers did not. In other words, the incumbents who complied with the invitation intervention were those with policy preferences closer to those of their electorate and so for whom the supply of programmatic information made electoral sense. Both incumbents and challengers, however, seem to have positively selected on competence. Overall, challenger candidates, with frequently less political sophistication and experience, were likely less successful at calculating the returns to participation or simply took a higher risk since they were likely lagging behind the incumbent.

9 Conclusion

We evaluate an intervention designed to elicit and disseminate programmatic promises from candidates to the House of Representatives in the 2017 elections in Liberia. Policy promises were elicited through debates from participating candidates and disseminated via radio rebroadcasting. The intervention electorally benefited the incumbents induced to attend their debates. These incumbents dominated their debates proposing policy platforms closer to those preferred by their electorates. Our results suggest that both the debate content and the mode of dissemination of such content mattered critically for this outcome. *Only* when relevant candidates attended their debates, voters paid attention to the debate information and subsequently engaged in more political information acquisition, and discussion and coordination with others. Incumbents induced to participate in the debates increased their radio exposure aided by increased demand from radio stations, while their deterred challengers reduced their on-the ground campaigning.

Our results point to the challenges of transitioning away from the low-quality, low-information equilibrium characterizing many developing democracies. By inducing a subset of self-selected candidates to compete on a more programmatic basis, our intervention had uneven electoral consequences particularly favoring incumbents. In this context, there may exist substantial returns to incumbency for some incumbents since challengers are usually less experienced and poorly equipped for programmatic competition. Incumbents, meanwhile, may possess informational advantages enabling them to better evaluate the return to participation in such initiatives. While this impact could be specific to the political and institutional context of Liberia, the point is broader. If such democratic initiatives are to enhance competition and contribute to shifting towards a more programmatic political equilibrium, then they must ensure that candidate incentives are aligned towards participation and that less experienced candidates receive additional training to face a more leveled playing field. How to best achieve this remains a question for future research.

References

Adena, M., Enikolopov, R., Petrova, M., Santarosa, V., and Zhuravskaya, E. (2015). Radio and the rise of the nazis in prewar germany. *Quarterly Journal of Economics*, 130(4):1885–1939.

Adida, C., Gottlieb, J., Kramon, E., and McClendon, G. (2017). Breaking the clientelistic voting equilibrium: The joint importance of salience and coordination. Working paper.

Afrobarometer (2015). Afrobarometer Data, Liberia, Round 6, 2014.

Afrobarometer (2018). Radio remains the most-used source of news in many african countries. Afrobarometer news release.

Al-Ubaydli, O., List, J. A., LoRe, D., and Suskind, D. (2017). Scaling for economists: Lessons from the non-adherence problem in the medical literature. *Journal of Economic Perspectives*, 31(4):125–144.

Arias, E. (2016). How does media influence social norms? a field experiment on the role of common knowledge. Working paper.

Arias, E., Balán, P., Larreguy, H. A., Marshall, J., and Querubín, P. (2017a). How social networks help voters coordinate around information provision to improve electoral accountability: Experimental evidence from mexico. Working paper.

Arias, E., Larreguy, H. A., Marshall, J., and Querubín, P. (2017b). Priors rule: When do malfeasance revelations help or hurt incumbent parties? Working paper.

Banerjee, A. V., Kumar, S., Pande, R., and Su, F. (2011). Do informed voters make better choices? experimental evidence from urban india. Working paper.

Banerjee, Abhijit, R. B. J. B. E. D. H. K. S. M. M. S. and Walton, M. (2017). From proof of concept to scalable policies: Challenges and solutions, with an application. *Journal of Economic Perspectives*, 31(4):73–102.

Barro, R. J. (1973). The control of politicians: An economic model. *Public Choice*, 14(1):19–42.

Besley, T. and Prat, A. (2006). Handcuffs for the grabbing hand? media capture and government accountability. *American Economic Review*, 96(3):720–736.

Bhandari, A., Larreguy, H., and Marshall, J. (2016). Able and mostly willing: An empirical anatomy of information's effect on voter efforts to hold politicians to account in senegal. Working paper.

Bidwell, K., Casey, K., and Glennerster, R. (2016). Debates: Voting and expenditure responses to political communication. Working Paper.

Bobonis, G., Gertler, P., Gonzalez-Navarro, M., and Nichter, S. (2017). Vulnerability and clientelism.

Bowles, J., Larreguy, H., and Liu, S. (2017). How weakly institutionalized parties monitor brokers in developing democracies: Evidence from post-conflict liberia. *Working Paper*.

Brierley, S., Kramon, E., and Ofosu, G. (2018). The moderating effect of debates on political attitudes. Working paper.

Brollo, F., Nannicini, T., Perotti, R., and Tabellini, G. (2013). The political resource curse. *The American Economic Review*, 103(5):1759–1796.

Chang, E. C., Golden, M. A., and Hill, S. J. (2010). Legislative malfeasance and political accountability. *World Politics*, 62(2):177–220.

Chong, A., De La O, A., Karlan, D., and Wantchekon, L. (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, C., Keefer, P., and Labonne, J. (2017). Incumbent advantage, voter information and vote buying. Working paper.

Cruz, C., Keefer, P., Labonne, J., and Trebbi, F. (2018). Making policies matter: Voter responses to campaign promises. Working paper.

de Figueiredo, M. F., Hidalgo, F. D., and Kasahara, Y. (2013). When do voters punish corrupt politicians? experimental evidence from brazil. Working paper.

Djankov, S., McLiesh, C., Nenova, T., and Shleifer, A. (2003). Who owns the media? *Journal of Law and Economics*, 46(2):341–381.

Dunning, T., Grossman, G., Humphreys, M., Hyde, S., and McIntosh, C. (forthcoming). *Metaketa I: The Limits of Electoral Accountability*. Cambridge University Press.

Enikolopov, R., Makarin, A., and Petrova, M. (2016). Social media and protest participation: Evidence from russia. Working Paper.

Enikolopov, R., Petrova, M., and Zhuravskaya, E. (2011). Media and political persuasion: Evidence from russia. *American Economic Review*, 101(7):3253–3285.

Fearon, J. D. (1999). Electoral accountability and the control of politicians: Selecting good types versus sanctioning poor performance. In Przeworski, A., Stokes, S., and Manin, B., editors, *Democracy, Accountability, and Representation*. Cambridge University Press.

Ferejohn, J. (1986). Incumbent performance and electoral control. *Public Choice*, 50(1):5–25.

Ferraz, C. and Finan, F. (2008). Exposing corrupt politicians: The effects of brazil's publicly released audits on electoral outcomes. *Quarterly Journal of Economics*, 123(2):703–745.

Fujiwara, T. and Wantchekon, L. (2013). Can informed public deliberation overcome clientelism? experimental evidence from benin. *American Economic Journal: Applied Economics*, 5(4):241–255.

García-Jimeno, C. and Yildirim, P. (2017). Matching pennies on the campaign trail: An empirical study of senate elections and media coverage. Working paper.

Glencorse, B. and Yealue, L. (2017). Liberia's democracy is failing its people. *Washington Post*.

Global Witness (2016). The deceivers.

Green, D. P. and Vasudevan, S. (2018). Diminishing the effectiveness of vote buying: Experimental evidence from a persuasive radio campaign in india. *Working paper*.

Grossman, G., Humphreys, M., and Sacramone-Lutz, G. (2014). I would like u wmp to extend electricity 2 our village. *American Political Science Review*, 108(3):688–705.

Grossman, G., Humphreys, M., and Sacramone-Lutz, G. (2017). Information technology and political engagement: Mixed evidence from uganda. Working paper.

