

# Who Debates, Who Wins? At-Scale Experimental Evidence on Debate Participation in a Liberian Election<sup>1</sup>

Jeremy Bowles<sup>2</sup> & Horacio Larreguy<sup>3</sup>

December 12, 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. By encouraging debate participation more intensely in some districts than in others, we shock the supply of programmatic information from leading candidates. We find substantively large effects on citizen learning, political engagement, and voter coordination concentrated in treated districts, where higher shares of leading candidates were induced to participate. In those districts, challengers decreased their on-the-ground campaigning efforts while incumbents increased their radio campaigning. The intervention electorally benefited incumbents, particularly those predicted to perform well in the debates and those whose policy priorities were more closely aligned with their constituents. Complying incumbents were more likely than their challengers to self-select into debate participation based on the quality of their policy platforms. The results point to the importance of understanding selection into the supply of programmatic information when evaluating the effects of its provision.

**JEL codes:** D72, O12

**Keywords:** accountability, information, selection

---

<sup>1</sup>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, Cesi Cruz, Patrick Francois, Thomas Fujiwara, Miriam Golden, Torben Iversen, Eric Kramon, Julien Labonne, Evan Lieberman, John Marshall, Vincent Pons, Pablo Querubin, Pia Raffler, Francesco Trebbi and Ariel White for useful discussions, and audiences at APSA, the Harvard Experiments Conference, the Harvard Political Economy of Development Working Group, Internews, Liberia Strategic Analysis, the NYU Center for Experimental Social Science, Toulouse School of Economics, Twaweza, USAID Liberia, University of British Columbia, Universitat Autònoma de Barcelona, Universitat Pompeu Fabra, the Wallis Institute Annual Conference 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).

<sup>2</sup>PhD Candidate, Department of Government, Harvard University. jbowles@g.harvard.edu.

<sup>3</sup>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 their actions 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.<sup>4</sup> However, this work often ignores the strategic calculus underlying the provision of programmatic information by candidates, who in clientelistic democracies can face strong incentives to reduce the amount and dissemination 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 their policy priorities by forcing them to commit to policy promises that affect sizable shares of the electorate (and are thus more difficult to renege) and then reporting those promises (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) and clientelism is engrained (Baland and Robinson, 2008; Robinson and Verdier, 2013; Anderson et al., 2015; Cruz et al., 2018b), democratic accountability suffers (Pande, 2011). One particularly effective way to enhance 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 their 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 impact of programmatic information is conditioned by which candidates have selected into supplying it.

In this paper we assess the electoral consequences of candidate selection into the supply of widely-disseminated programmatic information in the setting of Liberia, where clientelism is pervasive and the media sector is weak. We partnered with USAID and the NGO Internews to study the impact of randomized elements of a nationwide initiative to hold debates for all

---

<sup>4</sup>Ferraz and Finan (2008), Chang et al. (2010) and Larreguy et al. (2017) show that media reports on incumbent performance enhance electoral accountability. However, a series of localized information dissemination campaigns fail to consistently replicate these media effects (Banerjee et al., 2011; de Figueiredo et al., 2013; Chong et al., 2015; Bhandari et al., 2016; Dunning et al., 2019). 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).

73 House of Representatives seats ahead of the Liberian election of October 2017.<sup>5</sup> Beyond lawmaking, House members in Liberia control access to development funds as well as play key roles in the allocation and implementation of public goods, and thus voters have incentives to care about the policy priorities of the candidates and vote accordingly. However, historically, votes have been bought as often as won (Bowles et al., 2017). In an effort to improve democratic accountability, Internews organized 129 standardized debates, with at least one in each electoral district, to solicit the policy promises of the participating candidates. In the debates, the 59% of candidates who participated were asked a series of questions by moderating journalists on particular issues of local policy relevance, most often relating to district schools, primary healthcare facilities, and infrastructural investments. Rather than large townhall-style debates, the emphasis was on soliciting concrete policy platforms and promises from the candidates that would then be rebroadcast by community radio stations.

To shock the supply of policy promises, we randomly varied the intensity of invitation efforts to persuade candidates to participate in the debates. The decision to participate is risky, particularly in clientelistic settings where the returns to programmatic competition can be both limited and highly uncertain. Candidates who ‘win’ a debate may enjoy greater publicity and net electoral gains, but *ex ante* they risk performing poorly, revealing their policy priorities to be disconnected from their constituents and restricting their ability to target campaign promises to small groups of influential voters. These risks are especially pronounced for the leading candidates (incumbents and their challengers), who enjoy greater resources for campaigning, are adapted to the existing clientelistic equilibrium, and are more likely to be attacked by opponents to gain publicity.<sup>6</sup> Thus our intervention exerted substantial effort to mitigate leading candidates’ concerns regarding personal attacks and biased moderators and to highlight that the emphasis of the debates would be narrowly focused on policy proposals. By inducing leading candidates to participate in the debates, we intended to generate variation in the share of individuals exposed to their policy promises.

To analyze the effects of the intervention, we make use of a rich set of original data sources, including a nationwide panel survey of over 4,000 citizens, a survey of over 600 candidates who

---

<sup>5</sup>To our knowledge, this is the first time debates have been held for every office in an election in West Africa (Olukotun and Omotoso, 2017).

<sup>6</sup>Prominent examples abound of incumbents avoiding electoral debates, including President Jimmy Carter 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, “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.”

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 debate initiative and invitation intervention were successfully delivered. All debates were conducted, the invitation intervention generated substantively large treatment effects on the debate participation decision of predicted leading candidates, and radio stations broadcast the debates as contracted.

Around 25% of citizens in our survey heard their district debate via radio broadcast. Notably, citizens in treated districts were more exposed to debate content than those in control districts on both the extensive and intensive margin. As our qualitative accounts corroborate, citizens were more interested in, and responsive to, debate information when leading candidates were involved. This exposure to information about relevant candidates led to increased political engagement: in treated districts, citizens engaged in more political information acquisition, discussion, and coordination with others than in control districts. Increased exposure to the debates led citizens to reduce their uncertainty and update their beliefs about the competence and policy priorities of the leading candidates 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 debate when they participated.

In addition, we find that candidates reacted to the intervention in heterogeneous ways. Aided by an increase in demand from radio stations, incumbents increased their radio campaigning; challengers, meanwhile, reduced their on-the-ground campaigning. This suggests that incumbent debate participation had a deterrent effect on challengers' campaign efforts, as they performed worse in the debates than did the participating incumbents.

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, according to both self-reported and polling-station voting outcomes. Remarkably, 50% of incumbents in treated districts won re-election compared to 43% in control districts. These changes in electoral outcomes occurred particularly in districts where incumbents were predicted to perform well in the debates and had policy platforms that were better aligned with the priorities of their constituents. The results, therefore, suggest that shocking the supply of widely-disseminated policy promises by candidates led to broad increases in citizen exposure, engagement, and learning, but uneven electoral consequences that benefited incumbent candidates.

Our results point to the crucial importance of understanding candidate selection into the debates themselves and more generally the strategic decision by candidates to supply pro-

grammatic information. We show that complying incumbents were more likely to self-select into debate participation based on the proximity of their policy preferences to those of their constituents and the uncertainty around them. In other words, incumbents who complied with the intervention were those about whom citizens were initially most uncertain but for whom attending the debates to showcase their policy priorities made the most electoral sense. We find much weaker evidence of self-selection based on initial citizen perceptions of candidate competence. Complier challenger candidates, about whom citizens were also very uncertain, often lacked the political sophistication to correctly assess the returns to debate participation and on average experienced no gain from showing up. Importantly, this pattern of compliance mirrors the self-selection of candidates into debate participation in control districts. This suggests that the treatment effects of the invitation intervention are likely representative of those of the overall debate initiative.

This paper contributes to the expansive literature on information and accountability in democracy in several ways. Most importantly, our approach contrasts with previous debate interventions in that we allow debate participation to vary.<sup>7</sup> Previous studies ensured that all leading candidates were present in the debates in the selected constituencies where they were conducted (Bidwell et al., 2016; Platas and Raffler, 2017; Brierley et al., 2018). We show that candidate attendance cannot be taken for granted when scaling such initiatives at the national level—districts that were not assigned to the intensive-invitation intervention saw only 20% of the incumbents and 40% of the predicted challengers participating in the debate—and our results indicate that the effect of such initiatives is conditioned by the attendance decisions of those candidates. This difference in participation likely explains our finding that the intervention helped incumbent candidates induced to participate. All previous debate interventions 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 et al., 2017; Muralidharan and Niehaus, 2017). In particular, Al-Ubaydli et al. (2017) point to individual non-adoption of the treatment as a key factor explaining the difference in outcomes between small-scale and large-scale interventions. This lack of compliance is more of a problem for interventions subject to political economy considerations. While development

---

<sup>7</sup>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., 2019).

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 clientelistic equilibrium from which they often benefit. Our results then highlight the importance of selection into programmatic initiatives to understand their effects when scaled.

Our intervention also differs from previous work in its mode of delivery. We focus on debates that are broadcast 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 make voters aware that many others have also received a given piece of information, which in turn 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).<sup>8</sup> While we find sizable effects on voter discussion and coordination, Bidwell et al. (2016) find much weaker effects when debates were communicated via other methods, suggesting 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 undermines 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 debate screening,<sup>9</sup> our negative results for on-the-ground campaigning by challengers and positive results on radio campaigning 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 update from both valence and ideological campaign messages and vote accordingly. Closer to the context we study, Cruz et al. (2018a) show that, in the

---

<sup>8</sup>For example, see the mixed evidence from the plethora of studies of 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).

<sup>9</sup>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.

clientelistic democracy of the Philippines, voters who receive comparable information about campaign promises made by all candidates via leaflets 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. (2018a) is successful at clearly conveying comparable candidate campaign promises to voters to see if they update and vote accordingly, such is at the cost of departing from real campaigns where voters have to parse and compare campaign promises delivered by each candidate, as in the debates we study. In addition, our intervention is designed to capture the extent to which the results of information dissemination campaigns are conditioned by the decisions of candidates to supply that information, whereas Cruz et al. (2018a) chose not to vary the participation of candidates in providing information in order to isolate citizen updating and the drivers of their voting decisions.

Our results also contribute to recent work on political selection (Ferraz and Finan, 2011; Gagliarducci and Nannicini, 2013; Dal Bó et al., 2017; Gulzar and Khan, 2018). Although the empirical literature in this field is in its infancy, as noted by Dal Bó and Finan (2018), existing evidence suggests that voters broadly reward valence. Our results provide additional support for this conclusion, finding that in this context, citizens rewarded the candidates whose policy priorities best matched theirs rather than broad-based conceptions of competence.

Lastly, our methods and results also contribute to recent experimental work designed to reduce clientelistic campaign 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 Blattman et al. (2018), however, show that interventions designed to combat clientelistic and vote-buying practices as well as to facilitate competition on the basis of policy vary in their success. Our results underscore the existence of candidate-level variation in the suitability for programmatic competition. This suggests that recruiting citizens to run for office and training them in the articulation of policy platforms could be overlooked factors in shifting towards more programmatic political equilibria.

The paper proceeds as follows. In Section 2 we outline Liberia's political context, legislative candidacy, and electoral campaigning. In Section 3 we describe the debate initiative and our randomized intervention before providing descriptive evidence on the debates. Section 4 reports our hypotheses, and Section 5 describes the set of original data sources we draw on to evaluate them. Section 6 reports our estimation strategy. Section 7 presents results, which

we discuss in Section 8 when we analyze candidate selection into debate participation. We 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, in which each of 73 electoral districts elected a single representative for a six-year term in a first-past-the-post electoral system. Key responsibilities of House members include making laws, controlling access to development funds, and allocating and implementing public goods. 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).<sup>10</sup> The relatively low barriers to candidacy combined with a fragmented party environment mean that there are usually a large number of candidates for each office, making effective candidate selection particularly challenging. In the 2017 election there were 984 candidates for House of Representatives from 26 different political parties, with between 3 and 28 candidates per district.<sup>11</sup> Incumbents sought re-election in nearly 90% of the districts. As one editorial opined, “Rest assured that this is not a healthy expression of diverse opinions. Everyone wants a piece of the pie.” (Glencorse and Yealoe, 2017).

Once in office, legislator performance is often poor. On average, legislators attend slightly fewer than 80% of legislative sessions, but attendance is highly heterogeneous, with some lawmakers having close to perfect attendance while others attend as few as 45% of sessions (IREDD, 2016). Dissatisfaction with incumbent performance is widespread, with 67% of citizens mistrusting their representative, 38% believing that representatives never listen to what citizens say, and 68% disapproving of their incumbent’s performance (Afrobarometer, 2015). This dissatisfaction is not the result of citizens being unaware of who their legislator is—a remarkable 92% of our citizen survey correctly named their legislator—but rather the lack of mechanisms to hold politicians accountable for poor performance. Largely as a result of an

---

<sup>10</sup>In addition, representatives are often gatekeepers 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).

<sup>11</sup>Five candidates dropped out during the campaign season, meaning 979 candidates ended up on the ballot.

underdeveloped media sector, there is a paucity of credible information about political activity. Moreover, access to politicians is difficult due to both poor infrastructural conditions and a limited telecommunications network. For example, more than half of citizens report seeing their representatives only at election time, once every six years (USAID, 2018).

## 2.2 Who runs for office?

We draw on an original survey of 612 candidates to provide some descriptive evidence of candidate characteristics.<sup>12</sup> Throughout this paper we distinguish between three *predicted leading candidates* per district and *other candidates* for theoretical, practical, and measurement reasons. We further divide *predicted leading candidates* into *incumbents* and *predicted challengers*. The objective with this categorization 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 leading candidates*, those whose vote share ranked in the top three of their district, might be endogenous to our intervention. This assignment of 219 *predicted leading candidates*—64 *incumbents* and 155 *predicted challengers*—and 765 *other candidates* is described fully in Appendix A.1.

### I. Candidate Characteristics

	Age (1)	University educated (2)	Ran before (3)	Govt. job before (4)	NGO job before (5)	Advocacy experience (6)	Campaign expenditure (7)	Radio station (8)
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 I we provide descriptive statistics by candidate category. Candidates generally come from Liberia’s elite, with more education than an average citizen, and are overwhelmingly male (84%). Incumbents are 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 are less likely to have experience working for an NGO. Almost a third of all candidates

<sup>12</sup>The response rate is balanced across treatment groups both overall and within candidate categories.

have previously run for office, and nearly all report 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. The differences between incumbents and challengers 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 rallies in communities across each district at which candidates distribute gifts in cash or kind to generate support, while locally targeting generally non-credible policy promises. These promises usually involve building local schools and primary healthcare facilities, among other infrastructural investments, 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 campaign season, incumbents especially orchestrate mass turnout-buying and trucking of voters from the capital to their districts (Bowles et al., 2017). A USAID survey in 2015 indicated that 35% of respondents were personally given money in exchange for their vote while 49% of citizens believed that “many” or “almost everyone” accepted gifts from parties in exchange for their vote (USAID, 2015).

In this clientelistic context, candidates face few incentives to widely disseminate policy promises. Candidates themselves acknowledge differences in the *types* and *credibility* of policy promises delivered at local rallies versus over the radio, as Table II shows. Overall, candidates believe that promises made on the radio are more credible than those made at rallies, but they acknowledge the low likelihood of any campaign promise being kept. Interestingly, incumbents appear to be more sophisticated in this regard.

Because candidates lack incentives to publicize policy promises, the wide dissemination of policy platforms across districts is extremely rare.<sup>13</sup> As such, candidate campaigns broadly lack policy platforms, focusing instead on promises of local development (which rarely come to fruition) delivered through on-the-ground campaigning and vote buying. The absence of programmatic information is facilitated by a fractured media landscape. Radio stations are a

---

<sup>13</sup>One of the country’s most prominent newspapers, the *Daily Observer*, built a “promises tracker” ahead of the election where candidates could outline their policy platforms. No incumbent did this.

## II. Candidate Attitudes Towards Policy Promises

Candidate type	Different promises (1)	Rally credibility (2)	Radio credibility (3)
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 very likely to be fulfilled; ‘Radio credibility:’ candidate believes that promises made on radio are very likely to be fulfilled.

potentially important source of access to political 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 that has more than 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 I 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 debate initiative in the run-up to the October 2017 elections for House of Representatives to push back against Liberia’s clientelistic equilibrium and towards the beginnings of a programmatic one.<sup>14</sup> Internews partnered with several Liberian journalist associations to organize debates across different parts of the country.<sup>15</sup> 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 large number of candidates, multiple debates were held, generally on the same day, with candidates randomly assigned to a debate. The debates took place from mid-August to mid-September 2017. In total, 129 debates were held across all 73 districts.

<sup>14</sup>Other organizations held debates for the presidential race; those debates are not the focus here.

<sup>15</sup>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)



Debate venues were mostly administrative buildings, town halls, and schools. Every debate followed a simple and uniform structure. First, candidates were asked to outline their campaign promises. The moderator then posed the same questions to each candidate in turn, and each candidate was allowed three minutes to respond. 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 two or three questions based on research conducted by the moderator in the district about relevant local issues. Moderators intervened to prevent candidates from making personal attacks on other candidates.

Internews also organized the dissemination of the debate content to the electorate by partnering with at least one community radio station in each district that would broadcast and later rebroadcast the debate.<sup>16</sup> Internews chose 43 radio stations to rebroadcast the debates, making selections based on the signal strength of the station to maximize audience size and discounting any stations that were managed or owned by candidates running for office in that district.<sup>17</sup> Using geographical data on signal coverage and the geocoded 2008 census, we

<sup>16</sup>The audience in attendance at each debate was around 100 people. Election-related violence is a concern in Liberia and so, to minimize the risk of conflict, in-person audiences for the debates were purposefully kept small.

<sup>17</sup>The debates were broadcast by fewer than 73 radio stations since some had the ability to broadcast debates in more than one district. A few stations were discounted due to political affiliations, primarily because they could not be relied upon to replay the debates in full with no editing.

estimate that nearly 90% of the population was covered by a signal from the station broadcasting that district's debate.

To evaluate the impact of the supply of programmatic information by leading candidates, we randomized debate invitation effort at the district level.<sup>18</sup> We carefully designed the intervention 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 invitation activities already planned.

We also cross-randomized the extent of debate rebroadcasting, which ultimately had no effect because citizens were likely to hear their district debate even in districts without intensive rebroadcasting, likely reflecting the level of citizen interest in the debates. Because candidates were unaware of any differences in future rebroadcasting efforts,<sup>19</sup> we present results where we pool over rebroadcasting intensity for clarity of exposition.

### **3.2 Invitation intervention**

Whether to participate in a debate represents a strategic decision by candidates as to whether disseminating programmatic policy information will help or hurt their electoral bid. Candidates faced deep uncertainty regarding the returns to participation, particularly given the novelty of the initiative. In previous election cycles, policy debates were rare; instead, town hall meetings involving personal attacks, biased audiences, and candidate appeals to charisma were the norm. Candidates were frequently concerned that their character and competence would be attacked and moderators would be biased. Furthermore, poor performance in the debate might hurt their candidacy, while committing to a policy platform would restrict their ability to deviate on the campaign trail. Participation also involved non-trivial direct costs given the difficulties of travel in Liberia's rainy season, when many candidates resided in the capital.

Our invitation intervention was intended to target candidates' uncertainty about the costs and benefits of participating in the debate. By varying the intensity of our invitation efforts, we sought to introduce random variation in the level of debate participation across districts. Candidates in control districts were contacted by the relevant Liberian journalist association who invited them to attend and provided logistical information about the debate.

---

<sup>18</sup>Randomizing invitation effort at the candidate level, while cleaner experimentally, would have raised serious concerns in terms of ethics and fairness to candidates.

<sup>19</sup>According to our candidate survey, candidates believed that debates would be rebroadcast roughly two times, with no statistical difference by treatment assignment.

In treatment districts, this was augmented with three additional components aimed at persuading leading candidates to attend. First, we sent official invitations via email from Internews with USAID branding as far ahead of the debates as possible with logistical details and instructions for candidates to contact Internews if they had any doubts about the debates.<sup>20</sup> 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. These calls were designed to address any concerns candidates had about the debates and to clarify the objective, structure, and unbiasedness of the debates. Candidates were informed that questions would be asked about relevant local policy issues but were not provided specific questions. Third, we sent SMS reminders to all candidates on either the evening before or the morning of the debate with reminders and information on where to go. Thus, the email invitations emphasized the credibility of the debates and provided information about the structure of the debates, the phone calls served to further reduce uncertainty, and the text messages served to remind candidates about the debates.

While the intervention was at the district level, we did not expect homogeneity in the response of candidates within a district to the treatment. In particular, as highlighted by our pre-analysis plan, we expected that the treatment was more likely to affect the participation decision of leading candidates. These candidates faced greater potential electoral costs from poor debate performance and, given their greater resources, higher campaign opportunity costs to participation. Among these candidates, we expected that the intervention would particularly affect the participation of those who were most uncertain about the return to debate participation. 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.

### **3.3 Debate participation**

While the debates were well-attended overall, there was substantial variation across districts, ranging from 11% and 100%. Overall 59% of candidates participated, which broke down to 48% of incumbents compared to 60% of challengers and other candidates. As shown in Table III, the reasons cited for participation by our candidate survey respondents varied across

---

<sup>20</sup>USAID was the donor funding the debate 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.

candidate types. While non-incumbent candidates were most likely to cite their democratic duty, incumbents were far more likely to cite the opportunity to showcase their policy platforms to voters. All candidates viewed the debates as an opportunity to publicize their campaigns, but challengers put more emphasis than other candidates on the opportunity to demonstrate their competence. Only a small share of candidates pointed to radio broadcasting as a reason for their participation or admitted to attending in order to attack other candidates.

### III. Reasons Cited For Debate Participation

Candidate type	Duty (1)	Policies (2)	Competence (3)	Publicity (4)	Radio (5)	Attack (6)
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 allowed 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 why candidates were not present at their debate, although candidates predictably cite logistical issues rather than the electoral risks they faced in participation. Over 50% of non-participating candidates cited late notice, while 30% claimed that they did not receive an invitation. Nearly 20% mentioned road conditions to justify their absence.

As Table IV shows, the unbiased rules of debate moderation were generally kept and candidates were given equal time to outline their policy priorities. Candidates varied in how they emphasized their own qualifications during introductions, with incumbents focusing on their experience while challengers and other candidates highlighted their educational achievements. The most commonly cited policy priorities related to district primary schools, health facilities, and the quality and extent of roads. However, incumbents spoke at greater length about both the County Social Development Fund (CSDF) and the Legislative Support Project (LSP) funds. This makes sense as they have first-hand experience with these funds. Finally, incumbents were much more likely to both be attacked by other candidates and attack others, as their attendance seemed to act as a focal point for other debate participants.

Qualitative evidence from focus groups suggests that citizens were affected by candidate participation in the debates and found the presentation of concrete policy platforms to be a novelty. As one participant said, “Before the debate, the word ‘platform’ was a strange

#### IV. Transcript Descriptive Statistics

Candidate type	Intro words (1)	Education emphasis (2)	Experience emphasis (3)	CSDF words (4)	LSP words (5)	Attacked (6)	Attacker (7)
Incumbent	340.30	0.22	0.33	398.16	224.04	0.19	0.15
Challenger	351.99	0.30	0.26	284.66	218.00	0.04	0.04
Other	345.89	0.27	0.19	269.80	203.66	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; ‘Education emphasis’: candidate highlighted their education in introduction; ‘Experience emphasis’: candidate highlighted their experience 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.

word to me” (Vai Town, 26 September 2017). Many commented that the debates increased the accessibility of candidates, noting 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” (Foya, 20 September 2017). As a result, it is not surprising that citizens took note of 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 [the] other four candidates” (Massabolahun, 21 September 2017). Some even wanted debates to be mandatory: “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” (Vai Town, 26 September 2017).

Our qualitative evidence also suggests that the debates caused voters to change their assessment of the candidates. One participant stated, “The debate changed my attitude toward candidates and helped me discover the hidden secret of some candidates.” (Kolahun, 18 September 2017). Similarly, another participant mentioned, “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.” (Klay, 27 September 2017). Several focus groups pointed to the lack of specificity in candidates’ promises 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.” (Voinjama, 12 September). In particular, some respondents argued that the policy platforms of challengers were often weaker than those of the incumbents, as exemplified by one person’s comment that “I did not hear anything new from candidates contesting against the incumbent because the incumbent was already doing most of these things.” (Kolahun,

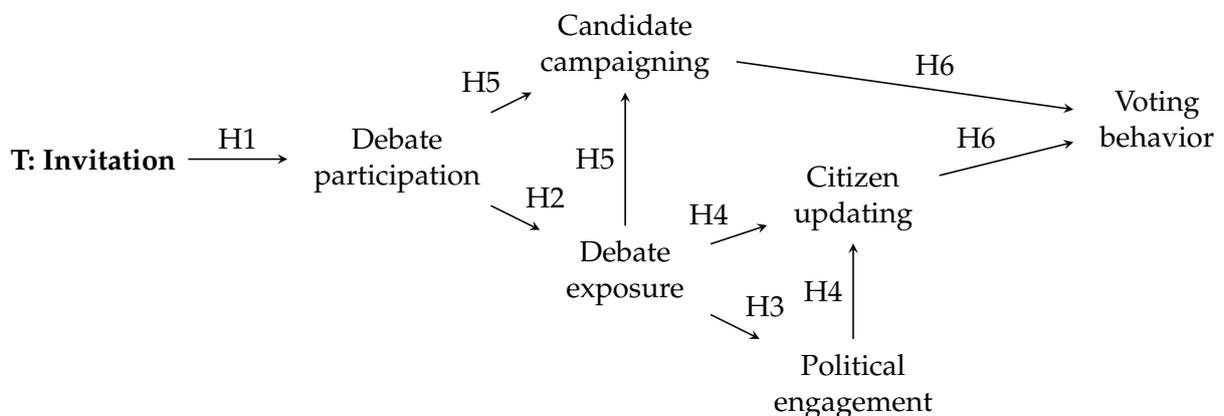
18 September 2017). Finally, citizens in our focus groups pointed to the importance of the mode of delivery for future accountability by highlighting that “the radio broadcast is a way of record-keeping. If you break your promise once elected, you will be on record for that.” (Klay, 27 September 2017).