Hicken, A., Leider, S., Ravanilla, N., and Yang, D. (2018). Temptation in Vote-Selling: Evidence from a Field Experiment in the Philippines. *Journal of Development Economics*, 131(March):1–14.

Holmstrom, B. (1999). Managerial incentive problems: A dynamic perspective. *Review of Economic Studies*, 66(1):169–182.

IREDD (2016). Salaries and allowances of lawmakers in the national budget.

Kamara, K. A. (2017). Media strengthening, public participation and democracy. In Garnett, T., editor, *Liberia Development Conference Anthology*, pages 176–191. USAID.

Keefer, P. (2007). Clientelism, credibility, and the policy choices of young democracies. *American Journal of Political Science*, 51(4):804–821.

Keefer, P. and Vlaicu, R. (2007). Democracy, credibility, and clientelism. *Journal of Law, Economics, and Organization*, 24(2):371–406.

Kendall, C., Nannicini, T., and Trebbi, F. (2015). How do voters respond to information? evidence from a randomized campaign. *American Economic Review*, 105(1):322–53.

Larreguy, H., Marx, B., Reid, O., and Blattman, C. (2018). A market equilibrium approach to reduce the incidence of vote-buying: Evidence from uganda.

Larreguy, H. A., Marshall, J., and Snyder, James M., J. (2017). Publicizing malfeasance: When the media structure facilitates electoral accountability in mexico. Working Paper.

Manacorda, M. and Tesei, A. (2016). Liberation technology: mobile phones and political mobilization in africa. Working paper.

Morris, S. and Shin, H. S. (2002). Social value of public information. *American Economic Review*, 92(5):1521–1534.

Muralidharan, K. and Niehaus, P. (2017). Experimentation at scale. *Journal of Economic Perspectives*, 31(4):103–124.

Olukotun, A. and Omotoso, S. A. (2017). *Political Communication in Africa*. Springer.

Pande, R. (2011). Can informed voters enforce better governance? experiments in low-income democracies. *Annual Review of Economics*, 3(1):215–237.

Platas, M. and Raffler, P. (2017). Meet the candidates: Information and accountability in primary and general elections. Working paper.

Spatz, B. J. and Thaler, K. M. (2018). Has liberia turned a corner? *Journal of Democracy*, 29(3):156–170.

USAID (2015). Liberia Electoral Access and Participation (LEAP) Survey. *United States Agency for International Development*.

USAID (2018). Liberia Electoral Access and Participation (LEAP) Survey. *United States Agency for International Development*.

Vicente, P. C. (2014). Is Vote Buying Effective? Evidence from a Field Experiment in West Africa. *The Economic Journal*, 124(574):F356–F387.

Wantchekon, L. (2003). Clientelism and voting behavior: Evidence from a field experiment in benin. *World Politics*, 55(3):399–422.

Wantchekon, L., Lopez-Moctezuma, G., Fujiwara, T., Lero, C. P., and Rubenson, D. (Forthcoming). Policy Deliberation and Voter Persuasion: Evidence from an Election in the Philippines. *American Journal of Political Science*, 131(March):1–14.

Yanagizawa-Drott, D. (2014). Propaganda and conflict: Evidence from the rwandan genocide. *Quarterly Journal of Economics*, 129(4):1947–1994.

Tables

Table 6: Descriptive statistics

	Mean	SD	Min	Max
A. District-level variables (n = 73)				
Scheduled debate week	4.18	1.39	1.00	8.00
Number of debates in district	2.08	0.66	1.00	4.00
Number of candidates (2017)	13.55	4.81	3.00	28.00
Incumbent ran in election (2017)	0.84	0.37	0.00	1.00
Share of repeat candidates (2017)	0.26	0.11	0.06	0.53
Log registered voters (2017)	10.23	0.40	9.27	11.06
1st voteshare (2011)	0.31	0.13	0.12	0.82
2nd voteshare (2011)	0.18	0.05	0.10	0.36
3rd voteshare (2011)	0.13	0.03	0.05	0.25
Voteshare HHI (2011)	0.19	0.11	0.07	0.69
Turnout (2011)	0.66	0.05	0.56	0.75
Log population density (2008)	−9.51	1.76	−11.91	−5.21
Share over 18 (2008)	0.48	0.02	0.43	0.54
Share with secondary education (2008)	0.15	0.05	0.04	0.28
Share with GSM coverage (2015)	0.71	0.30	0.01	1.00
Share owns a radio (2016)	0.74	0.12	0.38	1.00
Share gets radio news often (2016)	0.76	0.12	0.50	1.00
Avg. N radio stations covering each town (2016)	10.98	7.60	0.00	23.36
B. Individual-level variables (n = 4060)				
Male	0.75	0.43	0.00	1.00
Age	31.73	9.27	18.00	99.00
Completed primary school	0.07	0.26	0.00	1.00
Completed secondary school	0.29	0.46	0.00	1.00
Completed university	0.14	0.34	0.00	1.00

Sources: **District-level variables:** Debate variables from Internews. All 2017 and 2011 variables come from NEC. All 2008 variables come from 2008 Population and Housing Census. 'Share with GSM coverage' comes from Collins Mobile Coverage Explorer. 'Share owns a radio' and 'Share gets radio news often' come from Afrobarometer. 'Avg. N radio stations covering each town' comes from Internews. **Individual-level variables:** All come from researchers' panel survey.

Table 7: Candidate debate participation

	(1)	(2)	(3)
A. Share of candidates			
Invite	0.077**	0.065**	0.092***
	(0.034)	(0.030)	(0.033)
Control Mean	0.542	0.573	0.557
Observations	4060	4060	4060
B. Share of contenders			
Invite	0.201***	0.151***	0.220***
	(0.061)	(0.053)	(0.056)
Control Mean	0.442	0.525	0.480
Observations	4060	4060	4060
C. Incumbent			
Invite	0.212**	0.177**	0.234***
	(0.083)	(0.073)	(0.083)
Control Mean	0.280	0.372	0.299
Observations	4060	4060	4060
D. Share of challengers			
Invite	0.212***	0.144**	0.220***
	(0.074)	(0.063)	(0.067)
Control Mean	0.492	0.554	0.528
Observations	4060	4060	4060
E. Share of other candidates			
Invite	0.003	0.008	0.009
	(0.030)	(0.028)	(0.029)
Control Mean	0.562	0.583	0.575
Observations	3991	3991	3991
Weight	No	1/Obs	Reg/Obs

Outcome variables are the share of the respective set of candidates (all, predicted contenders, incumbent, predicted challenger, other candidate) who attended a debate out of all candidates in that district. Panels A-D have 4060 observations; Panel E has fewer due to only three candidates running in two districts (and hence no 'other candidates' defined).

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the district-level in parentheses.

Table 8: Debate exposure

	(1)	(2)	(3)
A. Standardized index			
Invite	0.245***	0.275***	0.342***
	(0.082)	(0.085)	(0.086)
Control Mean	-0.033	-0.024	-0.038
B. Change in heard debate			
Invite	0.080***	0.080***	0.102***
	(0.029)	(0.026)	(0.027)
Control Mean	0.084	0.082	0.082
C. Heard debate			
Invite	0.038*	0.035*	0.050**
	(0.022)	(0.020)	(0.021)
Control Mean	0.195	0.202	0.193
D. Number of times heard			
Invite	0.085*	0.104**	0.120***
	(0.045)	(0.046)	(0.045)
Control Mean	0.420	0.440	0.420
E. Learned about CSDF			
Invite	0.042*	0.061*	0.050*
	(0.024)	(0.032)	(0.027)
Control Mean	0.243	0.242	0.245
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs

Outcome variables: Panel A: a standardized index of variables in Panels B-D. Panel B: an indicator for whether the respondent had not heard their district debate at baseline but had at endline. Panel C: an indicator for whether the respondent had heard the debate at endline. Panel D: the number of times the respondent had heard the debate at endline. Panel E: standardized index of the change in how many factual questions about CSDF management respondents named between baseline and endline.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the district-level in parentheses.

Table 9: Political engagement

	(1)	(2)	(3)
A. Demand for political information			
Invite	0.169***	0.211***	0.202***
	(0.052)	(0.060)	(0.061)
Control Mean	-0.017	-0.044	-0.022
B. Debate coordination effects			
Invite	0.145**	0.150**	0.183***
	(0.062)	(0.058)	(0.058)
Control Mean	-0.051	-0.040	-0.067
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs

Outcome variables: Panel A: z-score index of variables measuring (1) change in how much respondents listened to the radio (2) change in how much they discussed politics with their friends (3) how much they accessed other sources of political information. Panel B: z-score index of variables measuring (1) how much respondents discussed the debate with friends (2) how much this discussion led to coordinating their vote choices.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the district-level in parentheses.

Table 10: Consequences of increased political engagement

	Turnout			**Whether to vote**			**Who to vote for**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
Invite	0.022***	0.018***	0.023***	0.019**	0.012	0.019**	0.034**	0.036**	0.042**
	(0.006)	(0.006)	(0.006)	(0.008)	(0.008)	(0.009)	(0.016)	(0.016)	(0.016)
Control Mean	0.701	0.699	0.699	0.067	0.076	0.073	0.124	0.128	0.124
Observations	5386	5386	5386	4060	4060	4060	4060	4060	4060
Weight	No	1/PS	Reg	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs

Outcome variables: Columns 1-2: the share of registered voters at the polling station-level who voted in the House of Representatives election, based on NEC data. Columns 3-5: an indicator for whether respondents stated that the debate changed their decision of whether to vote in the election. Columns 6-8: an indicator for whether respondents stated that the debate their decision of who to vote for.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the district-level in parentheses.

Table 11: Changes in certainty about candidates

	Certainty about competence			Certainty about policy		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Contenders						
Invite	0.122	0.134	0.107	0.238**	0.224**	0.196**
	(0.102)	(0.109)	(0.096)	(0.092)	(0.103)	(0.096)
Control Mean	0.284	0.256	0.275	-0.001	-0.021	0.000
Observations	12180	12180	12180	12180	12180	12180
B. Incumbent						
Invite	0.251*	0.263*	0.252**	0.275**	0.316**	0.313**
	(0.148)	(0.151)	(0.118)	(0.121)	(0.129)	(0.121)
Control Mean	0.087	0.039	0.070	-0.052	-0.074	-0.059
Observations	3496	3496	3496	3496	3496	3496
C. Challengers						
Invite	0.057	0.070	0.038	0.216**	0.184	0.152
	(0.102)	(0.113)	(0.107)	(0.095)	(0.114)	(0.104)
Control Mean	0.368	0.343	0.361	0.021	0.001	0.025
Observations	8684	8684	8684	8684	8684	8684
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs

DV in columns 1-3 is the standardized change in certainty respondents express about candidate competence between baseline and endline. DV in columns 4-6 is the standardized change in certainty respondents express about candidate priority issues between baseline and endline.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the district-level in parentheses.

Table 12: Updating about candidate competence and policy priorities

	Beliefs about competence			**Learning about policy**		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Contenders						
Invite	-0.016	-0.076	-0.028	0.059	0.055	0.068
	(0.066)	(0.075)	(0.067)	(0.061)	(0.082)	(0.071)
Control Mean	0.017	0.020	0.012	0.002	-0.029	-0.020
Observations	12180	12180	12180	12180	12180	12180
B. Incumbent						
Invite	0.109	0.074	0.105	0.090	0.129	0.092
	(0.083)	(0.101)	(0.093)	(0.065)	(0.092)	(0.074)
Control Mean	0.070	0.092	0.077	-0.001	-0.028	-0.016
Observations	3496	3496	3496	3496	3496	3496
C. Challengers						
Invite	-0.074	-0.140*	-0.090	0.039	0.028	0.065
	(0.071)	(0.083)	(0.073)	(0.066)	(0.089)	(0.082)
Control Mean	-0.005	-0.009	-0.015	0.004	-0.030	-0.021
Observations	8684	8684	8684	8684	8684	8684
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs

DV in columns 1-3 is the standardized change in citizen evaluations of candidate competence between baseline and endline. DV in columns 4-6 is the standardized change in the share of policy priorities of a given candidate that citizens correctly name between baseline and endline.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the district-level in parentheses.

Table 13: Candidate campaigning

	Ground			Radio		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Contenders						
Invite	-0.031**	-0.038***	-0.036**	0.013	0.029	0.018
	(0.014)	(0.014)	(0.015)	(0.021)	(0.021)	(0.021)
Control Mean	0.021	0.026	0.022	-0.016	-0.032	-0.020
Observations	12168	12168	12168	12180	12180	12180
B. Incumbent						
Invite	-0.027	-0.037	-0.024	0.082**	0.088**	0.092**
	(0.023)	(0.026)	(0.027)	(0.037)	(0.041)	(0.042)
Control Mean	0.108	0.101	0.097	0.120	0.073	0.099
Observations	3492	3492	3492	3496	3496	3496
C. Challengers						
Invite	-0.037*	-0.045**	-0.046**	-0.024	-0.004	-0.017
	(0.019)	(0.017)	(0.019)	(0.027)	(0.026)	(0.027)
Control Mean	-0.016	-0.004	-0.010	-0.074	-0.074	-0.070
Observations	8676	8676	8676	8684	8684	8684
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs

Outcome variable in columns 1-3 is a z-score index of how often candidates (1) visited (2) distributed leaflets (3) bought votes in respondents' communities during campaigning. DV in columns 4-6 is z-score of how often respondents heard candidate on the radio in the two weeks before the election.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the district-level in parentheses.

Table 14: Vote switching

	Main effect			Interaction term: Std. Performance			Interaction term: Std. Policy match		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Contenders									
Invite	-0.002	-0.002	-0.007	-0.002	-0.005	-0.002	-0.002	-0.002	-0.007
	(0.015)	(0.015)	(0.015)	(0.012)	(0.014)	(0.012)	(0.015)	(0.015)	(0.015)
Interaction				0.080***	0.077***	0.081***	-0.009**	-0.005	-0.008*
				(0.010)	(0.012)	(0.011)	(0.004)	(0.005)	(0.004)
Invite × Interaction				0.001	0.001	0.001	0.012*	0.010	0.011*
				(0.012)	(0.014)	(0.012)	(0.006)	(0.006)	(0.006)
Control Mean	0.149	0.151	0.151	0.149	0.151	0.151	0.149	0.151	0.151
Observations	12180	12180	12180	12180	12180	12180	12180	12180	12180
B. Incumbent									
Invite	0.044	0.028	0.037	0.068**	0.044	0.066**	0.041	0.023	0.033
	(0.041)	(0.040)	(0.042)	(0.031)	(0.035)	(0.032)	(0.040)	(0.038)	(0.040)
Interaction				0.022	0.020	0.019	-0.012	-0.015	-0.016
				(0.019)	(0.024)	(0.020)	(0.009)	(0.010)	(0.010)
Invite × Interaction				0.048**	0.055**	0.051**	0.038***	0.050***	0.047***
				(0.024)	(0.026)	(0.024)	(0.014)	(0.015)	(0.015)
Control Mean	0.197	0.201	0.204	0.197	0.201	0.204	0.197	0.201	0.204
Observations	3496	3496	3496	3496	3496	3496	3496	3496	3496
C. Challengers									
Invite	-0.029	-0.011	-0.018	-0.034*	-0.019	-0.021	-0.030	-0.012	-0.019
	(0.022)	(0.021)	(0.023)	(0.018)	(0.018)	(0.018)	(0.022)	(0.021)	(0.023)
Interaction				0.079***	0.069***	0.074***	-0.007	-0.001	-0.005
				(0.012)	(0.015)	(0.013)	(0.005)	(0.006)	(0.006)
Invite × Interaction				-0.028**	-0.019	-0.022	-0.000	-0.006	-0.004
				(0.013)	(0.016)	(0.014)	(0.007)	(0.008)	(0.008)
Control Mean	0.128	0.132	0.128	0.128	0.132	0.128	0.128	0.132	0.128
Observations	8684	8684	8684	8684	8684	8684	8684	8684	8684
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs

Outcome variable is an indicator for whether respondent switched their vote choice towards either a predicted contender (Panel A), the incumbent (Panel B) or a predicted challenger (Panel C) between the baseline and endline surveys. Columns 1-3 show the main effects, Columns 4-6 include interactions of treatment assignment with standardized candidate-level measures of debate performance, and Columns 7-9 include interactions with standardized respondent-candidate-level measures of preference alignment measured at baseline. Section 5.3 explains these interaction terms further.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the district-level in parentheses.

Table 15: Polling station-level voting outcomes

	Main effect			Interaction term: Std. Performance			Interaction term: Std. Policy match		
(Invitation, Rebroadcasting)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Contenders									
Invite	-0.011	-0.011	-0.010	-0.024*	-0.023*	-0.021*	-0.017	-0.017	-0.016
	(0.010)	(0.009)	(0.010)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)
Interaction				0.047***	0.047***	0.050***	-0.037**	-0.034**	-0.039**
				(0.010)	(0.010)	(0.010)	(0.017)	(0.016)	(0.018)
Invite × Interaction				0.014	0.012	0.012	0.012	0.002	0.016
				(0.012)	(0.013)	(0.012)	(0.020)	(0.020)	(0.020)
Control Mean	0.152	0.153	0.153	0.152	0.153	0.153	0.152	0.153	0.153
Observations	16165	16165	16165	16165	16165	16165	16030	16030	16030
B. Incumbent									
Invite	0.012	0.011	0.010	0.022	0.022	0.022	0.019	0.019	0.019
	(0.025)	(0.025)	(0.026)	(0.018)	(0.016)	(0.018)	(0.028)	(0.026)	(0.028)
Interaction				0.036**	0.037**	0.039**	0.018	0.023	0.024
				(0.015)	(0.014)	(0.015)	(0.039)	(0.038)	(0.039)
Invite × Interaction				0.088***	0.088***	0.085***	0.001	0.002	-0.004
				(0.025)	(0.022)	(0.026)	(0.039)	(0.037)	(0.040)
Control Mean	0.245	0.249	0.249	0.245	0.249	0.249	0.245	0.249	0.249
Observations	4762	4762	4762	4762	4762	4762	4627	4627	4627
C. Challengers									
Invite	-0.020	-0.019	-0.018	-0.019	-0.020*	-0.018	-0.026**	-0.023*	-0.025*
	(0.012)	(0.012)	(0.013)	(0.012)	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)
Interaction				0.033***	0.031**	0.034***	-0.046*	-0.039*	-0.051**
				(0.012)	(0.014)	(0.011)	(0.023)	(0.021)	(0.024)
Invite × Interaction				0.019	0.021	0.019	0.006	-0.007	0.013
				(0.018)	(0.020)	(0.017)	(0.021)	(0.020)	(0.021)
Control Mean	0.113	0.111	0.111	0.113	0.111	0.111	0.113	0.111	0.111
Observations	11403	11403	11403	11403	11403	11403	11403	11403	11403
Weight	No	1/PS	Reg	No	1/PS	Reg	No	1/PS	Reg

Outcome variable is an indicator for whether respondent switched their vote choice towards either a predicted contender (Panel A), the incumbent (Panel B) or a predicted challenger (Panel C) between the baseline and endline surveys. Columns 1-2 show the main effects, Columns 3-4 include interactions of treatment assignment with standardized candidate-level measures of performance, and Columns 7-9 include interactions with standardized candidate-level measures of preference alignment with citizen priorities measured at baseline. Section 5.3 explains these interaction terms further.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the district-level in parentheses.

A Appendix

A.1 Classifying candidates as contenders

For each candidate, we constructed an indicator variable for whether the candidate was predicted to be a serious contender. We constructed this indicator as follows, in a sequential fashion until there were three per district: (1) if a candidate was the incumbent; (2) if the candidate ran in the 2011 election and placed 2nd or 3rd; and (3) if the candidate was from a top party. We defined top parties as, sequentially, the incumbent Unity Party (UP), Coalition for Democratic Change (CDC), Liberty Party (LP), the Alternative National Congress (ANC) and the All Liberia Party (ALP). This process resulted in three selected candidates in all districts. These *predicted contenders* are then split into two groups: whether the candidate is the *incumbent* or whether they are a *predicted challenger*, i.e., a non-incumbent predicted contender. The incumbent ran in 64/73 (88%) of races, and so in the remaining 9 districts all three of these candidates are coded as challengers. One additional incumbent ran in a new district and is consequently coded as a challenger. Validating our indicator for top candidates with actual electoral results, we find that in 50% of cases our predicted contenders came in the top three in their district, and in 71% of cases came in the top 5. Given our aim to identify a set of relevant candidates who had plausible chances at electoral success and voters would be interested in, we consider the exercise to be successful.

We show in Table A9 that, using the 'actual' contenders who placed in the top three in the election—whether *actual contenders, winner* or *actual challengers*—generates a similarly strong first stage on debate participation. Using this alternative categorization generates a set of qualitatively similar results, albeit with a more restricted sample of only those candidates who were both *predicted* and *actual* contenders in the citizen-candidate level analysis. However, given the effects we find on voting outcomes, we consider it likely that the definition of *actual contenders* is endogenous to our intervention. These additional results are available on request.

A.2 Balance

We report balance on pre-treatment covariates at the district, individual, polling station and candidate levels. Balance is assessed by estimating Equation 1 for each covariate as an outcome, but omitting the individual-level $\mathbf{X}_i$ and district-level $\mathbf{Z}_d$ as controls. Across the different specifications, we present the coefficient on the treatment indicator $Invite$ to test for evidence of imbalance between treatment groups. For district-level specifications, we report specifications where districts are unweighted and where we weight by the number of registered voters, and use heteroskedasticity-robust standard errors. For individual-level specifications, we restrict to the the endline survey sample and consider two types of outcomes. First, we assign district-level outcomes to individuals in this sample. Second, we use individual-level covariates collected in the survey itself. We use weights as described in Section 6.2 and cluster at the district level. For the polling station-level specifications, we assign district-level outcomes to each polling station in that district. We present an unweighted specification, one where we

weight by the inverse of the number of polling stations in that district and one where we weight by the number of registered voters in that polling station, and again cluster at the district level. Lastly, for the candidate-level specification we present an unweighted specification assessing balance on characteristics drawn from our candidate survey, as well as a weighted specification where we weight by the inverse of the number of responding candidate types in a given district. We refer throughout to imbalance on the unweighted specification since patterns of limited imbalance are generally shared irrespective of weighting schemes.

In Table A2 we report balance at the district level. In the unweighted specification, 0 (0) out of 18 covariates are imbalanced at the 5% (10%) level. In Table A3 we report balance in the endline survey sample when we assign district-level covariates to respondents. In the unweighted specification we find that 2 (2) out of 18 covariates are imbalanced. In Table A4 we report balance in the endline survey sample using individual-level covariates. In the unweighted specification, 1 (1) out of 4 covariates are imbalanced. In Table A5 we report balance at the polling station level. We find imbalance on 0 (1) covariates out of 18 covariates are imbalanced.