## 4 Hypotheses

Our goal is to assess the consequences of a shock to the supply of policy promises by leading candidates, to be subsequently disseminated by the media, in a context characterized by limited programmatic competition. We evaluate the impact of our randomized intervention on candidate participation in the debates, citizen exposure to the debates, citizen beliefs about candidates’ competence and policy promises, citizen political engagement, candidate campaigning, and voting behavior.

We refer to high-intensity invitation districts as treated districts.<sup>21</sup> We present the relationships among our pre-registered hypotheses in Figure 2. 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, particularly predicted leading candidates, would attend 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, allaying any concerns and reducing

<sup>21</sup>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 hypotheses are largely unchanged despite pooling the rebroadcasting treatments.

uncertainty, and persuading them to participate. Our first hypothesis, therefore, was that this added recruitment effort would lead 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 leading candidates, both incumbents and predicted challengers, in each race. As mentioned earlier, incumbents and leading 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.

**H2.** Citizens would be more exposed to the debates in treated districts.

We anticipated that more citizens in treated districts would be exposed to the debates than in control districts because the participation of a higher share of leading candidates should increase citizen interest in the debate. In principle this should operate on both the extensive margin, increasing the probability of citizens hearing the debate, and the intensive margin, with citizens paying more attention to the content conditional on hearing it. This hypothesis, while important, is not obvious because the campaigning season is frenetic 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, especially when leading candidates attended the debates. Further, with each debate rebroadcast an average of six times leading up to the election, we expected that citizens would have sufficient opportunity to listen to the content if they wanted.

**H3.** Citizens would increase their political engagement in treated districts.

Due to the novelty of the information conveyed to voters through the debate initiative, we expected broad treatment effects on political engagement. In particular, we expected that exposure to relevant political information from leading candidates would increase citizen demand for even more political information, both through their social networks and from the media. Further, based on the large body of research on the political role of media, we hypothesized that the broadcasting of information about leading candidates would lead to increased discussion of the debates with others, influencing common beliefs about the candidates. This should amplify the direct exposure effects of individuals to the debate content and lead to coordinated vote choices.

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

We anticipated that increased exposure to the debates and subsequent political information acquisition in treated districts would lead citizens to learn more about the participating candidates. With so many candidates and political parties competing, we expected that listening to a set of candidates answer standardized questions would offer an unusual opportunity to acquire and compare information on their policy positions and competence. This updating could either result from shifting citizens' beliefs about candidate competence and policy priorities or by reducing the uncertainty around them. We conjectured, therefore, that citizen evaluations of the leading candidates induced to attend the debates would be affected by treatment as a result of an increase in both the share of those candidates supplying information, and the demand for political information more broadly.

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

In contrast to localized debate interventions, we hypothesized that the ability of candidates to increase their on-the-ground campaigning efforts in response to increased debate attendance by leading candidates 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 where voters were more exposed to the debate. Further, since candidates were unaware of rebroadcasting plans, candidates might have had little time to respond to the rebroadcasts themselves.

In fact, debates could crowd out on-the-ground campaigning by leading candidates who were persuaded to participate. By attending, candidates commit themselves—at some level—to campaigning on a more programmatic basis rather than focusing on the distribution of cash and promises at the village level. In addition, if debate participation benefited those candidates who were persuaded to attend the debates, their opponents might then rationally reduce campaigning effort.

**H6a.** Citizens in treated districts would be more likely to vote for well-performing candidates.

**H6b.** Citizens in treated districts would be more likely to vote for candidates who were aligned with their policy preferences.

Ultimately, we anticipated that treatment assignment would lead to changes in citizen vote choices. If citizens in treated districts were more exposed to debate content, and therefore to

the policy promises of the higher share of leading candidates participating, then we expected two key heterogeneous results. First, that citizens would more frequently vote for candidates who performed better in the relevant debate. Second, given the programmatic focus of the initiative, that better-aligned candidates would accrue more votes in treated districts compared to control. These heterogeneities stemmed from our belief that debate participation would not have uniform net returns across candidates.

## 5 Data

### 5.1 Data sources

Our primary data source is a panel survey of 4,060 registered voters conducted in 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 A1 indicate, the sample is older, more male, and better educated than the average Liberian.

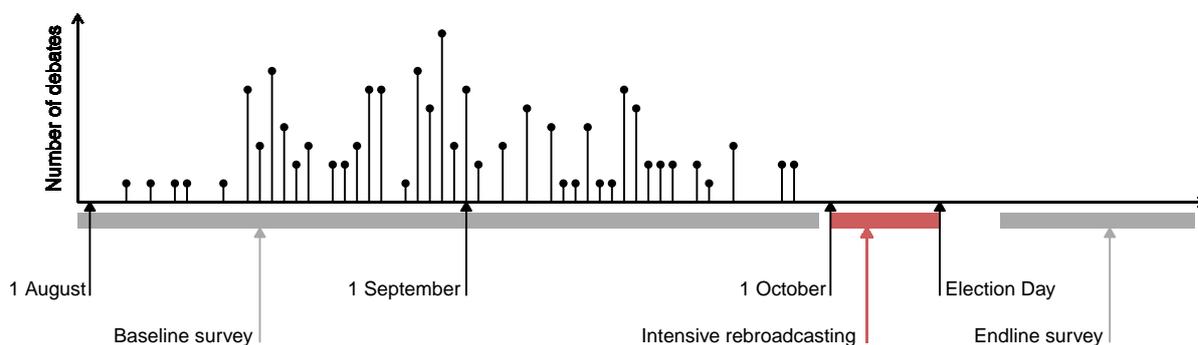
In Figure 3 we show a timeline of the debates and data collection. Our survey began in early August, right before the first debates. Most data collection was completed by early September but concluding the baseline survey in several electoral districts took several more weeks.<sup>22</sup> The overlap of the baseline survey and the live debates is not a major concern. First, we control for any potential baseline debate exposure using the date on which respondents were interviewed. Second, for variables that were collected only for 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 of the intervention 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 of the mechanism behind our differential effects by incumbency. Third, we use

---

<sup>22</sup>Since the cellphone-number sample was stratified at the county-level, sampling in some districts proved difficult, especially when the county contained both urban and rural districts since, in these cases, most calls went to those in the urban districts, and so achieving a sufficient sample in the rural districts took longer than anticipated.

Figure 3. Timeline of Debates Initiative



debate transcripts from each debate. Internews partnered with the Daily Observer newspaper and hired trained journalists to transcribe each debate. Fourth, we conducted a survey of more than 50 radio stations to gather more descriptive evidence about these stations and to validate their frequency of rebroadcasting the debates. This survey included radio stations who were not contracted for rebroadcasting to assess the extent of unintended rebroadcasting of debate content.

## 5.2 Outcome variables

To assess whether the invitation intervention, the debates themselves, and the debate rebroadcasting were properly implemented, we exploit two main pieces of data. For candidate debate participation, we use data from administrative debate reports as well as debate transcripts. 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, unedited and on schedule. We complement this data with responses to our survey of radio stations to assess whether contracted and non-contracted stations also rebroadcast 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 while details on their construction are in Appendix A.4. Whenever relevant, we aggregate related outcome variables using standardized z-scores as described by Kling et al. (2007). To measure debate exposure, we asked respondents whether they heard the debate as well as factual questions about the debate to confirm their exposure. We assess effects on political

engagement using survey questions about individual demand for political information and coordination within their social networks. We measure turnout using administrative data. In Table A2 we provide descriptive statistics for all the raw respondent level outcome variables.

To assess individual beliefs about the candidates in their districts, we ask respondents about only three *predicted leading candidates*, as described in Section 2.2. 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 campaign efforts of each of these predicted leading candidates and respondent beliefs about their competence and policy platforms. For all respondent-candidate dyads, we split the analysis into the incumbent and a pooling of the predicted challengers. Ultimately we are interested in respondent vote choices, which we validate using polling station-level data. In Table A3, we provide descriptive statistics for all the raw respondent-candidate level outcome variables.

Because we had no control over the local issues that would be discussed in the debates, all questions about policy priorities were open ended. These include both the questions about the policy priorities of the voters as well as their perceptions about the priorities of the *predicted leading candidates*. To analyze these priorities, open survey responses were coded by independent coders with no knowledge of treatment assignment.

### 5.3 Interaction variables

As stated in H6a and H6b, we expected that voting outcomes would be affected by two key variables: candidate performance in the debate and the extent of alignment between citizens' and candidates' policy priorities. In our survey analysis, we asked citizens about who they thought won their district debate. This performance measure, however, is naturally correlated with treatment assignment and so in the main analysis we construct a *predicted* candidate-level measure of debate performance,<sup>23</sup> and use this to predict performance for the full sample of candidates irrespective of actual participation. In contrast to other debate studies (Bidwell et al., 2016; Platas and Raffler, 2017), using an expert panel to measure performance of all candidates was unfeasible since we lack universal candidate attendance.

We measure the extent of preference alignment between respondents and candidates using

---

<sup>23</sup>We do this by flexibly estimating a LASSO model of the debate performance outcomes of participating candidates on pre-determined covariates including their gender, incumbency status, party, baseline citizen assessments of their competence, policy priorities, and certainty regarding both competence and policy priorities.

data from our baseline survey in which we asked respondents to name their top three policy priorities in their district as well as to name what they believe are the top three policy priorities for each of the three predicted leading candidates. We aggregate this latter measure to the district-level across respondents to create a measure of that candidate’s policy priorities. We then calculate the share of a respondent’s top issues that are shared with each candidate to create a measure of preference alignment. We also created a district-level version of preference alignment where we calculate the average of this variable at the district-level. However, because our sample is not representative of the district as a whole, we consider the district-level measure of preference alignment to be a noisy measure of the match between candidates’ policy priorities and all voters in the district. We provide these results for completeness.

## 6 Estimation

### 6.1 Treatment assignment and balance

We randomly assigned all 73 districts into one of two treatment conditions: low invitation effort (the control group) or high invitation effort (the treatment group). To assign treatment conditions, we first 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. Second, we blocked on a set of pre-treatment covariates at the district-level to maximize power.<sup>24</sup> This strategy generated 19 blocks, with 3 or 4 districts per block. We then randomly assigned 38 districts to the treatment and 35 to control. Pre-treatment covariates at the district, individual, polling station, and candidate levels are well-balanced across treatment groups. 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:

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

where  $y_{icd}$  is the outcome for respondent  $i$  regarding candidate  $c$  in district  $d$ .<sup>25</sup>  $T$  is an indi-

<sup>24</sup>We blocked on variables described in Panel A of Table A1.

<sup>25</sup>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,  $y_{cd}$ .

cator for districts assigned to treatment.  $\eta_b$  are randomization block fixed effects and  $\theta_e$  are survey enumerator fixed effects. Throughout, we include both district-level covariates  $\mathbf{Z}_d$  and individual-level covariates  $\mathbf{X}_i$  for which we provide descriptive statistics in Panels A and B of Table A1, respectively. While we have good balance on these covariates we include them to improve precision. We cluster standard errors at the district level, 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').

We consider an analogous specification to that of Equation (1) for electoral outcomes available at the polling station-level. In this case, instead of individual-level controls, we control for polling station-level variables listed in Panel C of Table A1. At the polling station-level, we report specifications which are unweighted, 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  $\Delta y_{icd}$ . 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 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.

Our coefficient of interest in Equation (1) is  $\beta$ , which recovers the reduced-form or intent-to-treat treatment effect of the invitation intervention. While the treatment effect of debate participation, or the instrumental variable (IV) estimate that we would recover by instrumenting debate participation with assignment to treatment, is of great theoretical interest, some of our estimates indicate that the exclusion restriction condition that is required for the IV estimate to be unbiased is unlikely to hold. Lastly, while reduced-form estimates recover the effects of the intervention for complying candidates, in Section 8 we argue that, due to the nature of selection into debate participation, these estimates speak to the effect of the debate initiative more generally.

## 7 Results

Our results suggest rich and consistent consequences of the debate initiative, and in particular the invitation intervention, on political outcomes. First, we show that the intervention was successful as candidates, especially leading candidates, were more likely to attend their debates in treated districts. Second, we find strong evidence that citizens were more exposed to the debates in these districts. Third, we find increased political engagement in treated districts. These results echo the qualitative evidence from our focus groups, that citizens were more interested in the debate content in districts where the most relevant candidates participated.

Fourth, we show that citizens learned more about incumbent candidates' policy priorities and competence, and to a lesser extent about challengers' priorities, in treated districts than they did in control districts. There is 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 suggesting that incumbents dominated their debates. Fifth, challengers decreased their on-the-ground campaigning efforts and incumbents increased their radio campaigning in treated districts. Finally, as a result, in treated districts citizens voted more often for their incumbent, especially when they performed well or matched baseline voter policy preferences. We present results here using indices standardized such that the control group mean is zero with standard deviation one, but we provide treatment effects on individual outcome variables in the Appendix.

### 7.1 Intervention implementation

We start by testing H1 to assess whether the intervention had its intended effect. Table V reports treatment effects on the attendance of candidates at the debates.<sup>26</sup> 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 attending the debates in treated districts. In Panel B we show that incumbents were 21.2 pp (76%) more likely to attend in treated districts, and in Panel C we show that predicted challenger candidates were 21.2 pp (43%) more likely to attend. Lastly, Panel D reports no treatment effect on other candidates, reinforcing our expectation that the invitation intervention would mainly affect the debate participation decision of leading

---

<sup>26</sup>In Appendix Table A13 we show that there is no difference in how often the debates were rebroadcast. This is measured either based on the radio monitors who tuned into each contracted rebroadcast, or based on our radio station survey that included radio stations not contracted to rebroadcast but recorded as present in a district debate.

## V. Candidate Debate Participation

	(1)	(2)	(3)
<b>A. Share of candidates</b>			
<i>Invite</i>	0.077** (0.034)	0.065** (0.030)	0.092*** (0.033)
Control Mean	0.542	0.573	0.557
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
<b>B. Incumbent</b>			
<i>Invite</i>	0.212** (0.083)	0.177** (0.073)	0.234*** (0.083)
Control Mean	0.280	0.372	0.299
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
<b>C. Share of challengers</b>			
<i>Invite</i>	0.212*** (0.074)	0.144** (0.063)	0.220*** (0.067)
Control Mean	0.492	0.554	0.528
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
<b>D. Share of other candidates</b>			
<i>Invite</i>	0.003 (0.030)	0.008 (0.028)	0.009 (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, incumbent, predicted challenger, other candidate) who attended a debate out of all candidates in that district. Panels A-D have 4060 observations; Panel D 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.

candidates.<sup>27</sup>

## 7.2 Exposure to debates

We next report the effects of treatment assignment on our measures of citizen exposure to the debates (H2), which should reflect that citizens are more interested in the debates when leading candidates attend. In Panel A of Table VI, we use a standardized index of our measures of direct exposure to the debates (whether respondents heard the debate and how often they heard them). The results indicate that citizens in treated districts had exposure 0.255 standard deviations (sd) higher than those in control districts.<sup>28</sup>

In Panel B, we use an index reflecting factual knowledge about the debates themselves, including whether the respondent's stated debate winner actually attended the debate, and the share of candidates and leading candidates respondents claims participated. The results suggest that citizens in treated districts had factual knowledge of the debates 0.113 sd higher than citizens in control districts. Given that our endline survey began around a month after the election, the persistence of this effect suggests meaningful differences in debate exposure. In Panel C we use an index reflecting factual knowledge about a national policy issue, the management of County Social Development Funds, which was asked in every debate. We find treatment effects of around 0.1 sd on correctly learning about management issues of these poorly-understood funds. This suggests that the debates were relatively effective at conveying policy information, particularly when leading candidates participated. We provide results disaggregating the components of these indices in Table A14. These results provide strong support for the idea that the invitation intervention meaningfully affected citizen exposure to the debate content because more leading candidates attended.<sup>29</sup>

## 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 leading candidates, the invitation

---

<sup>27</sup>In Appendix Table A12, we further show that whether the eventual election winner attended, and the share of actual challengers attended was substantially affected by treatment assignment, as expected.

<sup>28</sup>Importantly for our ability to pool districts assigned to different rebroadcasting intensity, as we show in Appendix A.3, there were no differential treatment effects on citizen exposure to the debates when pooling districts which were cross-randomized to different rebroadcasting intensities or when restricting to districts assigned to high invitation intensity.

<sup>29</sup>This is also consistent with the theory that more deliberation between relevant candidates will increase voters' attention (Wantchekon et al., 2018).

## VI. Debate Exposure

	(1)	(2)	(3)
<b>A. Debate listening index</b>			
<i>Invite</i>	0.255*** (0.086)	0.281*** (0.087)	0.353*** (0.089)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
<b>B. Debate knowledge index</b>			
<i>Invite</i>	0.113** (0.057)	0.112** (0.052)	0.147*** (0.053)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
<b>C. Policy knowledge index</b>			
<i>Invite</i>	0.097* (0.055)	0.143* (0.075)	0.116* (0.062)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs

Outcome variables: Panel A: a standardized index of (1) indicator for whether the respondent had not heard their district debate at baseline but had at endline (2) indicator for whether the respondent had heard the debate at endline (3) the number of times the respondent had heard the debate at endline. Panel B: a standardized index of (1) indicator for whether the respondent's stated debate winner actually attended the debate (2) share of candidates respondent claims participated (3) share of predicted leading candidates respondent claims participated. Panel C: a standardized index of the change in how many factual questions about CSDF management respondents answered correctly 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.

intervention affected the political engagement of citizens and ultimately their turnout (H3). Panel A of Table VII reports 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 by listening to the radio, discussing with friends, and seeking other forms of political information.

#### VII. Political Engagement

	(1)	(2)	(3)
<b>A. Political information demand index</b>			
<i>Invite</i>	0.169*** (0.052)	0.208*** (0.060)	0.201*** (0.061)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
<b>B. Debate coordination index</b>			
<i>Invite</i>	0.149** (0.063)	0.153** (0.059)	0.190*** (0.060)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
<b>C. Polling station-level turnout</b>			
<i>Invite</i>	0.015*** (0.005)	0.012** (0.005)	0.015*** (0.005)
Control Mean	0.701	0.699	0.699
Observations	5386	5386	5386
Weight	No	1/PS	Reg

Outcome variables: Panel A: standardized 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: index of variables measuring (1) how much respondents discussed the debate with friends (2) how much this discussion led to coordinating their vote choices. Panel C: polling station-level turnout in House of Representatives election.

Specifications using citizen survey outcomes 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. Specifications using polling station-level data are estimated using OLS and include block FE, district-level and polling station-level controls. For weighted specifications, 'PS' is the number of polling stations in that district and 'Reg' is the number of registered voters at that polling station. \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01. Standard errors clustered at the district-level in parentheses.

In line with these results, in Panel B there are strong treatment effects on a standardized coordination index capturing whether citizens discussed the debates with others and whether they believed these discussions led them to coordinate their vote choices. Specifically, citizens exhibit a 0.149 sd increase in coordination in treated districts. Both panels then provide strong

evidence that, when exposed to debates in which relevant leading candidates participated, citizens demanded more political information and coordinated their vote choices. We provide results decomposing these indices in Table A15.

Next, we provide evidence that the citizen engagement caused by the invitation intervention ultimately led to higher turnout. In Panel C of Table VII, we use administrative polling station-level data on turnout. Turnout in the House of Representatives election was on average 1.5 pp higher in treated districts than the control mean of 70.1 percent (2%). Overall, the results provide strong evidence that greater candidate participation in the debates increased political engagement of citizens, which in turn generated actual electoral mobilization.

#### **7.4 Effects on beliefs about candidates**

We next analyze the effect of the intervention on beliefs about the competence and policy priorities of predicted leading candidates (H4). 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 incumbents, shown in Panel 1.A. of Table VIII. The results indicate that citizens in treated districts became significantly more certain about incumbent competence and priority issues, with standardized treatment effect sizes of 0.178 and 0.169 sd, respectively. In Panel 1.B. there is no evidence that citizens in treated districts became more certain about the competence of challengers, and mixed 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 leading candidates, shown in Panel 2. These treatment effects suggest positive updating regarding incumbents and negative updating for their challengers, but the estimates are imprecise. Panel 2.A. reports suggestive sizable, but statistically insignificant, treatment effects that citizens feel more positive about incumbent competence and that they have learned more about their policy priorities. Specifically, for both outcomes, the treatment effect is around 0.1 sd. In contrast, Panel 2.B. indicates that citizens updated negatively about the competence of their challengers and they did not learn about the policy priorities of those candidates. The treatment effect on challenger competence is 0.1 sd and marginally significant.

The results provide evidence that aligns with the focus group evidence in Section 3.3. Inducing the participation of leading candidates increased certainty regarding their competence and priority issues among their electorate. Indeed, as we show in Section 8, the candidates

VIII. Updating About Candidates

1. Uncertainty	Certainty about competence			Certainty about policy		
	(1)	(2)	(3)	(4)	(5)	(6)
<b>A. Incumbent</b>						
<i>Invite</i>	0.178* (0.105)	0.186* (0.107)	0.179** (0.084)	0.169** (0.074)	0.195** (0.080)	0.192** (0.074)
Observations	3496	3496	3496	3496	3496	3496
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs
<b>B. Challengers</b>						
<i>Invite</i>	0.037 (0.066)	0.046 (0.074)	0.025 (0.070)	0.139** (0.061)	0.118 (0.073)	0.098 (0.067)
Observations	8684	8684	8684	8684	8684	8684
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs
2. Levels	Beliefs about competence			Learning about policy		
	(1)	(2)	(3)	(4)	(5)	(6)
<b>A. Incumbent</b>						
<i>Invite</i>	0.098 (0.075)	0.066 (0.089)	0.093 (0.082)	0.089 (0.065)	0.125 (0.089)	0.091 (0.073)
Observations	3496	3496	3496	3496	3496	3496
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs
<b>B. Challengers</b>						
<i>Invite</i>	-0.078 (0.075)	-0.147* (0.088)	-0.096 (0.077)	0.038 (0.065)	0.027 (0.086)	0.063 (0.080)
Observations	8684	8684	8684	8684	8684	8684
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs

Panels 1.A and 1.B: the outcome variable in columns 1-3 is the standardized change in certainty respondents express about candidate competence between baseline and endline, and in columns 4-6 is the standardized change in certainty respondents express about candidate priority issues between baseline and endline. Panels 2.A and 2.B: the outcome variable in columns 1-3 is the standardized change in respondent perception about candidate competence between baseline and endline, and in columns 4-6 is the standardized learning that respondents reflect 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.

induced to participate by the intervention were those about whom citizens were initially most uncertain. This increase in certainty was concentrated with respect to the incumbent, who spoke significantly more on the national policy issue questions for which they possessed more experience. Similarly, the suggestive positive treatment effects on citizens' perceptions about incumbent competence but negative perceptions about challenger competence indicate that incumbents were likely to outperform challengers when they were induced to participate. Overall the evidence supports that incumbents but not their challengers benefited both in terms of increased voter information about their policy priorities and somewhat in terms of perceptions of their competence.

## 7.5 Candidate response and campaigning

Next, we consider treatment effects on candidates' campaign efforts (H5). In Table IX, we report results on standardized indices of survey responses regarding "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' towns. The radio campaigning index captures candidate presence on the radio.

In Panel A, 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 B reports 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. We provide results disaggregating the on-the-ground campaigning index in Table A16.