In Table A6 we assess evidence of imbalance on incumbent quality by treatment assignment. In the unweighted specification, we find imbalance on 0 (0) covariates out of 3. In Table A8 we test for imbalance at the candidate-level using our survey of candidates who ran in the election. Importantly, in Column 1 we demonstrate balanced response rates to our post-election survey across all candidates, contenders, incumbents and challengers across treatment groups. Using the full sample of candidates, we find imbalance on 1 (2) covariates out of 8. Restricting to predicted contenders we find imbalance on 2 (2) covariates. Restricting to incumbents we find imbalance on 0 (1) covariates. Restricting to predicted challengers we find imbalance on 2 (3) covariates. We consider balance at the candidate-level to be good particularly given our primarily descriptive employment of this data.

Overall we find little evidence of aggregate imbalance—whether on political or non-political variables—and, when applicable, we control for the variables we blocked on throughout the analysis to deal with whatever imbalance that might exist.

A.3 Divergences from Pre-Analysis Plan

This study was pre-registered with EGAP (ID: 20171024AA) and AEA (ID: AEARCTR-0002553) under the title "Turning Up, Tuning In, Turning Out: Experimental Evidence from Liberia." Pre-registration took place before endline data collection and any data analysis. In this section we describe the limited differences between our PAP and the final paper, as well as the logic behind them.

A.3.1 Data and estimation

We reorganized some of categorizations of variables from the PAP to fit into more coherent groupings. This comprised combining 'Knowledge about candidates' and 'Beliefs over candidate competence' into

'Effects on beliefs about candidates;' and 'Voter coordination' and relevant parts of 'Debate exposure' into 'Effects on political engagement'.

As we discuss in the paper, we cross-randomized a separate intervention to vary the intensity of debate rebroadcasting which ultimately had little effect. Since this additional intervention had no effect and candidates were unaware of rebroadcasting plans, we pool over rebroadcasting intensity for clarity of exposition. Importantly, however, we made no multiplicative hypotheses – rather, all our hypotheses were with respect to the overall *intensity* of the debates initiative and focused on those districts assigned to both high invitation intensity and high invite intensity. We can demonstrate that all our key results also hold under this factorial design, but pooling rebroadcasting loses relatively little granularity and gains substantially in clarity. The estimating equation we use in the paper is closest to what we called our 'base specification' in our PAP (Equation 5).

We additionally pre-registered the possibility of constructing an individual-level instrument for the debate attendance of candidates, leveraging random assignment of candidates to debates with the incumbent and at different times of day in districts where more than one debate was held. We found such an instrument to be underpowered due to the number of districts which only ended up holding one debate and so do not report results using it. We also pre-registered a local regression discontinuity design (Equation 8) leveraging quasi-random assignment to respondents being interviewed before or after the live debate in their district at baseline, but lacked sufficient within-district variation to pursue this. Finally, we pre-registered the use of one-tailed tests but report two-tailed tests throughout to be conservative.

We did not pre-register outcomes relating to overall turnout in the election nor the distinction we use between 'on-the-ground' and 'radio' campaigning by candidates. While our pre-registered hypotheses make reference to the distinction between incumbents and challenger candidates, we did not pre-register the differential compliance of incumbents with the invitation intervention that we discuss in Section 8.

A.3.2 Hypotheses

We reorganised and grouped many of our pre-registered hypotheses, which were generally made with reference to individual outcome variables, into more coherent aggregated clusters. Out of the 27 hypotheses we pre-registered, results directly testing 18 of them are presented in the final paper.[45] The nine missing hypotheses fall into three categories. First, we do not report results relating to the hypotheses using within-district variation in whether citizens at baseline were interviewed before or after their district debate had been broadcast for the first time due to the lack of variation mentioned above.

Second, we do not report results for our pre-registered set of hypotheses relating to citizen attitudes towards the media and the electoral process. We anticipated that citizens in districts assigned to more

[45] Broadly we aggregated hypotheses from 'Debate exposure and knowledge about candidates' and 'Beliefs about candidates' into H4; hypotheses from 'Preferences and voting behavior' into H6a/H6b; 'Media consumption, attitudes, and institutions' and 'Debate exposure and knowledge about candidates' into H3.

intensive debates would update positively about the neutrality and contribution of the media to the electoral process due to the novelty and unbiasedness of the debate structure. We found little systematic evidence of this happening – we interpret this to be consistent with the campaigning response of incumbent candidates campaigning more aggressively on the radio in these districts where they were more likely to participate. We report these results for completeness in Table A13. The final paper therefore contains substantially less emphasis on the intervention affecting media credibility than our PAP.

A.4 Variable construction

In this appendix we document the construction of all variables used in the analysis. Unless otherwise noted, these variables come from our panel survey of citizens where we refer to specific items in our baseline and endline survey instruments using the format *wave-question*, where *wave* is represented by *B* (baseline) or *E* (endline) and *question* is simply the question on the relevant instrument. Both survey instruments can be found online at http://egap.org/registration/2899.

As described in Section 6, whenever we asked the same question in both baseline and endline we use the difference as an outcome. We preserve whether variables are discrete or continuous. For indices, we simply sum relevant components and standardize the index to have mean zero and standard deviation of one.

Table 7:

- *Share of candidates*: share of the total candidates in that district who participated in their district debate. Source: debate reports.
- *Share of contenders*: share of the predicted contenders in that district (see Appendix A.1) who participated in their district debate. Source: debate reports.
- *Incumbent*: indicator for whether incumbent participated in their district debate. Source: debate reports.
- *Share of challengers*: share of the predicted challengers in that district (see Appendix A.1) who participated in their district debate. Source: debate reports.
- *Share of other candidates*: share of non-predicted candidates in that district (see Appendix A.1) who participated in their district debate. Source: debate reports.

Table 8:

- *Standardized index*: standardized index of below variables.
- *Change in heard debate*: indicator variable for whether the respondent heard their district debate between baseline (B-Q7) and endline (E-Q14) surveys.
- *Heard debate*: indicator variable for whether the respondent heard their district debate at endline (E-Q14).

- *Number of times heard*: continuous variable for the number of times respondents reported hearing their district debate at endline (E-Q15).
- *Learned about CSDF*: change in the share of correct factual responses about CSDF management that citizens named between baseline and endline. These questions related to identifying the individual overall responsible for CSDF decisions (B-Q9, E-Q9), the requirements to involve citizens in decisions (B-Q10. E-Q10) and the requirements to report allocation decisions (B-Q11, E-Q11).

Table 9:

- *Demand for political information*: standardized index of: respondents' change in listening to radio between baseline (B-Q6, E-Q6); how frequently respondents sought political information from non-radio sources such as newspapers, television and the internet (E-Q7); change in how frequently respondents discussed political issues with friends, family, neighbors and other members of the community (B-Q8, E-Q8).
- *Debate coordination effects*: standardized index of: how frequently respondents discussed the debate content with others (E-Q19); whether this discussion led respondents, along with others, to agree on one particular candidate to vote for (E-Q20).

Table 10:

- *Turnout*: polling station-level turnout, defined as number of votes cast in House of Representatives election divided by the total number of registered voters at that polling station. Source: National Elections Commission.
- *Whether to vote*: whether respondents stated that the debates changed their mind over whether to vote in the election (E-Q18).
- *Who to vote for*: whether respondents stated that the debates changed their mind over who to vote for in the election (E-Q18).

Table 11:

- *Certainty about competence*: standardized change in how sure respondents were about the competence of specific predicted contender candidates between baseline (B-Q22, B-Q24, B-Q26) and endline (E-Q34, E-Q36, E-Q38).
- *Certainty about issues*: standardized change in how sure respondents were about the priority issues of specific predicted contender candidates between baseline (B-Q16, B-Q18, B-Q20) and endline (E-Q28, E-Q30, E-Q32).

Table 12:

- *Beliefs about competence*: standardized change in how competent respondents believe specific predicted contender candidates were between baseline (B-Q21, B-Q23, B-Q25) and endline (E-Q33, E-Q35, E-Q37).
- *Learning about policy*: standardized change in the share of candidate priority issues that citizens name between baseline (B-Q15, B-Q17, B-Q19) and endline (E-27, E-29, E-31). We define candidate priorities using the aggregate of citizen beliefs over a given candidate's priorities measured in the baseline survey.

Table 13:

- *Ground*: standardized index of survey responses to questions about how often specific predicted contender candidates distributed leaflets or posters in their community (E-Q41.1, E-Q41.2, E-Q41.3), made campaign visits to their community (E-Q41.1, E-Q41.2, E-Q41.3) and how frequently other people in their community voted for a given candidate in exchange for money, food or other gifts (E-Q40.1, E-Q40.2, E-Q40.3).
- *Radio*: standardized measure how how frequently respondents heard candidates on the radio in the two weeks before the election (E-Q39.1, E-Q39.3, E-Q39.5).

Table 14:

- *Main effect*: indicator for whether a respondent did not name a specific predicted contender candidate at baseline (B-Q27) but did at endline (E-Q45).
- *Interaction: Performance*: jackknife measure of debate performance of a specific predicted contender, defined as the share of other citizens in the respondent's district who named that candidate as their debate winner (E-Q17).
- *Interaction: Priority match*: measure of preference alignment between respondent and a specific predicted contender candidate. Defined as the share of the three priority issues the respondents name in their districts at baseline (B-Q13) that are shared with the priorities of a given candidate based on aggregating citizen perceptions of that candidates' priorities at baseline (B-Q15, B-Q17, B-Q19).

Table 15:

- *Main effect*: vote share of given candidate at the polling station-level, defined as the number of votes received by candidates divided by number of registered voters at that polling station. Source: National Elections Commission.
- *Interaction: Performance*: district-level analog of jackknife measure, defined as the share of respondents in a given district who name a given candidate (out of all candidates running for office) as their debate winner (E-Q17).
- *Interaction: Performance*: defined as the average proportion of top priorities respondents in a given district (B-Q13) share with a given candidate's top three priority issues (B-Q15, B-Q17, B-Q19).

A.5 Regressions underlying selection results

The test statistics relating to Table 5 come from estimating the following regression at the candidate-level, where we use non-participating incumbents as the excluded category:

$$y_{cd} = \beta_0 + \beta_1 P + \sum_g \beta_2 g + \sum_g \beta_3 (g \times P) + \epsilon_{cd} \tag{A1}$$

Where we regress outcome y for candidate c in district d onto a constant, an indicator P for participating in the debate, a set of indicators g for whether the candidate is either a 'challenger' or 'other candidate', and the interaction of these indicators with debate participation. We cluster standard errors at the district level. We present results in Table A1 which are the basis for the figures cited in the discussion.

Table A1: Candidate selection regressions

	(1)	(2)	(3)
Challenger	0.078*	0.049***	-0.157***
	(0.042)	(0.016)	(0.053)
Other	0.062*		
	(0.034)		
P	0.158***	0.081***	0.148*
	(0.058)	(0.020)	(0.076)
Challenger × P	-0.154**	-0.086***	-0.057
	(0.069)	(0.024)	(0.074)
Other × P	-0.156**		
	(0.060)		
Constant	0.191***	0.401***	3.487***
	(0.032)	(0.016)	(0.053)
Observations	612	219	219
Intercept	0.191	0.401	3.487

All specifications are estimated using OLS using Equation A1. Outcome variables in Column 1 and 2 are priority alignment measures between candidate and citizens in their district as per Section 8. Outcome variable in Column 3 is mean of candidate competence reported by citizens at baseline. Non-participating incumbents are the excluded category. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the district-level in parentheses.

Table A2: District-level balance

	VS of 1st place (2011)		**VS of 1st place (2011)**		**VS of 3rd place (2011)**	
Invite	-0.011	-0.015	-0.005	-0.008	-0.000	-0.001
	(0.024)	(0.025)	(0.010)	(0.010)	(0.007)	(0.006)
Control Mean	0.322	0.309	0.178	0.183	0.129	0.129
	Turnout (2011)		**Share ran in 2011**		**VS HHI (2011)**	
Invite	-0.004	-0.004	-0.012	-0.019	-0.006	-0.009
	(0.009)	(0.008)	(0.024)	(0.024)	(0.020)	(0.020)
Control Mean	0.661	0.667	0.258	0.266	0.198	0.189
	N. candidates (2017)		**Incumbent ran (2017)**		**Number of debates**	
Invite	-0.509	-0.670	0.037	0.016	-0.083	-0.131
	(0.833)	(0.928)	(0.079)	(0.076)	(0.126)	(0.140)
Control Mean	13.634	14.780	0.850	0.883	2.115	2.256
	Log pop. dens. (2008)		**Log reg. voters (2017)**		**GSM coverage (2016)**	
Invite	0.254	0.160	0.056	0.060	0.032	0.005
	(0.365)	(0.370)	(0.061)	(0.055)	(0.062)	(0.055)
Control Mean	-9.847	-9.108	10.214	10.351	0.645	0.773
	Share owns radio (2015)		**Share radio news (2015)**		**Radio coverage (2016)**	
Invite	-0.019	-0.034	-0.017	-0.038	-0.109	0.024
	(0.025)	(0.022)	(0.026)	(0.025)	(1.347)	(1.384)
Control Mean	0.755	0.773	0.767	0.780	10.051	12.830
	Debate week		**Share sec. ed. (2008)**		**Share 18+ (2008)**	
Invite	0.028	-0.030	0.004	0.003	0.000	0.001
	(0.215)	(0.209)	(0.010)	(0.010)	(0.004)	(0.004)
Control Mean	3.980	3.979	0.137	0.154	0.483	0.487
Observations	73	73	73	73	73	73
Weight	None	Reg	None	Reg	None	Reg

Descriptions of all variables can be found in Table 6.

All specifications are estimated using OLS and include block FE. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the district-level in parentheses.

Table A3: Individual-level balance (district covariates)

	VS of 1st place (2011)			**VS of 1st place (2011)**			**VS of 3rd place (2011)**		
Invite	0.007	-0.009	-0.014	-0.011	-0.005	-0.008	-0.000	-0.000	-0.001
	(0.025)	(0.021)	(0.021)	(0.009)	(0.008)	(0.009)	(0.007)	(0.006)	(0.005)
Control Mean	0.298	0.308	0.308	0.185	0.182	0.183	0.129	0.128	0.129
	Turnout (2011)			**Share ran in 2011**			**VS HHI (2011)**		
Invite	0.001	-0.004	-0.004	-0.043**	-0.012	-0.019	0.007	-0.005	-0.007
	(0.007)	(0.007)	(0.007)	(0.019)	(0.021)	(0.020)	(0.021)	(0.017)	(0.017)
Control Mean	0.665	0.659	0.667	0.275	0.270	0.266	0.181	0.189	0.188
	N. candidates (2017)			**Incumbent ran (2017)**			**Number of debates**		
Invite	-0.642	-0.549	-0.710	0.010	0.034	0.014	-0.117	-0.091	-0.139
	(0.757)	(0.709)	(0.795)	(0.068)	(0.068)	(0.066)	(0.108)	(0.108)	(0.120)
Control Mean	15.084	13.963	14.833	0.900	0.858	0.883	2.291	2.152	2.265
	Log pop. dens. (2008)			**Log reg. voters (2017)**			**GSM coverage (2016)**		
Invite	0.207	0.246	0.149	0.072	0.055	0.059	0.029	0.032	0.004
	(0.317)	(0.312)	(0.317)	(0.052)	(0.052)	(0.047)	(0.050)	(0.053)	(0.047)
Control Mean	-8.975	-9.611	-9.096	10.342	10.215	10.354	0.795	0.695	0.774
	Share owns radio (2015)			**Share radio news (2015)**			**Radio coverage (2016)**		
Invite	-0.041**	-0.020	-0.035*	-0.035	-0.017	-0.039*	-0.172	-0.138	-0.010
	(0.020)	(0.022)	(0.019)	(0.025)	(0.022)	(0.022)	(1.222)	(1.159)	(1.190)
Control Mean	0.771	0.750	0.774	0.778	0.767	0.780	13.552	11.275	12.895
	Debate week			**Share sec. ed. (2008)**			**Share 18+ (2008)**		
Invite	-0.035	0.031	-0.026	0.006	0.004	0.002	-0.001	0.000	0.001
	(0.193)	(0.184)	(0.179)	(0.008)	(0.009)	(0.009)	(0.004)	(0.004)	(0.003)
Control Mean	4.162	4.148	3.980	0.160	0.146	0.154	0.488	0.485	0.487
Observations	4061	4061	4061	4061	4061	4061	4061	4061	4061
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs

Descriptions of all variables can be found in Table 6.

All specifications are estimated using OLS and include block FE. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the district-level in parentheses.

Table A4: Individual-level balance (individual covariates)

	Survey date			**Education**		
Invite	-1.063**	-0.796	-1.006*	0.067	0.102	0.081
	(0.523)	(0.572)	(0.573)	(0.086)	(0.084)	(0.075)
Control Mean	71.801	71.153	71.422	6.586	6.447	6.534
	Age			**Male**		
Invite	0.163	0.359	0.225	0.011	0.025	0.021
	(0.439)	(0.464)	(0.437)	(0.015)	(0.016)	(0.016)
Control Mean	31.728	32.103	31.877	0.746	0.744	0.740
Observations	4061	4061	4061	4061	4061	4061
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs

All specifications are estimated using OLS and include block FE. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the district-level in parentheses.

Table A5: Polling station-level balance

	VS of 1st place (2011)			**VS of 1st place (2011)**			**VS of 3rd place (2011)**		
Invite	-0.015	-0.011	-0.015	-0.009	-0.005	-0.009	-0.001	-0.000	-0.001
	(0.021)	(0.021)	(0.021)	(0.009)	(0.008)	(0.009)	(0.005)	(0.006)	(0.005)
Control Mean	0.308	0.309	0.309	0.183	0.184	0.184	0.129	0.129	0.129
	Turnout (2011)			**Share ran in 2011**			**VS HHI (2011)**		
Invite	-0.004	-0.004	-0.004	-0.016	-0.012	-0.017	-0.009	-0.006	-0.008
	(0.007)	(0.007)	(0.007)	(0.021)	(0.021)	(0.020)	(0.017)	(0.017)	(0.018)
Control Mean	0.665	0.667	0.667	0.265	0.266	0.266	0.188	0.189	0.189
	N. candidates (2017)			**Incumbent ran (2017)**			**Number of debates**		
Invite	-0.652	-0.509	-0.681	-0.050	-0.019	-0.052	-0.125	-0.083	-0.131
	(0.771)	(0.716)	(0.798)	(0.067)	(0.070)	(0.066)	(0.118)	(0.109)	(0.121)
Control Mean	14.659	14.802	14.802	0.883	0.890	0.890	2.240	2.257	2.257
	Log pop. dens. (2008)			**Log reg. voters (2017)**			**GSM coverage (2016)**		
Invite	0.177	0.254	0.154	0.062	0.056	0.060	0.010	0.032	0.004
	(0.322)	(0.314)	(0.320)	(0.048)	(0.052)	(0.047)	(0.048)	(0.053)	(0.047)
Control Mean	-9.224	-9.098	-9.098	10.325	10.352	10.352	0.753	0.775	0.775
	Share owns radio (2015)			**Share radio news (2015)**			**Radio coverage (2016)**		
Invite	-0.031	-0.019	-0.033*	-0.037*	-0.017	-0.038*	-0.011	-0.109	-0.076
	(0.019)	(0.022)	(0.019)	(0.021)	(0.022)	(0.022)	(1.202)	(1.158)	(1.201)
Control Mean	0.770	0.773	0.773	0.778	0.780	0.780	12.496	12.931	12.931
	Debate week			**Share sec. ed. (2008)**			**Share 18+ (2008)**		
Invite	-0.042	0.028	-0.048	0.002	0.004	0.002	0.002	0.000	0.001
	(0.176)	(0.185)	(0.178)	(0.009)	(0.009)	(0.009)	(0.003)	(0.004)	(0.003)
Control Mean	4.009	3.988	3.988	0.152	0.154	0.154	0.486	0.487	0.487
Observations	5386	5386	5386	5386	5386	5386	5386	5386	5386
Weight	No	1/PS	Reg	No	1/PS	Reg	No	1/PS	Reg

Descriptions of all variables can be found in Table 6.

All specifications are estimated using OLS and include block FE. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the district-level in parentheses.

Table A6: Incumbent balance

	Attendance		**Absent**		**Distant**	
(Invitation, Rebroadcasting)	(1)	(2)	(3)	(4)	(5)	(6)
Invite	0.034	0.019	-0.003	0.013	-0.025	-0.023
	(0.032)	(0.030)	(0.020)	(0.024)	(0.028)	(0.024)
Mean	0.791	0.807	0.117	0.114	0.079	0.066
Observations	73	73	73	73	73	73
Weight	None	Reg	None	Reg	None	Reg

Outcome variables are plenary session attendance measures taken from legislator scorecards for 2016. Legislators either attend, are absent, or are away from Monrovia for each plenary session.

All specifications are estimated using OLS and include block FE. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the district-level in parentheses.

Table A7: Candidate-level balance

	Response (1)	Age (2)	Univ. ed. (3)	Radio (4)	Male (5)	Ran before (6)	Gov job (7)	Advocacy (8)	NGO (9)
A. All candidates									
Invite	-0.028	-0.379	0.012	0.020	0.009	-0.024	-0.028	0.064**	0.065*
	(0.031)	(1.064)	(0.049)	(0.013)	(0.027)	(0.032)	(0.030)	(0.025)	(0.036)
Control Mean	0.63	48.48	0.55	0.03	0.85	0.30	0.34	0.83	0.38
Observations	984	608	612	612	612	612	612	612	612
B. Contenders									
Invite	0.034	3.063**	-0.004	0.003	0.074	0.103	-0.059	0.076	0.184**
	(0.072)	(1.494)	(0.086)	(0.043)	(0.061)	(0.075)	(0.072)	(0.046)	(0.083)
Control Mean	0.57	49.02	0.66	0.08	0.87	0.49	0.39	0.84	0.28
Observations	219	128	131	131	131	131	131	131	131
C. Incumbents									
Invite	0.139	2.394	0.061	-0.030	0.121	0.000	-0.394*	-0.121	0.061
	(0.131)	(4.085)	(0.210)	(0.200)	(0.174)	(.)	(0.212)	(0.174)	(0.255)
Control Mean	0.42	55.31	0.69	0.15	0.77	1.00	0.69	0.92	0.23
Observations	64	31	31	31	31	31	31	31	31
D. Challengers									
Invite	0.010	3.875**	-0.005	-0.002	0.027	0.134	0.036	0.155**	0.175*
	(0.083)	(1.655)	(0.106)	(0.052)	(0.067)	(0.102)	(0.086)	(0.065)	(0.103)
Control Mean	0.64	47.31	0.65	0.06	0.90	0.35	0.31	0.81	0.29
Observations	155	99	102	102	102	102	102	102	102

Panel A presents balance tests for the full set of candidates in survey, Panel B restricts to predicted contender candidates, Panel C restricts to incumbent candidates, Panel D restricts to predicted challenger candidates. Outcome variables are: response rate to survey; age in years; indicator for whether candidate completed university; indicator for whether they own or manage a radio station; indicator for candidate being male; indicator for candidate having run for office before; indicator for candidate having a government job before; indicator for candidate having advocacy experience; indicator for candidate working for an NGO before.
All specifications are estimated using OLS and include block FE. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the district-level in parentheses.