This pattern of results suggests that the presence of incumbents in the debates deterred challengers, who spent less time campaigning in person around their district in treated districts than did challengers in control districts. Moreover, incumbent debate performance led incumbents to increase their use of radio campaigning to complement the rebroadcasting of the debates. Suggestively, in our candidate survey, 77% of incumbents believed that the debates led radio stations to issue invitations for interviews more frequently, while only 63% of challengers believed they received more invitations. In addition, 68% of incumbents report that the debates led them to change their campaign strategy whereas only 52% of challengers reported a change. Restricting the data to candidates who said the debates affected their mode of campaigning, 60%

## IX. Candidate Campaigning

	Ground			Radio		
	(1)	(2)	(3)	(4)	(5)	(6)
<b>A. Incumbent</b>						
<i>Invite</i>	-0.032 (0.025)	-0.044 (0.029)	-0.030 (0.030)	0.082** (0.037)	0.087** (0.040)	0.092** (0.042)
Observations	3496	3496	3496	3496	3496	3496
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs
<b>B. Challengers</b>						
<i>Invite</i>	-0.037* (0.019)	-0.045** (0.017)	-0.046** (0.019)	-0.025 (0.028)	-0.004 (0.027)	-0.018 (0.028)
Observations	8684	8684	8684	8684	8684	8684
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs

Outcome variable in columns 1-3 is a standardized index of how often candidates (1) visited (2) distributed leaflets (3) bought votes in respondents' communities during campaigning. Outcome variable in columns 4-6 is a standardized measure of how often respondents heard candidates 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.

of incumbents said they affected their radio appearances while only 40% said they affected their on-the-ground campaigning. By contrast, only 27% of challengers said the debates affected their radio appearances compared to 73% who said they affected their on-the-ground campaigning.

### 7.6 Voting Behavior

Our results indicate that the intervention benefited incumbents induced to participate the most. While both incumbent and challenger candidates were more likely to attend the debates in treated districts, politically motivated citizens learned more and felt 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 these changes ultimately affected voting behavior.<sup>30</sup>

Table X presents treatment effects on vote choice using both our citizen survey (Panel 1) and polling station-level electoral returns (Panel 2). Panel 1 presents a set of specifications testing for whether there were treatment effects on vote choice, defined as whether the respondent reported

<sup>30</sup>Appendix Table A17 reports that citizens in treated districts show no change in the reasons they cited for their vote choice in the baseline versus endline surveys, which thus cannot account for our results on voting.

## X. Voting Outcomes

	Main effect			Std. Performance			Interaction term: Std. Policy match		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<b>1. Respondent-level</b>									
<b>1.A. Incumbent</b>									
<i>Invite</i>	0.045**	0.051**	0.035*	0.049**	0.054**	0.039*	0.041**	0.045**	0.032
	(0.020)	(0.023)	(0.020)	(0.020)	(0.023)	(0.020)	(0.019)	(0.023)	(0.019)
<i>Invite</i> × $\widehat{\text{Std. performance}}$				0.067*	0.059	0.039			
				(0.038)	(0.036)	(0.031)			
<i>Invite</i> × Std. policy match							0.042***	0.051***	0.052***
							(0.016)	(0.017)	(0.018)
Control Mean	0.278	0.282	0.288	0.278	0.282	0.288	0.278	0.282	0.288
Observations	3496	3496	3496	3496	3496	3496	3496	3496	3496
<b>1.B. Challengers</b>									
<i>Invite</i>	-0.048***	-0.039***	-0.036**	-0.051***	-0.039**	-0.039**	-0.048***	-0.039***	-0.036**
	(0.014)	(0.013)	(0.014)	(0.017)	(0.015)	(0.016)	(0.014)	(0.013)	(0.014)
<i>Invite</i> × $\widehat{\text{Std. performance}}$				-0.015	-0.002	-0.026			
				(0.046)	(0.039)	(0.043)			
<i>Invite</i> × Std. policy match							-0.006	-0.004	-0.003
							(0.007)	(0.008)	(0.008)
Control Mean	0.156	0.156	0.156	0.156	0.156	0.156	0.156	0.156	0.156
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
<b>2. Polling station-level</b>									
<b>2.A. Incumbent</b>									
<i>Invite</i>	0.042*	0.037	0.041*	0.045*	0.036*	0.044*	0.045*	0.041*	0.043*
	(0.024)	(0.023)	(0.024)	(0.023)	(0.021)	(0.023)	(0.023)	(0.022)	(0.023)
<i>Invite</i> × $\widehat{\text{Std. performance}}$				0.082***	0.099***	0.081***			
				(0.027)	(0.030)	(0.026)			
<i>Invite</i> × Std. policy match							-0.016	-0.017	-0.017
							(0.028)	(0.029)	(0.029)
Control Mean	0.246	0.250	0.250	0.246	0.250	0.250	0.246	0.250	0.250
Observations	4618	4618	4618	4618	4618	4618	4618	4618	4618
<b>2.B. Challengers</b>									
<i>Invite</i>	-0.029**	-0.031**	-0.028**	-0.033**	-0.036***	-0.031**	-0.029**	-0.029**	-0.029**
	(0.012)	(0.012)	(0.012)	(0.013)	(0.012)	(0.013)	(0.012)	(0.013)	(0.012)
<i>Invite</i> × $\widehat{\text{Std. performance}}$				-0.034	-0.031	-0.037			
				(0.027)	(0.029)	(0.026)			
<i>Invite</i> × Std. policy match							-0.000	-0.009	0.005
							(0.019)	(0.017)	(0.019)
Control Mean	0.113	0.112	0.112	0.113	0.112	0.112	0.113	0.112	0.112
Observations	11385	11385	11385	11385	11385	11385	11385	11385	11385
Weight	No	1/PS	Reg	No	1/PS	Reg	No	1/PS	Reg

Panels 1.A. and 1.B: Outcome variable is an indicator for whether respondent expressed voting for either the incumbent (Panel A) or a predicted challenger (Panel B) at endline. Columns 1-3 report the main effects, Columns 4-6 include interactions of treatment assignment with standardized candidate-level measures of predicted debate performance, and Columns 7-9 include interactions with standardized respondent-candidate-level measures of preference alignment measured at baseline. Panels 2.A. and 2.B: Outcome variables are vote shares over the number of registered voters for the incumbent (Panel A) or predicted challengers (Panel B) using polling station-level electoral data. We use the district-level analogues of the interaction terms used in Panel 1. Section 5.3 explains these interaction terms further.

Specifications using citizen survey outcomes 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. Specifications using polling station-level data are estimated using OLS and include block FE, district-level and polling station-level controls. For weighted specifications, 'PS' is the number of polling stations in that district and 'Reg' is the number of registered voters at that polling station. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors clustered at the district-level in parentheses.

voting for a given candidate at endline.<sup>31</sup> Columns 1-3 present the main effects of the treatment assignment. Columns 4-6 present specifications in which we interact treatment assignment with the standardized measure of predicted 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 1.A., focusing on the incumbent, we find strong positive treatment effects on vote choice. Incumbents experienced a 4.5 pp increase in respondent vote choice in treated districts. Moreover, the significantly positive interactions with both debate performance and policy priority alignment indicate that this effect is greater among incumbents who were predicted to perform well in the debates and whose policy priorities align with those of their electorates. In contrast, focusing on predicted challengers in Panel 1.B., there are broadly negative main effects and little evidence of significant interactions. Challengers experienced a significant 4.8 pp drop among respondents' vote choice in treated districts.

In Panel 2 of Table X, we use polling station-level data to augment our survey results. These results reinforce our findings. In Panel 2.A., we find that incumbent vote share in treated districts was 4.2 pp higher in treated districts than in control districts, albeit somewhat more noisily estimated. Incumbent vote share is larger in districts where their predicted debate performance was higher. However, we find no significant interaction with our district measure of policy priority overlap here, which we attribute to the fact that our sample of district respondents might be a noisy measure of the true district policy priority overlap. In Panel 2.B., we continue to find that challenger vote share in treated districts is 2.9 pp smaller than in control districts, and there is no evidence of interactions with predicted debate performance or overlap in policy priorities between challengers and citizens in their districts. The polling station results broadly support the survey results, and suggest positive electoral consequences for incumbents in treated districts but negative consequences for their challengers.

## 8 Discussion

Our results suggest that electoral gains accrued to incumbent candidates in treated districts, particularly when they performed well in their debates and showed that their policy priorities matched those of their constituents. Importantly, the results from our survey of vote choice

---

<sup>31</sup>We pre-registered using an outcome variable defined as vote switching towards a given candidate. However, since few respondents indicated a concrete vote choice at baseline, using either measure produces qualitatively identical results. Focusing on endline responses allows us to compare voter responses to polling station outcomes.

and polling station outcomes mirror actual election outcomes: 50% of incumbents in treated districts won re-election compared to 43% in control districts. This difference is not surprising given that more than 35% of races are decided by narrow winning margins of less than five percentage points. But, in a context where approval of incumbent performance is generally low, as discussed in Section 2.1, these results might seem surprising. 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%—whereas the additional invitation effort increased this number to around 50% in treated districts. This leaves a sizable share of incumbents who did not attend.

We show that these incumbents positively self-selected into debate participation to a greater extent than challenger candidates, particularly in terms of the degree of their policy alignment with citizens. We focus on standardized candidate-level measures of (1) baseline policy alignment with citizens in the district; (2) baseline citizen uncertainty about candidate policy priorities; (3) baseline citizen assessments of candidate competence; (4) baseline citizen uncertainty about candidate competence. Following Abadie (2003) and Angrist and Fernandez-Val (2013), we first compute the average of those variables for the various compliance groups: compliers (candidates induced to participate in the debates due to the invitation intervention), always-takers (candidates who would have participated absent the intervention), and never-takers (candidates who would not have participated even with the intervention).

In Panel A of Table XI we provide this descriptive evidence for incumbent candidates. We find that always-taker incumbents were best aligned with the policy priorities of citizens in their districts. Complier incumbents were aligned better than never-taker incumbents, but worse than always-takers. Importantly, however, and consistent with the idea that debate intervention allayed concerns about the risk of debate attendance, there was much less certainty about the policy priorities of complier incumbents than either always-taker or never-taker incumbents. We find a less stark difference in terms of self-selection on competence across groups, but again we find that there was less certainty about the competence of the compliers at baseline.

Focusing on challenger candidates in Panel B, we find a much less clear pattern of self-selection into the debates. While always-takers do not seem to self-select relative to never-takers in terms of their policy alignment with citizens, compliers *negatively* do. Similarly, while always-takers positively self-selected in terms of competence, the positive self-selection of compliers is significantly greater. As with incumbents, there was much less certainty about the policy priorities and competence of complying challengers than either always-taker or never-taker

challengers.

## XI. Characterizing Compliers

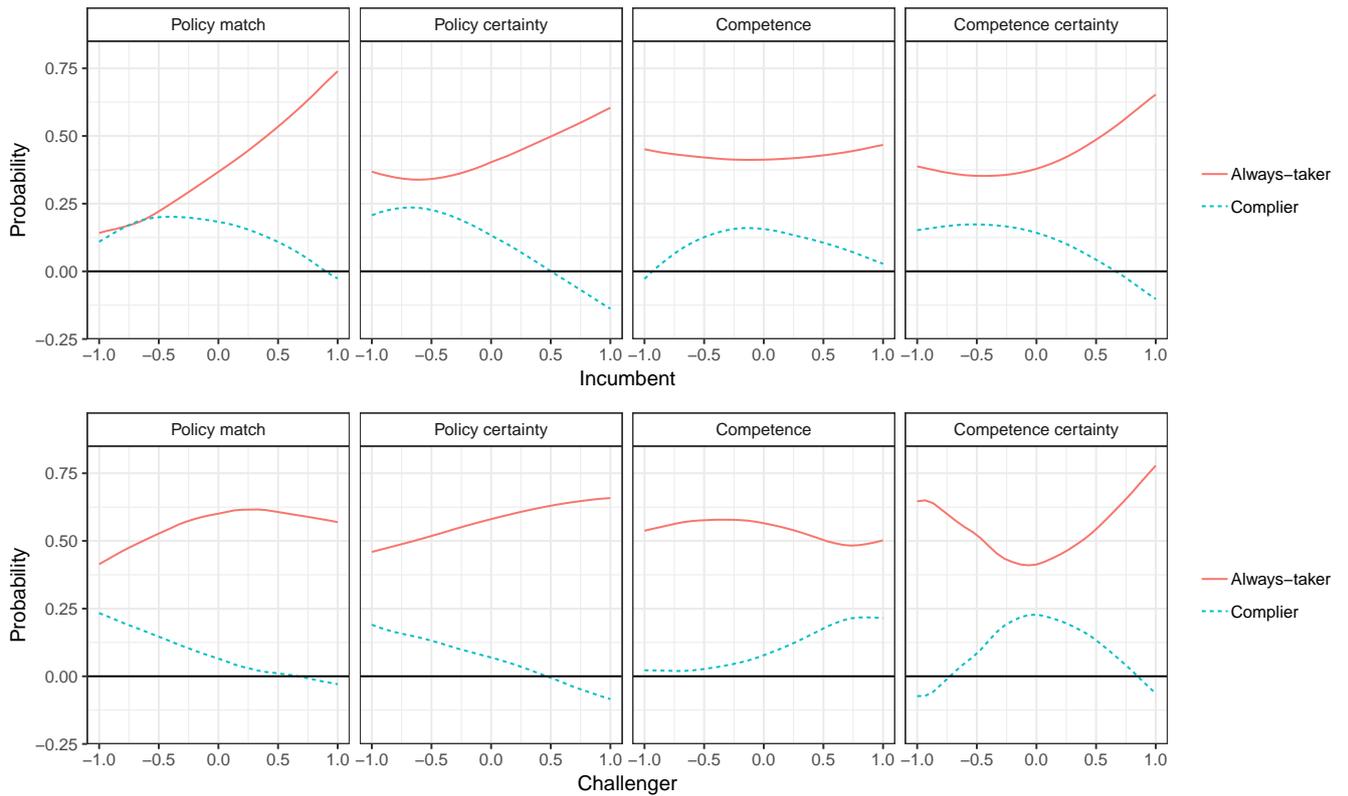
Variable	Compliers (1)	Always-takers (2)	Never-takers (3)
<b>A. Incumbents</b>			
Policy match	-0.22	0.60	-0.47
Policy certainty	-1.00	0.14	0.06
Competence	-0.04	0.08	-0.07
Competence certainty	-1.10	0.29	-0.01
<b>B. Challengers</b>			
Policy match	-1.38	0.17	0.14
Policy certainty	-1.19	0.12	0.10
Competence	0.45	0.06	-0.18
Competence certainty	-0.50	0.19	-0.19

We standardize each variable within the relevant candidate type. Estimates are the kappa-weighted means for various compliance groups by incumbency status: compliers, always-takers and never-takers, following Abadie (2003) and Angrist and Fernandez-Val (2013).

We provide additional graphical evidence regarding the nature of candidate selection into the debates using non-parametric regressions to examine the probability of a given candidate being an always-taker or complier across different values of the baseline characteristics we focus on. Figure 4 presents these results in solid red for always-takers and in dashed blue for compliers. The plots in the top panel again suggest strongly positive self-selection among always-taker incumbents based on their policy alignment with constituents. Among candidates in districts assigned to the control, incumbents with policy alignment 1 sd above the mean participated at rates of around 75%. Those with policy alignment 1 sd below mean participated less than 20% of the time. With respect to complying incumbents, the plots suggest that the invitation intervention induced the participation of incumbents at intermediate levels of policy alignment, which is natural given that those incumbents with the highest levels of policy alignment were likely always-takers. In turn, we continue to observe no self-selection based on competence. With respect to certainty about both candidate policies and competence, again we corroborate a strong positive self-selection among always-taker incumbents in that those candidates were relatively well-known to citizens along both dimensions.

The plots in the bottom panel, in turn, confirm a substantially weaker self-selection of always-taker challengers on both candidate policy alignment and competence. With respect to complying challengers, in contrast with complying incumbents, we see a negative self-selection on policy alignment and a positive selection on competence. However, as with complying

Figure 4. Complier Status by Baseline Candidate Characteristics



*Note:* Optimal bandwidths based on Calonico et al. (2018).

incumbents, we observe that there is substantially less certainty about both their policy priorities and competence.

Overall, the evidence indicates that incumbent candidates were much better at self-selecting into debate participation than challengers; they seemed better able to recognize when their policy priorities aligned with those of their constituents and when constituents were uncertain of this alignment. This effect is clear for both always-takers and compliers, which suggests that the treatment effects of the invitation intervention are likely representative of those of the overall debate initiative. In turn, across candidate types, selection on perceptions of overall competence was a lot weaker. Given the specific focus of the debates on policy issues, it appears that incumbents were better equipped to correctly assess the returns to debate participation.<sup>32</sup> Challenger candidates, who frequently had less political sophistication and experience, were less successful at calculating the returns to participation.

<sup>32</sup>It is worth considering whether this sophistication is a product of prior selection or experience accrued over incumbency. While underpowered to conduct a regression discontinuity design to account for selection into incumbency, this exercise provides suggestive evidence that it is experience, rather than selection into incumbency, that accounts for our observed results.

## 9 Conclusion

We evaluate an intervention intended to vary the participation of candidates in a nationwide debate initiative that was designed to elicit and disseminate programmatic promises from candidates running for a seat in Liberia's House of Representatives in 2017. Policy promises were elicited from candidates participating in debates and disseminated via radio rebroadcasting. The intervention electorally benefited the incumbents who were induced to attend their debates. These incumbents self-selected into the debates when their policy priorities were aligned with their constituents, and thus they dominated their debates by proposing policy platforms closer to those preferred by their electorates. Our results suggest that both the debate content and the mode of dissemination mattered critically for this outcome. When relevant candidates attended their debates voters paid substantially more attention to the debate information and subsequently engaged in more political information acquisition, 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 challengers reduced their on-the-ground campaigning.

Our results point to the challenges of transitioning away from the low-accountability equilibrium characterizing many clientelistic democracies. By inducing a subset of self-selected candidates to compete on a more programmatic basis, our intervention had uneven electoral consequences that favored incumbents. In this context, there may exist substantial returns to incumbency for some incumbents, as 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 and may fade in future electoral cycles as debates become less novel and candidate entry evolves (Le Pennec and Pons, 2018), the point is broader. If such democratic initiatives are to enhance competition and contribute to shifting from a clientelistic 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

- Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics*, 113(2):231–263.
- 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.
- Anderson, S., Francois, P., and Kotwal, A. (2015). Clientelism in indian villages. *American Economic Review*, 105(6):1780–1816.
- Angrist, J. D. and Fernandez-Val, I. (2013). Extrapolate-ing: External validity and overidentification in the late framework. In *Advances in Economics and Econometrics: Volume 3, Econometrics: Tenth World Congress*, volume 51, page 401. Cambridge University Press.
- 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.
- Baland, J.-M. and Robinson, J. A. (2008). Land and power: Theory and evidence from chile. *American Economic Review*, 98(5):1737–1765.

- Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., Shotland, M., 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.
- Banerjee, A. V., Kumar, S., Pande, R., and Su, F. (2011). Do informed voters make better choices? experimental evidence from urban india. Working paper.
- 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.
- Blattman, C., Larreguy, H., Marx, B., and Reid, O. (2018). A market equilibrium approach to reduce the incidence of vote-buying: Evidence from uganda.
- 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.
- Calonico, S., Cattaneo, M. D., and Farrell, M. H. (2018). On the effect of bias estimation on coverage accuracy in nonparametric inference. *Journal of the American Statistical Association*, pages 1–13.
- 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. (2018a). Making policies matter: Voter responses to campaign promises. Working paper.
- Cruz, C., Labonne, J., and Querubín, P. (2018b). Politician family networks and electoral outcomes: Evidence from the philippines. *American Economic Review*, 107(10):3006–3037.
- Dal Bó, E. and Finan, F. (2018). Progress and perspectives in the study of political selection. *Annual Review of Economics*, (0).
- Dal Bó, E., Finan, F., Folke, O., Persson, T., and Rickne, J. (2017). Who becomes a politician? *Quarterly Journal of Economics*, 132(4):1877–1914.
- 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., McIntosh, C., and Nellis, G. (2019). *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.
- Ferraz, C. and Finan, F. (2011). Motivating politicians: The impacts of monetary incentives on quality and performance. Working paper.

- 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.
- Gagliarducci, S. and Nannicini, T. (2013). Do better paid politicians perform better? disentangling incentives from selection. *Journal of the European Economic Association*, 11(2):369–398.
- 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*.
- Gulzar, S. and Khan, M. Y. (2018). Barriers to political entry: Experimental evidence from local government elections in pakistan. 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.
- Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.

- Larreguy, H. A., Marshall, J., and Snyder, James M., J. (2017). Publicizing malfeasance: When the media structure facilitates electoral accountability in Mexico. Working Paper.
- Le Pennec, C. and Pons, V. (2018). Time, not TV debates, changes voters' mind: Evidence from 61 elections across the world. 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.
- Robinson, J. A. and Verdier, T. (2013). The political economy of clientelism. *Scandinavian Journal of Economics*, 115(2):260–291.
- 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. (2018). 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.

## A Appendix

### A.1 Classifying candidates as leading candidates

For each candidate we constructed an indicator variable for whether the candidate was a *predicted leading candidate*. 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 leading candidates* 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 leading candidate. 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 leading candidates came in the top three in their district, and in 71% of cases came in the top five. 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 A12 that, using the ‘actual’ leading candidates who placed in the top three in the election—whether *actual leading candidates*, *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* leading candidates in the citizen-candidate level analysis. However, given the effects we find on voting outcomes, we consider it likely that the definition of *actual leading candidates* 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  $X_i$  and district-level  $Z_d$  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, first we assign district-level outcomes to each polling station in that district and second we use polling station-level variables using the fact that 90% of polling places in 2017 were also used in the 2011 election. For all new polling places we assign district-level averages. 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 A4 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 A5 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 A6 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 A7 we report balance at the polling station level. We find evidence of imbalance on 0 (1) covariates out of 18 covariates.

In Table A9 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 A10 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, incumbents and challengers across treatment groups. Using the full sample of candidates, we find imbalance on 1 (2) covariates out of 8. 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 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 (with either two or ten rebroadcasts) 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 rebroadcasting 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 power. In Table A19 we show that the rebroadcasting intervention did not lead to significantly positive effects on debate exposure in either the full sample or the sample restricted to respondents in intensive-invitation districts, in spite of being correctly implemented. As discussed, this lack of effects stems from number of citizens who heard their district debates when they were only broadcast a few times.

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. Additionally, in our PAP we pre-registered the use of a jackknife measure of debate performance; given that actual debate performance is likely a function of treatment assignment we instead use the predicted measure of performance described in Section 5.3.

While our pre-registered hypotheses make reference to the distinction between incumbents and challenger candidates, particularly differential treatment effects of the intervention on their participation, we did not pre-register the descriptive analysis we perform regarding their intervention compliance behavior 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 19 of them are presented in the final paper.<sup>33</sup> The eight missing hypotheses fall into two 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 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

---

<sup>33</sup>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.

radio in these districts where they were more likely to participate. We report these results for completeness in Table A18. 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 standardize each component such that units in the control group have zero mean and standard deviation of one then average them, following Kling et al. (2007).

Table V:

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

Table VI:

- *Debate listening index*: standardized index of:
  - *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).

- *Debate knowledge index*: standardized index of:
  - *Debate winner attended debate*: indicator for whether respondent’s named debate winner actually attended the debate (E-Q17).
  - *Stated share of participating candidates*: continuous variable for the share of candidates in respondent’s district they recall participating in debate (E-Q16).
  - *Stated share of participating leading candidate*: continuous variable for the share of leading candidates in respondent’s district they recall participating in debate (E-Q39.2, E-Q39.4, E-Q39.6).
- *Policy knowledge index*: standardized index of:
  - *Manager of CSDF*: change in whether respondents switch towards correctly identifying Representatives as the primary controller of CSDF (B-Q9, E-Q9).
  - *CSDF reporting requirement*: change in whether respondents correctly switch towards believing that it is a legal obligation to report CSDF allocation (B-Q10, E-Q10).
  - *CSDF citizen involvement requirement*: change in whether respondents correctly switch towards believing that it is a legal obligation to involve citizens in CSDF allocation decisions (B-Q11, E-Q11).

Table VII:

- *Political information demand index*: standardized index of:
  - *Change in radio listening*: Respondents’ change in listening to radio between baseline (B-Q6, E-Q6).
  - *Demand for non-radio information sources*: How frequently respondents sought political information from non-radio sources such as newspapers, television and the internet (E-Q7).
  - *Change in political discussion with friends*: Change in how frequently respondents discussed political issues with friends, family, neighbors and other members of the community (B-Q8, E-Q8).
- *Debate coordination index*: standardized index of:
  - *Discussed debate with friends*: How frequently respondents discussed the debate content with others (E-Q19).

- *Discussion led to coordination*: Whether this discussion led respondents, along with others, to agree on one particular candidate to vote for (E-Q20).
- *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.

Table VIII:

- *Certainty about competence*: standardized change in how sure respondents were about the competence of specific predicted leading 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 leading candidates between baseline (B-Q16, B-Q18, B-Q20) and endline (E-Q28, E-Q30, E-Q32).
- *Beliefs about competence*: standardized change in how competent respondents believe specific predicted leading 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 IX:

- *Ground*: standardized index of:
  - Whether candidate distributed leaflets or posters in respondent's community (E-Q41.1, E-Q41.2, E-Q41.3).
  - Whether candidate made campaign visits to respondent's community (E-Q41.1, E-Q41.2, E-Q41.3).
  - 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 frequently respondents heard candidates on the radio in the two weeks before the election (E-Q39.1, E-Q39.3, E-Q39.5).