Table A8: Candidate-level balance (weighted)

	Response (1)	Age (2)	Univ ed (3)	Radio (4)	Male (5)	Ran before (6)	Gov job (7)	Advocacy (8)	NGO (9)
A. All candidates									
Invite	-0.045	-0.171	-0.013	0.030*	0.025	-0.031	-0.018	0.043	0.057
	(0.033)	(0.899)	(0.056)	(0.017)	(0.031)	(0.036)	(0.039)	(0.027)	(0.037)
Mean	0.63	48.48	0.55	0.03	0.85	0.30	0.34	0.83	0.38
Observations	984	608	612	612	612	612	612	612	612
B. Contenders									
Invite	0.011	3.554**	0.037	-0.009	0.076	0.126	-0.051	0.062	0.170*
	(0.058)	(1.690)	(0.090)	(0.051)	(0.060)	(0.086)	(0.081)	(0.051)	(0.087)
Mean	0.67	49.02	0.66	0.08	0.87	0.49	0.39	0.84	0.28
Observations	219	128	131	131	131	131	131	131	131
C. Challengers									
Invite	-0.037	4.918***	0.042	0.008	0.008	0.153	0.054	0.144**	0.165
	(0.064)	(1.815)	(0.105)	(0.064)	(0.068)	(0.106)	(0.094)	(0.068)	(0.101)
Mean	0.78	47.31	0.65	0.06	0.90	0.35	0.31	0.81	0.29
Observations	155	99	102	102	102	102	102	102	102

Panel A presents balance tests for the full set of candidates in survey, Panel B restricts to predicted contender candidates, Panel C restricts to incumbent candidates, Panel D restricts to predicted challenger candidates. Outcome variables are: response rate to survey; age in years; indicator for whether candidate completed university; indicator for whether they own or manage a radio station; indicator for candidate being male; indicator for candidate having run for office before; indicator for candidate having a government job before; indicator for candidate having advocacy experience; indicator for candidate working for an NGO before.
All specifications are estimated using OLS and include block FE. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the district-level in parentheses.

Table A9: First stage: Invitation (supplementary)

	(1)	(2)	(3)
A. Share of actual contenders			
Invite	0.220***	0.162***	0.220***
	(0.059)	(0.053)	(0.057)
Control Mean	0.431	0.515	0.458
B. Election winner			
Invite	0.253**	0.202**	0.275***
	(0.097)	(0.089)	(0.093)
Control Mean	0.501	0.520	0.474
C. Share of actual challengers			
Invite	0.267***	0.176**	0.237***
	(0.068)	(0.068)	(0.068)
Control Mean	0.488	0.572	0.525
Observations	4060	4060	4060
D. Share of actual other candidates			
Invite	0.018	0.029	0.036
	(0.039)	(0.037)	(0.038)
Control Mean	0.563	0.584	0.584
Observations	3991	3991	3991
Weight	No	1/Obs	Reg/Obs

Outcome variables are the share of the respective set of candidates (actual contenders, winner, actual challenger) who attended a debate out of all candidates in that district. Actual contenders are defined as candidates who ranked in the top three in their race in the election. Actual other candidates are those who did not rank in the top three. Panels A-C have 4060 observations; Panel D has fewer due to only three candidates running in two districts (and hence no 'actual other candidates' defined).

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the district-level in parentheses.

Table A10: Debate exposure (supplementary)

	(1)	(2)	(3)
A. Debate winner attended debate			
Invite	0.075**	0.078***	0.096***
	(0.031)	(0.029)	(0.029)
Control Mean	0.291	0.297	0.283
B. Accuracy of N participants			
Invite	-0.103***	-0.107***	-0.109***
	(0.035)	(0.029)	(0.035)
Control Mean	0.311	0.315	0.309
C. Change in attended debate			
Invite	-0.055*	-0.071*	-0.058**
	(0.028)	(0.037)	(0.029)
Control Mean	0.129	0.129	0.123
D. Attended debate			
Invite	0.010	0.015	0.019
	(0.010)	(0.012)	(0.011)
Control Mean	0.151	0.153	0.143
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs

Outcome variables: Panel A: an indicator for whether the respondents' stated debate winner actually attended the debate. Panel B: the deviation of how many participants the respondents report in the debates compared to actual number (negative coefficient means more accurate). Panel C: change in whether respondent reports attending debate between baseline and endline. Panel D: indicator for whether respondent reports attending debate at endline.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the district-level in parentheses.

Table A11: Debate winner

	(1)	(2)	(3)
A. Debate winner is contender			
Invite	0.016	0.016	0.021
	(0.022)	(0.021)	(0.022)
Control Mean	0.222	0.208	0.211
B. Debate winner is incumbent			
Invite	0.013	0.013	0.020
	(0.016)	(0.018)	(0.017)
Control Mean	0.112	0.096	0.097
C. Debate winner is challenger			
Invite	0.003	0.005	0.001
	(0.020)	(0.019)	(0.019)
Control Mean	0.110	0.112	0.114
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs

Outcome variables are indicators for whether respondent reports their debate winner as a predicted contender, incumbent or predicted challenger.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the district-level in parentheses.

Table A12: Vote choice reason

	Campaign promises			**Expectations**		
	(1)	(2)	(3)	(4)	(5)	(6)
Invite	0.006	0.002	0.000	0.018	0.005	0.024
	(0.010)	(0.012)	(0.012)	(0.019)	(0.021)	(0.018)
Control Mean	0.033	0.036	0.035	0.067	0.064	0.063
	Competence			**Experience**		
	(1)	(2)	(3)	(4)	(5)	(6)
Invite	0.013	0.016	0.001	-0.024	-0.001	-0.008
	(0.030)	(0.032)	(0.033)	(0.026)	(0.030)	(0.028)
Control Mean	0.196	0.209	0.201	0.262	0.254	0.258
Observations	4060	4060	4060	4060	4060	4060
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs

Outcome variable is an indicator for whether respondents switched towards citing candidate campaign promises, expectated policy by the candidate, candidate competence or candidate experience as their main reason for their vote choice.
All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the district-level in parentheses.

Table A13: Broader consequences

	(1)	(2)	(3)
A. Media: trust and bias			
Invite	-0.007	-0.003	-0.004
	(0.013)	(0.013)	(0.013)
Control Mean	0.003	-0.001	-0.005
B. Media: helps democracy			
Invite	-0.003	-0.004	-0.015
	(0.020)	(0.026)	(0.025)
Control Mean	0.021	0.018	0.023
C. Electoral attitudes			
Invite	0.010	0.013	0.004
	(0.026)	(0.031)	(0.029)
Control Mean	-0.018	-0.015	-0.019
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs

Outcome variables are all z-score indices. Panel A: extent to which the media (1) was unbiased during election (2) gave equal coverage of candidates (3) is trustworthy. Panel B: media (1) helps select competent representatives (2) ensures representatives reflect views of voters. Panel C: elections (1) help select competent representatives (2) ensure representatives reflect views of voters.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the district-level in parentheses.