Table X:

- Panel 1:
  - *Main effect*: indicator for whether a respondent named a specific predicted leading candidate as their vote choice at endline (E-Q45).
  - *Interaction: Performance*: measure of predicted debate performance of a specific predicted leading candidate, generated by flexibly estimating observed measure of debate performance and predicting out-of-sample on candidates who did not participate.
  - *Interaction: Priority match*: measure of preference alignment between respondent and a specific predicted leading 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).
  
- Panel 2:
  - *Main effect*: Vote share of candidate at polling station-level.
  - *Interaction: Performance*: District-level analogue of respondent-level interaction.
  - *Interaction: Priority match*: District-level analogue of respondent-level interaction.

## A1. Descriptive Statistics

	Mean	SD	Min	Max
<b>A. District-level variables (<math>n = 73</math>)</b>				
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>B. Individual-level variables (<math>n = 4060</math>)</b>				
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
<b>C. Polling station-level variables (<math>n = 5386</math>)</b>				
Number of registered voters in PS (2017)	405.12	74.87	12.00	505.00
Number of PS in VRC	3.57	1.89	1.00	9.00
VRC added in 2017	0.10	0.31	0.00	1.00
Number of registered voters in VRC (2011)	1422.96	770.45	45.00	3995.00
Number of PS in VRC (2011)	3.36	1.59	1.00	9.00
Turnout (2011)	0.63	0.09	0.14	1.01
Share of invalid votes (2011)	0.07	0.03	0.00	0.35
PS covered by partner radio station	0.85	0.36	0.00	1.00
1/N radio stations covering PS	0.05	0.03	0.00	0.11
Urban PS	0.33	0.47	0.00	1.00

Table presents descriptive statistics relating to all control variables used in the regression analyses. *Sources:* **District-level variables:** Debate variables from Internews. All 2017 and 2011 variables come from National Elections Commission (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. **Polling station-level variables:** Radio station variables come from Internews. All other variables come from NEC. 90% of polling stations were in locations where a polling place (Voter Registration Center, VRC) existed in 2011. For new polling stations we assign district-level averages.

A2. Descriptive Statistics (Respondent level outcomes)

	Mean	SD	Min	Max
<b>Table V</b>				
Share of candidates attended debate	0.59	0.22	0.11	1.00
Incumbent attended debate	0.35	0.48	0.00	1.00
Share of challengers attended debate	0.60	0.37	0.00	1.00
<b>Table VI</b>				
Heard debate between baseline and endline	0.09	0.29	0.00	1.00
Heard debate at endline	0.21	0.41	0.00	1.00
Number of times heard debate	0.46	1.06	0.00	24.00
Debate winner attended debate	0.32	0.47	0.00	1.00
Stated share of participating candidates	0.12	0.29	0.00	2.53
Stated incumbent debate participation	0.16	0.34	0.00	1.00
Switches to correct CSDF controller	0.25	0.43	0.00	1.00
Switches to correct requirement for CSDF reporting	0.25	0.43	0.00	1.00
Switches to correct requirement for CSDF citizen engagement	0.25	0.43	0.00	1.00
<b>Table VII</b>				
Change in radio listening	0.26	2.17	-6.00	6.00
Demand for non-radio information sources	5.01	2.15	0.00	7.00
Change in political discussion with friends	-0.07	2.30	-6.00	6.00
Discussed debate with friends	1.01	1.61	0.00	5.00
Discussion led to coordination	1.06	1.76	0.00	5.00

Table presents descriptive statistics relating to all unstandardized, disaggregated outcomes variables used in the respondent-level regressions as outcome variables.

### A3. Descriptive Statistics (Respondent-candidate level outcomes)

	Incumbent				Challengers			
	Mean	SD	Min	Max	Mean	SD	Min	Max
<b>Table VIII</b>								
Change in certainty of candidate competence	0.08	1.40	-4.00	4.00	0.36	1.53	-4.00	4.00
Change in certainty of candidate priority issues	-0.03	1.63	-4.00	4.00	-0.02	1.58	-4.00	4.00
Change in assessment of candidate competence	0.28	1.44	-4.00	4.00	0.12	1.23	-4.00	4.00
Change in share of candidate priorities named	0.02	0.41	-1.00	1.00	0.03	0.42	-1.00	1.00
<b>Table IX</b>								
Frequency of hearing candidate on radio	2.95	1.14	0.00	5.00	2.67	1.13	0.00	5.00
Candidate distributed leaflets in locality	0.93	0.25	0.00	1.00	0.90	0.30	0.00	1.00
Candidate visited locality	0.81	0.39	0.00	1.00	0.74	0.44	0.00	1.00
Frequency of candidate vote buying	2.28	1.39	1.00	5.00	2.10	1.28	1.00	5.00
<b>Table X (Panel 1)</b>								
Votes for candidate	0.31	0.46	0.00	1.00	0.15	0.35	0.00	1.00
Measure of predicted debate performance	0.27	0.18	0.00	0.70	0.12	0.15	0.00	0.83
Measure of policy match	0.42	0.27	0.00	1.00	0.42	0.28	0.00	1.00
<b>Table X (Panel 2)</b>								
Vote share	0.26	0.22	0.00	0.99	0.11	0.16	0.00	0.99
Measure of predicted debate performance	0.25	0.19	0.00	0.68	0.12	0.13	0.00	0.75
Measure of policy match	0.43	0.09	0.16	0.59	0.43	0.09	0.00	0.64

Table presents descriptive statistics relating to all unstandardized, disaggregated outcomes variables used in the respondent-candidate level regressions as outcome variables.

#### A4. District-level Balance

	<b>VS of 1st place (2011)</b>		<b>VS of 1st place (2011)</b>		<b>VS of 3rd place (2011)</b>	
<i>Invite</i>	-0.011 (0.024)	-0.015 (0.025)	-0.005 (0.010)	-0.008 (0.010)	-0.000 (0.007)	-0.001 (0.006)
Control Mean	0.322	0.309	0.178	0.183	0.129	0.129
	<b>Turnout (2011)</b>		<b>Share ran in 2011</b>		<b>VS HHI (2011)</b>	
<i>Invite</i>	-0.004 (0.009)	-0.004 (0.008)	-0.012 (0.024)	-0.019 (0.024)	-0.006 (0.020)	-0.009 (0.020)
Control Mean	0.661	0.667	0.258	0.266	0.198	0.189
	<b>N. candidates (2017)</b>		<b>Incumbent ran (2017)</b>		<b>Number of debates</b>	
<i>Invite</i>	-0.509 (0.833)	-0.670 (0.928)	0.037 (0.079)	0.016 (0.076)	-0.083 (0.126)	-0.131 (0.140)
Control Mean	13.634	14.780	0.850	0.883	2.115	2.256
	<b>Log pop. dens. (2008)</b>		<b>Log reg. voters (2017)</b>		<b>GSM coverage (2016)</b>	
<i>Invite</i>	0.254 (0.365)	0.160 (0.370)	0.056 (0.061)	0.060 (0.055)	0.032 (0.062)	0.005 (0.055)
Control Mean	-9.847	-9.108	10.214	10.351	0.645	0.773
	<b>Share owns radio (2015)</b>		<b>Share radio news (2015)</b>		<b>Radio coverage (2016)</b>	
<i>Invite</i>	-0.019 (0.025)	-0.034 (0.022)	-0.017 (0.026)	-0.038 (0.025)	-0.109 (1.347)	0.024 (1.384)
Control Mean	0.755	0.773	0.767	0.780	10.051	12.830
	<b>Debate week</b>		<b>Share sec. ed. (2008)</b>		<b>Share 18+ (2008)</b>	
<i>Invite</i>	0.028 (0.215)	-0.030 (0.209)	0.004 (0.010)	0.003 (0.010)	0.000 (0.004)	0.001 (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 A1.

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.

### A5. Individual-level Balance (district covariates)

	VS of 1st place (2011)			VS of 1st place (2011)			VS of 3rd place (2011)		
<i>Invite</i>	0.007 (0.025)	-0.009 (0.021)	-0.014 (0.021)	-0.011 (0.009)	-0.005 (0.008)	-0.008 (0.009)	-0.000 (0.007)	-0.000 (0.006)	-0.001 (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)		
<i>Invite</i>	0.001 (0.007)	-0.004 (0.007)	-0.004 (0.007)	-0.043** (0.019)	-0.012 (0.021)	-0.019 (0.020)	0.007 (0.021)	-0.005 (0.017)	-0.007 (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		
<i>Invite</i>	-0.642 (0.757)	-0.549 (0.709)	-0.710 (0.795)	0.010 (0.068)	0.034 (0.068)	0.014 (0.066)	-0.117 (0.108)	-0.091 (0.108)	-0.139 (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)		
<i>Invite</i>	0.207 (0.317)	0.246 (0.312)	0.149 (0.317)	0.072 (0.052)	0.055 (0.052)	0.059 (0.047)	0.029 (0.050)	0.032 (0.053)	0.004 (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)		
<i>Invite</i>	-0.041** (0.020)	-0.020 (0.022)	-0.035* (0.019)	-0.035 (0.025)	-0.017 (0.022)	-0.039* (0.022)	-0.172 (1.222)	-0.138 (1.159)	-0.010 (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)		
<i>Invite</i>	-0.035 (0.193)	0.031 (0.184)	-0.026 (0.179)	0.006 (0.008)	0.004 (0.009)	0.002 (0.009)	-0.001 (0.004)	0.000 (0.004)	0.001 (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 A1.

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.

### A6. Individual-level balance (individual covariates)

	Survey date			Education		
<i>Invite</i>	-1.063** (0.523)	-0.796 (0.572)	-1.006* (0.573)	0.067 (0.086)	0.102 (0.084)	0.081 (0.075)
Control Mean	71.801	71.153	71.422	6.586	6.447	6.534
	Age			Male		
<i>Invite</i>	0.163 (0.439)	0.359 (0.464)	0.225 (0.437)	0.011 (0.015)	0.025 (0.016)	0.021 (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.

A7. Polling Station-level Balance (district covariates)

	<b>VS of 1st place (2011)</b>			<b>VS of 1st place (2011)</b>			<b>VS of 3rd place (2011)</b>		
<i>Invite</i>	-0.015 (0.021)	-0.010 (0.021)	-0.015 (0.021)	-0.009 (0.009)	-0.005 (0.008)	-0.009 (0.009)	-0.001 (0.005)	-0.000 (0.006)	-0.001 (0.005)
Control Mean	0.308	0.309	0.309	0.183	0.184	0.184	0.129	0.129	0.129
	<b>Turnout (2011)</b>			<b>Share ran in 2011</b>			<b>VS HHI (2011)</b>		
<i>Invite</i>	-0.004 (0.007)	-0.004 (0.007)	-0.004 (0.007)	-0.016 (0.020)	-0.012 (0.021)	-0.018 (0.020)	-0.008 (0.017)	-0.006 (0.017)	-0.008 (0.018)
Control Mean	0.665	0.667	0.667	0.266	0.266	0.266	0.188	0.188	0.188
	<b>N. candidates (2017)</b>			<b>Incumbent ran (2017)</b>			<b>Number of debates</b>		
<i>Invite</i>	-0.648 (0.771)	-0.506 (0.715)	-0.677 (0.798)	-0.050 (0.067)	-0.019 (0.070)	-0.053 (0.066)	-0.124 (0.118)	-0.083 (0.109)	-0.130 (0.121)
Control Mean	14.651	14.794	14.794	0.883	0.890	0.890	2.239	2.256	2.256
	<b>Log pop. dens. (2008)</b>			<b>Log reg. voters (2017)</b>			<b>GSM coverage (2016)</b>		
<i>Invite</i>	0.178 (0.322)	0.255 (0.314)	0.154 (0.320)	0.062 (0.048)	0.056 (0.052)	0.060 (0.047)	0.010 (0.048)	0.032 (0.053)	0.005 (0.047)
Control Mean	-9.223	-9.097	-9.097	10.325	10.352	10.352	0.753	0.775	0.775
	<b>Share owns radio (2015)</b>			<b>Share radio news (2015)</b>			<b>Radio coverage (2016)</b>		
<i>Invite</i>	-0.031 (0.019)	-0.019 (0.022)	-0.033* (0.019)	-0.037* (0.021)	-0.017 (0.022)	-0.038* (0.022)	-0.001 (1.201)	-0.100 (1.157)	-0.065 (1.199)
Control Mean	0.770	0.773	0.773	0.778	0.780	0.780	12.485	12.919	12.919
	<b>Debate week</b>			<b>Share sec. ed. (2008)</b>			<b>Share 18+ (2008)</b>		
<i>Invite</i>	-0.041 (0.176)	0.029 (0.185)	-0.046 (0.178)	0.003 (0.009)	0.004 (0.009)	0.002 (0.009)	0.002 (0.003)	0.000 (0.004)	0.001 (0.003)
Control Mean	4.007	3.987	3.987	0.152	0.154	0.154	0.486	0.487	0.487
Observations	5383	5383	5383	5383	5383	5383	5383	5383	5383
Weight	No	1/PS	Reg	No	1/PS	Reg	No	1/PS	Reg

Descriptions of all variables can be found in Table A1.

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.

### A8. Polling Station-level Balance (PS covariates)

	N. PS in VRC (2017)			N. PS in VRC (2011)		
<i>Invite</i>	-0.099 (0.247)	0.008 (0.226)	-0.121 (0.249)	-0.114 (0.185)	-0.011 (0.183)	-0.142 (0.186)
Control Mean	3.605	3.807	3.807	3.399	3.571	3.571
	New VRC			Urban		
<i>Invite</i>	-0.014 (0.014)	-0.008 (0.016)	-0.015 (0.013)	0.067 (0.058)	0.079 (0.056)	0.063 (0.059)
Control Mean	0.111	0.101	0.101	0.299	0.326	0.326
	Reg. Voters (2017)			Reg. Voters (2011)		
<i>Invite</i>	2.247 (6.170)	4.668 (6.416)	0.686 (4.775)	-53.204 (90.770)	0.239 (89.758)	-67.706 (90.943)
Control Mean	403.792	418.480	418.480	1439.977	1526.086	1526.086
	Turnout (2011)			Invalid votes (2011)		
<i>Invite</i>	0.014 (0.009)	0.011 (0.009)	0.015* (0.009)	0.002 (0.003)	0.004 (0.003)	0.002 (0.003)
Control Mean	0.620	0.620	0.620	0.066	0.066	0.066
	Radio covered			Radio intensity		
<i>Invite</i>	-0.016 (0.054)	0.012 (0.052)	-0.025 (0.054)	0.001 (0.004)	0.003 (0.004)	-0.000 (0.004)
Control Mean	0.848	0.858	0.858	0.046	0.045	0.045
Observations	5383	5383	5383	5383	5383	5383
Weight	No	1/PS	Reg	No	1/PS	Reg

Descriptions of all variables can be found in Table A1. Analysis throughout is at the polling station-level, where multiple polling stations exist within a single location called a VRC. 1780/2080 VRCs existed in the 2011 election; for these 300 new VRCs we assign district-level averages.

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.

### A9. Incumbent Balance

	Attendance		Absent		Distant	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Invite</i>	0.034 (0.032)	0.019 (0.030)	-0.003 (0.020)	0.013 (0.024)	-0.025 (0.028)	-0.023 (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.

A10. Candidate-level Balance

	Response (1)	Age (2)	Univ. ed. (3)	Radio (4)	Male (5)	Ran before (6)	Gov job (7)	Advocacy (8)	NGO (9)
<b>A. All candidates</b>									
<i>Invite</i>	-0.028 (0.031)	-0.379 (1.064)	0.012 (0.049)	0.020 (0.013)	0.009 (0.027)	-0.024 (0.032)	-0.028 (0.030)	0.064** (0.025)	0.065* (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>B. Incumbents</b>									
<i>Invite</i>	0.139 (0.131)	2.394 (4.085)	0.061 (0.210)	-0.030 (0.200)	0.121 (0.174)	0.000 (.)	-0.394* (0.212)	-0.121 (0.174)	0.061 (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
<b>C. Challengers</b>									
<i>Invite</i>	0.010 (0.083)	3.875** (1.655)	-0.005 (0.106)	-0.002 (0.052)	0.027 (0.067)	0.134 (0.102)	0.036 (0.086)	0.155** (0.065)	0.175* (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 unweighted balance tests for the full set of candidates in survey, Panel B restricts to incumbent candidates, Panel C 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.

A11. 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)
<b>A. All candidates</b>									
<i>Invite</i>	-0.045 (0.033)	-0.171 (0.899)	-0.013 (0.056)	0.030* (0.017)	0.025 (0.031)	-0.031 (0.036)	-0.018 (0.039)	0.043 (0.027)	0.057 (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>B. Challengers</b>									
<i>Invite</i>	-0.037 (0.064)	4.918*** (1.815)	0.042 (0.105)	0.008 (0.064)	0.008 (0.068)	0.153 (0.106)	0.054 (0.094)	0.144** (0.068)	0.165 (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

In this table we weight observations by the inverse of the number of responding candidate types in a given district. Since there is one incumbent per district, for incumbents this would be identical to Panel B of Table A10 and thus we omit it. Panel A presents balance tests for the full set of candidates in survey, Panel B 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.

A12. Candidate Debate Participation (supplementary)

	(1)	(2)	(3)
<b>A. Election winner</b>			
<i>Invite</i>	0.253** (0.097)	0.202** (0.089)	0.275*** (0.093)
Control Mean	0.501	0.520	0.474
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
<b>B. Share of actual challengers</b>			
<i>Invite</i>	0.267*** (0.068)	0.176** (0.068)	0.237*** (0.068)
Control Mean	0.488	0.572	0.525
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
<b>C. Share of actual other candidates</b>			
<i>Invite</i>	0.018 (0.039)	0.029 (0.037)	0.036 (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 (winner, actual challenger) who attended a debate out of all candidates in that district. Actual challengers are defined as candidates who ranked in the top three in their race in the election but were not the incumbent. Actual other candidates are those who did not rank in the top three. Panels A and B have 4060 observations; Panel C 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.

A13. Debate Rebroadcasting

	(1)	(2)	(3)
<b>A. Radio monitors</b>			
<i>Invite</i>	0.387 (0.998)	-0.700 (0.860)	-0.206 (0.974)
Control Mean	5.230	5.618	5.466
<b>B. Radio survey</b>			
<i>Invite</i>	1.055 (1.044)	0.326 (0.906)	0.746 (0.954)
Control Mean	7.473	7.702	7.698
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs

Outcome variables are, in Panel A, the number of contracted rebroadcasts confirmed by radio monitors and, in panel B, Number of rebroadcasts based on survey of radio stations, including those not contracted to rebroadcast but recorded as being present in the debate venue.

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.

A14. Debate Exposure (supplementary)

	(1)	(2)	(3)
<b>A. Debate listening index</b>			
<b>(1) Change in heard debate</b>			
<i>Invite</i>	0.080*** (0.029)	0.080*** (0.026)	0.102*** (0.027)
Control Mean	0.084	0.082	0.082
<b>(2) Heard debate</b>			
<i>Invite</i>	0.038* (0.022)	0.035* (0.020)	0.050** (0.021)
Control Mean	0.195	0.202	0.193
<b>(3) Number of times heard</b>			
<i>Invite</i>	0.085* (0.045)	0.104** (0.046)	0.120*** (0.045)
Control Mean	0.420	0.440	0.420
<b>B. Debate knowledge index</b>			
<b>(1) Debate winner attended debate</b>			
<i>Invite</i>	0.075** (0.031)	0.078*** (0.029)	0.096*** (0.029)
Control Mean	0.291	0.297	0.283
<b>(2) Stated share of participating candidates</b>			
<i>Invite</i>	0.023 (0.015)	0.024 (0.015)	0.031** (0.014)
Control Mean	0.111	0.115	0.111
<b>(3) Stated share of participating leading candidates</b>			
<i>Invite</i>	0.030* (0.017)	0.026* (0.016)	0.039** (0.016)
Control Mean	0.145	0.153	0.148
<b>C. Policy knowledge index</b>			
<b>(1) Manager of CSDF</b>			
<i>Invite</i>	0.041 (0.029)	0.053 (0.038)	0.044 (0.032)
Control Mean	0.237	0.231	0.243
<b>(2) CSDF reporting requirement</b>			
<i>Invite</i>	-0.011 (0.032)	0.018 (0.035)	-0.008 (0.033)
Control Mean	0.247	0.249	0.247
<b>(3) CSDF citizen involvement requirement</b>			
<i>Invite</i>	0.094** (0.040)	0.113** (0.051)	0.114** (0.045)
Control Mean	0.246	0.246	0.244
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs

All outcome variables are described in Section A.4.

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.

A15. Political Engagement (supplementary)

	(1)	(2)	(3)
<b>A. Political information demand index</b>			
<b>(1) Change in radio listening</b>			
<i>Invite</i>	0.285** (0.134)	0.451*** (0.164)	0.370** (0.160)
Control Mean	0.233	0.213	0.232
<b>(2) Demand for non-radio information sources</b>			
<i>Invite</i>	0.143* (0.075)	0.168* (0.085)	0.179** (0.083)
Control Mean	4.970	4.800	4.932
<b>(3) Change in political discussion with friends</b>			
<i>Invite</i>	0.510*** (0.189)	0.516** (0.226)	0.491** (0.193)
Control Mean	-0.124	-0.106	-0.117
<b>B. Debate coordination index</b>			
<b>(1) Discussed debate with friends</b>			
<i>Invite</i>	0.207** (0.102)	0.222** (0.095)	0.275*** (0.095)
Control Mean	0.936	0.953	0.908
<b>(2) Discussion led to coordination</b>			
<i>Invite</i>	0.282** (0.107)	0.285*** (0.101)	0.343*** (0.103)
Control Mean	0.956	0.974	0.928
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs

All outcome variables are described in Section A.4.

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.

A16. Candidate Campaigning (supplementary)

	(1)	(2)	(3)
<b>A. Incumbent</b>			
<b>(1) Candidate visited community</b>			
<i>Invite</i>	-0.027 (0.019)	-0.038 (0.023)	-0.032 (0.023)
Control Mean	0.823	0.824	0.820
<b>(2) Candidate distributed leaflets</b>			
<i>Invite</i>	-0.012 (0.010)	-0.014 (0.012)	-0.012 (0.011)
Control Mean	0.935	0.933	0.932
<b>(3) Vote buying for candidate</b>			
<i>Invite</i>	0.035 (0.039)	0.059 (0.039)	0.070* (0.041)
Control Mean	2.258	2.223	2.231
Observations	3492	3492	3492
<b>B. Challenger</b>			
<b>(1) Candidate visited community</b>			
<i>Invite</i>	-0.024** (0.012)	-0.019* (0.010)	-0.024* (0.012)
Control Mean	0.756	0.753	0.753
<b>(2) Candidate distributed leaflets</b>			
<i>Invite</i>	-0.013** (0.006)	-0.016** (0.007)	-0.017** (0.007)
Control Mean	0.905	0.908	0.907
<b>(3) Vote buying for candidate</b>			
<i>Invite</i>	-0.008 (0.038)	-0.025 (0.039)	-0.021 (0.037)
Control Mean	2.109	2.145	2.132
Observations	8676	8676	8676
Weight	No	1/Obs	Reg/Obs

All outcome variables are described in Section A.4.

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.

A17. Vote Choice Reason

	Campaign promises			Expectations		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Invite</i>	0.006 (0.010)	0.002 (0.012)	0.000 (0.012)	0.018 (0.019)	0.005 (0.021)	0.024 (0.018)
Control Mean	0.033	0.036	0.035	0.067	0.064	0.063
	Competence			Experience		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Invite</i>	0.013 (0.030)	0.016 (0.032)	0.001 (0.033)	-0.024 (0.026)	-0.001 (0.030)	-0.008 (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, expectedated 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.

A18. Broader Consequences

	(1)	(2)	(3)
<b>A. Media: trust and bias</b>			
<i>Invite</i>	-0.007 (0.013)	-0.003 (0.013)	-0.004 (0.013)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
<b>B. Media: helps democracy</b>			
<i>Invite</i>	-0.003 (0.020)	-0.004 (0.026)	-0.015 (0.025)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
<b>C. Electoral attitudes</b>			
<i>Invite</i>	0.010 (0.027)	0.013 (0.031)	0.004 (0.029)
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.

A19. Rebroadcasting Intervention

	(1)	(2)	(3)
<b>A. Full sample</b>			
<i>Rebroadcast</i>	0.149 (0.093)	0.113 (0.088)	0.145 (0.102)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
<b>B. Respondents in intensive-invite districts</b>			
<i>Rebroadcast</i>	-0.021 (0.156)	-0.021 (0.128)	0.021 (0.156)
Observations	2252	2252	2252
Weight	No	1/Obs	Reg/Obs

Outcome variable is our standardized index of debate exposure. In Panel A, we show no overall effects on debate exposure. In Panel B, we show no effects on debate exposure if we restrict to respondents in those districts assigned to high invitation intensity.

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.